Vol. 26, No. 3, 2022 | DOI: 10.26794/2587-5671 | ISSN 2587-5671 (Print) ISSN 2587-7089 (Online)

FINANCE: THEORY AND PRACTICE

Scientific and practical journal. Former title: Bulletin of the Financial University. Published since 1997

The edition is reregistered in the Federal Service for communication, informational technologies and media control: PI No. FS77-70021 of May 31, 2017

Publication frequency — 6 issues per year

Founder: Financial University

The aim of the journal is the scientific discussion of topical issues in the field of finance and related fields

The journal is Scopus indexed, in the Russian Science Citation Index (RSCI), in the core of the RSCI, in the Web of Science — Russian Science Citation Index (RSCI), is included in the listing of periodicals recommended by the Higher Attestation Commission for the publication of the main results of the postgraduate and doctoral dissertations.

All articles of journal Finance: Theory and Practice are published with a digital object identifier (DOI).

The journal is distributed only by subscription Subscription index 82140 in the consolidated catalogue "The Press of Russia". The journal is publicly available (Open Access) on the website http://financetp.fa.ru/jour/index

Manuscripts are to be submitted to the editorial office in electronic form: **vestnikfinu@mail.ru**

Minimal size of the manuscript: 4 ths words; optimal — 6 ths words.

The editorial makes a mandatory expertise (review, scientific and stylistic editing) of all the materials to be published in the journal.
More information on publishing terms is at: **financetp.fa.ru**

CONTENTS

FINANCE: THEORY AND PRACTICE

Scientific and practical journal

Vol. 26, No. 3, 2022

Editor-in-Chief –
M.A. Fedotova
Head of Scientific Journals Editorial Department –
V.A. Shadrin
Managing editor –
I.S. Dovgal
Translator –
E.V. Logutova
Reference Manager –
V.M. Alekseev
Proofreader –
S.F. Mihaylova
Design, make up –
S.M. Vetrov

Editorial address:
53, Leningradsky prospekt, office 5.4
Moscow, 125993
tel.: **+7 (499) 553-10-71 (internal 10-79)**
E-mail: **vestnikfinu@mail.ru**
Site: **financetp.fa.ru**

Subscription in editorial office
tel: **+7 (499) 553-10-71 (internal 10-80)**
e-mail: **sfmihajlova@fa.ru**
S.F. Mihaylova

Signed for press on 29.06.2022
Format 60 x 84 1/8.
Size 33,75 printer sheets.
Order № 626.
Printed by Publishing House of the Financial University (51, Leningradsky prospekt)

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-6-18
UDC 330.88(045)
JEL B26, E61, H50

Public Goals and Government Expenditures: Are the Solutions of the "Modern Monetary Theory" Realistic?

I.M. Stepnov[a], **J.A. Kovalchuk**[b], **M.V. Melnik**[c], **T. Petrovic**[d]
[a, b] MGIMO University, Moscow, Russia; [a, c] Financial University, Moscow, Russia;
[b] Moscow Aviation Institute, Moscow, Russia; [d] Innovation Project Management Center — IPM Center Ltd, Pancevo, Serbia

ABSTRACT

The **relevance** is determined by the presence of discussions among economists and political scientists about the state regulation role, especially during crises when governments are expected to provide an economic miracle solution or new measures that are not thought about during periods of growth. The **subject** of the study is government expenditures, which should be fairly distributed by governments in crisis times with a clear vision and understanding of public goals, for the achievement of which it is necessary to undertake economic changes. Therefore, the **purpose** of the study is to generalize the results of the discussion about the possibility of changing the economic "mainstream" to the "modern monetary theory" (MMT) and its prospects in public finance management. The research **methods** include systematic, critical, comparative and cross-country analysis and verification of the realism of the MMT's proposals. The **results** of the study made it possible to analyze the economic solutions of the new theory based on the political subtext of the theory itself: is it possible to entrust governments with an unlimited budget filled with the issue of national currency; has there been a need to update public goals; what are the prospects for full employment with government funding of jobs; is it possible for society to subject taxes to increase or is there enough transparency and collection; is inflation dangerous in the current situation and in the future; is the MMT really an economic recipe for politicians focused on achieving public goals, including ecology and inequality. It is **concluded** that all-round criticism and all-round approval are equally unacceptable in relation to the new theory from the standpoint of practical state regulation. Some elements of the MMT have already been implemented in various countries, both successfully and with negative consequences. Therefore, when considering the MMT in unity with public goals, the criterion of choice should not be the quality of economic decisions, not their elegance and validity, but enforceability in accordance with urgent and strategic tasks that need to be solved by politicians.

Keywords: modern monetary theory; government expenditures; public goals; values; economic policy; fiscal policy; money issue; employment; crisis; inflation

For citation: **Stepnov I.M., Kovalchuk J.A., Melnik M.V., Petrovic T. Public goals and government expenditures: Are the solutions of the "modern monetary theory" realistic?** ***Finance: Theory and Practice.*** **2022;26(3):6-18. DOI: 10.26794/2587-5671-2022-26-3-6-18**

INTRODUCTION

Crises and renewal of economic theories

A new wave of active discussion of the "Modern Monetary Theory" (hereinafter referred to as "MMT") coincided with the growth of economic and social challenges facing the governments of almost all countries. The crises provoke mass discussions of new economic theories and influence policymakers' decisions to change the economic strategy of society.

Modern discussion has been caused by four books that give the most complete understanding of MMT, and in many respects, when disclosing its provisions, we turn to them:

- Wray L. R. Modern monetary theory: A primer on macroeconomics for sovereign money systems [1];
- Wray L. R., et al. Public service employment: A path to full employment [2];
- Mitchell W., et al. Macroeconomics [3];
- Kelton S. The deficit myth: Modern monetary theory and the birth of the people's economy [4].

The terms "monetarism" or "monetary policy" have long been accepted in Russian-language articles and documents, but in the domestic scientific discussion MMT is defined as "modern monetary theory" (for example, [5]), and not "modern monetarism theory". Therefore, in order to avoid ambiguity, we use the wording "modern monetary theory" adopted in Russian publications.

The concept of MMT after the crisis of 2007–2009 has not received enough discussion over the past ten years, but has proved to be in demand in two current situations: in the political struggleы in the United States (including the ideology of the Green New Deal) and in the global fight against the pandemic, more precisely, with the socio-economic consequences of the pandemic. During the pandemic crisis the governments of many countries (regardless of the state of their economies) made decisions on financial assistance to support households and businesses from the state budget (*Fig. 1*), which was not even discussed during the financial crisis. In general, the situation with the pandemic has exacerbated the discussion on the need to improve public administration, postponing for a short time the issues of decarbonization of the economy. According to S. A. Afontsev, "the problem of state regulation, unlike the problems of individual choice or the investment behavior of a company, is not purely economic, but economic and political in nature ... this is not surprising, that the mainstream economic models turn out to be inadequate in this case: they completely ignore the political side of the problem, which is why the study built on their basis inevitably turns out to be one-sided" [6, p. 74]. The same happened (or rather, is happening) with the "modern monetary theory" — a new solution for monetarists.

MMT is a heterodox theory and therefore is attractive [7]. According to MMT a government does not need to worry about budget constraints and excessive inflation until society achieves its goals of full employment, first-class healthcare, and other public goals.

This raises two questions (more precisely, two doubts):

1) if MMT is good, then why hasn't it been adopted by policymakers yet?

2) if MMT is bad and its shortcomings are obvious, then why have leading economists joined the discussion (for example, G. Mankiw, P. Krugman)?

After all, MMT is to define both political goals (what the government spends its resources on), such as employment, and means for those goals (how the government finances its expenditures), such as money printing. Without trying to give an unambiguous answer to the questions posed, we have defined the aim of this paper as finding out whether a "modern monetary theory", which is no longer so modern, can benefit society since many of its provisions have been known for decades [5].

Undoubtedly, the coronavirus pandemic fully revealed the lack of public administration tools in the framework of

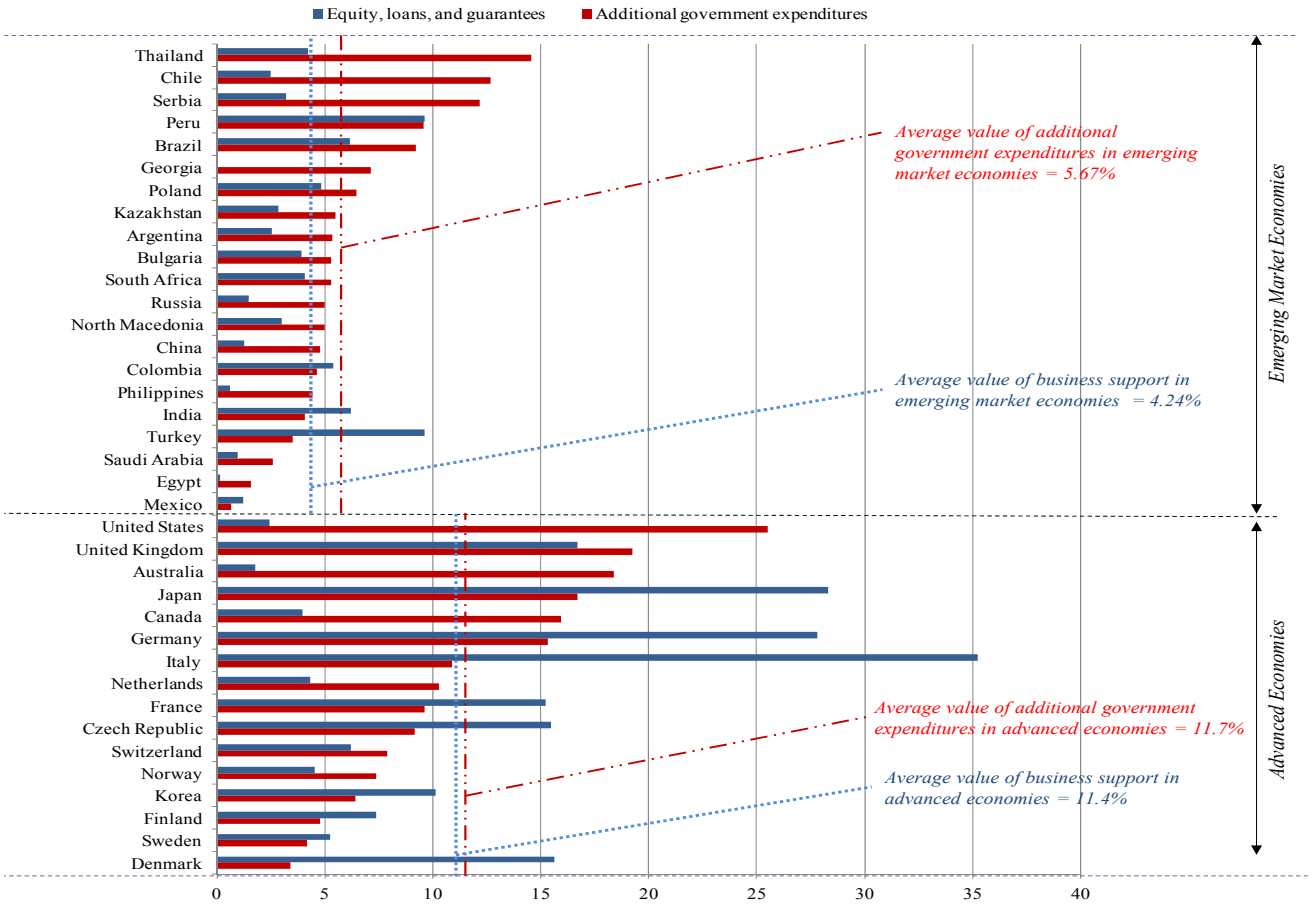


Fig. 1. **Additional government expenditures January 2020 — October 2021, % of the country's GDP**

Source: compiled by the authors based on the data: IMF Fiscal Affairs Department. Fiscal Monitor Database of Country Fiscal Measures in Response to the COVID-19 Pandemic. October 2021. URL: https://www.imf.org/en/Topics/imf-and-covid19/Fiscal-Policies-Database-in-Response-to-COVID-19.

the choice between "life" and "livelihoods" [8] and required a compromise between conflicting public values, and different approaches in different countries.

The economic impact of the (still ongoing) pandemic has shown that price signals are no longer the main determinants of supply and demand, limiting the effectiveness of traditional theories' recommendations. As a result, the economy entered a state far from equilibrium [9]. Policymakers have been given a choice: to focus on initial responses that are to help mitigate current challenges or to focus on accelerating post-pandemic economic recovery.

At this point in time, MMT is claimed to have in store drastic measures to respond to current economic challenges such as public debt and unemployment. Therefore, in the ongoing discussion (M. Adil, A. Baker, R. Murphy, L. Baronian, G. Brady, G. Epstein, Y. Huang, G. Menkiw, G. Tavlas and others [7, 9–19]) active attempts have been made to assess the MMT provisions. Obviously, the criterion for such an assessment should not be the uniqueness of economic models, but their political feasibility. In other words, the question is whether MMT can help to get out of the crisis with minimal losses and not face rising inflation and unemployment in the future, not burden the next generations with the need to fulfill debt obligations — i.e. realize public goals or disavow them?

The foundations of MMT include five key provisions [13].

First, sovereign governments can issue as much currency as they want, and have

unlimited ability to fulfill promised future payments.

Second, unlimited public spending is allowed until the goals are met.

Third, the budget deficit, which precedes tax payments.

Fourth, a growing national debt is not dangerous for the economy as long as inflation is contained.

A fifth feature, the Internet shaped the spread of MMT and its influence on public opinion in the absence of the traditional channels of economic policy making.

It is worth noting that the recently published textbook "Macroeconomics" [3] does not contain sections on the theory of economic growth and the analysis of aggregate demand and aggregate supply. Therefore, we agree with C. Mann [20] that MMT presents both correct and simplified and incorrect conclusions, for example, in relation to the financial sector and financial intermediation.

Thus, in practical terms, it is not the economic decisions that served as the basis of the new theory that are of interest (more precisely, a new combination of well-known provisions in a holistic concept, although it has significant gaps), but political overtones: the new theory capable of ensuring the full implementation of public goals, or its effectiveness is not as high as its proponents try to prove. Is this true?

RESEARCH RESULTS

Economic decision making in political markets and public goals

The complexity of economic policy is largely determined by the heterogeneity of the participants in social and economic processes in society, according to S.G. Kirdina-Chandler [21], due to the lack of microeconomic ideas about economic entities and their relationships with the inefficiency of macroeconomic analysis. MMT uses this contradiction and tries to build a large economy on microeconomic approaches [10], excluding the need for economic growth, which, in our opinion, is controversial.

For a long time, welfare and economic growth were treated as synonyms, with the latter being seen as the main (if not the only) source of growth in welfare, which aimed policymakers at the pursuit of economic growth, without giving any opportunity to think about the alternative: is it possible to increase welfare without economic growth?

According to V.A. Mau [22], MMT uses its own vision of economic growth, but, like the well-known neoclassical theories of economic development and the theory of macroeconomic equilibrium, it does not provide for the growth of welfare without economic growth. MMT, focusing on left-wing politics, chooses the basis of demand stimulation mechanisms as a source of economic growth, which allows us to conclude that it is opposed to supply-side economics during the period of dominance of liberal economic doctrine [22]. MMT supporters argue that economic growth can be zero without consequences for the social welfare [4]. The latter is an unsubstantiated and contradictory assertion.

The transfer of economic decision-making to political markets is also facilitated by the term "public goal", which is quite often used by proponents of the MMT [1]. The fundamental proposition of MMT is the idea that public spending plans should be developed on the basis of specific socio-economic powers given to governments, and not on the technical measurement of any financial outcome, such as a balanced government budget. This shift from evaluating the performance of governments to measuring the effectiveness of economic measures in society is an important feature of the new theory and is not always justified.

Governments, according to the creators of MMT, are called upon to serve a public goal and to reconcile private interests with a public goal. The public goal is associated with an increase in the standard of living, and L. Wray directly raises the question of how to define a public goal: "What is the public purpose? It is not easy to define or identify. One of the basic functions of any social

organization are to provide the necessary food, clothing, shelter, education, health care, legal framework, and socialization for the survival of society" [1, p. 196].

The use of the concept of a public goal is pragmatic and creates an attraction for a new theory (not always justified) since public purposes can justify or motivate a particular policy and provide at least a benchmark (not always a clear one) against which the effectiveness of public administration can be judged. Therefore, despite the objectivity of a number of economic prerequisites, the political prescriptions of the MMT are not fully justified [16].

Is there a demand to review goals and values in modern society?

Defining public goals depends largely on the socio-cultural context in a particular society and, therefore, may change over time. The existing set of values in society is stable only for a short period of time, therefore, considering the possibilities of MMT, one should understand: is there a demand for changes in society, what are its goals, and why is "modern monetary policy" not completely rejected? One of the answers to the last question is the assumption that MMT was most accepted by those segments of society whose interests were not included in the current political tasks of the society.

The traditionally socially oriented economy, dealing with issues of values, does not cover all of their dynamics, largely leaving the demands of society for the future even at the goal-setting stage. This approach shapes the growth of inequality in society, and from this standpoint, MMT attracts followers with promises to reverse the process of growing inequality. Interestingly, MMT proponents tend to frame the theory in a positive way, using the terms "employment" and "social welfare" instead of overcoming unemployment and inequality, which is confirmed by comparing the macroeconomic performance of countries with different levels of development and income of the population (*Table 1*).

V. A. Mau notes that a new paradigm of socio-economic policy is now being formed, which will dominate in the foreseeable future, since in history "there is no final state, eternal happiness and truths acquired forever" [22, p. 6]. That is why, spreading the conclusions of S. A. Afontsev [23, p. 377] about the current situation, we agree that even in the case of the hypothetical adoption of MMT, "the ongoing transformations should lead not only to an increase in the efficiency of the economic system but also to the formation of a political mechanism, which provides prospects for their continuation and development", which has not yet been confirmed by MMT proponents.

In MMT, examples of the public goal in its current state so far boil down to two messages: the idea of universal job security and the idea of unlimited government expenditures (including the implementation of the "green economy"). Such a vision is still ideal and rather vague but nevertheless exists as a pragmatic guideline by which policymakers understand the attractiveness of the new approach.

When considering political and public goals, it is necessary to take into account the asymmetry of policy effectiveness, when monetary decisions are well received by politicians, but to a lesser extent by society. It is the unpopularity of measures during periods of unemployment and inflation that is another incentive for a positive perception of MMT.

In the context of the pandemic crisis, many governments are trying to form a mixed fiscal and credit policy. Governments seek to simultaneously support workers and provide loans and financial support to corporations. Special programs to accelerate the economic recovery will also require new spending. Assessing the attractiveness of MMT, it should be agreed that if the above decisions were made within the MMT framework, then such approaches would not require a search for sources of financing in the short term. The main discussion was to be about the criteria for the profitability of

Table

Macroeconomic indicators of countries group, 2020

Indicators	World	Advanced economies	Emerging market and developing economies	Low income countries
GDP growth rate,%	–3.267	–4.713	–2.198	1.23
GDP per capita, current prices, USD dollars	–	44190.926	15076.61	2000.70
GDP based on purchasing-power-parity (PPP) share of world total, %	no data available	42.502	57.498	1.29
Unemployment rate,%	5.68	6.621	5.92	5.32
Employment, %	57.24	56.11	63.01	50.42
General government total expenditure, % of GDP	–	47.387	33.977	10.66

Source: compiled by the authors based on the data: IMF. World Economic Outlook Database. URL: https://www.imf.org/en/Publications/WEO/weo-database/2021/April/weo-report?a=1&c=001,110,163,998,200,511,&s=NGDP_RPCH, PPPPC, PPPSH, LUR, LE, GGX_NGDP, D_NGDPD,&sy=2019&ey=2020&ssm=0&scsm=1&scc=0&ssd=1&ssc=0&sic=0&sort=subject&ds=,&br=1; The World Bank. DataBank. World Development Indicator. URL: https://databank.worldbank.org/source/world-development-indicators#

such investments, i.e. future tax revenues [9]. Some policymakers accept MMT provisions precisely because the new theory would allow them to increase the social attractiveness of political decisions without increasing the collection of taxes that were traditionally considered necessary to finance such expenditures, for example, during the premiership of M. Thatcher in the UK.

Only practice can show whether the choice of such policymakers is correct if at least one country completely adopts MMT. But at present, states are increasingly competing not so much due to cheap labor or an abundance of natural resources, but due to the quality of public administration [22]. In our opinion, it is impossible to assess, based only on the provisions of the theory, whether MMT will provide an opportunity to improve the quality of public administration, focusing specifically on public goals.

Therefore, in the case of hypothetical adoption of MMT ideas by policymakers, they should not only set public goals but also seek a compromise between private and public goals, which the new theory does not cover.

Can the state be entrusted with unlimited budget expenditures?

MMT supporters argue that expenditures should not be limited to budgetary goals: "A budget deficit is not an issue as long as the deficit does not lead to inflation" [4, p. 4]. Supporters of financial conservatism often say that there is no "magic money tree" and often argue: "How can we afford it?" [24, p. 33]. MMT suggests that the "magic money tree" is the state, emphasizing its fundamental conclusion that money is a product of the state, not the market.

Within the framework of the theories of monetarism and the theory of rational

expectations presented by M. Friedman, and R.-E. Lucas, the state budget deficit is the main fear leading to the abandonment of many programs to expand jobs, improve healthcare, and make a transition to a green economy, etc. At the same time, we should not forget that sovereigns can default in many ways: including inflating liabilities or depreciating the currency [20], avoiding responsibility for exorbitant government deficits. These trends, according to S. Kelton, limit social development, since any ambitious proposal inevitably leads to discussions about how to find money to pay for it [4].

Until MMT includes in its toolkit a legally approved set of rules for MMT performers, ideal goals will not be able to keep performers endowed with unlimited powers from possible distortions in the practical implementation of the new theory.

Linking unlimited funding and monetary sovereignty, the authors of MMT put forward four conditions for success:

- the country should print its own currency [1, p. 19];
- debt should be measured in its own currency [1, p. 145];
- the currency should be freely exchanged [1, p. 145];
- the exchange rate must be stable without conversion into the currency of another country or gold [1, p. 18–19].

Monetary sovereignty largely corresponds to the global trend of "representations and political imperatives based on the absolute value of sovereignty and the equally absolute predominance of national (country) interests" [25, p. 20].

MMT proponents believe governments should fund spending until the country achieves its goals of full employment, first-class healthcare for all citizens, and other desirable policies. This approach requires government regulations to act as a set of restrictions for policymakers in order to protect the public from abuse of the powers given to the policymakers.

In addition to political issues, the question arises of how the government will manage an unlimited budget. G. Mankiw agrees that the government can always print money to pay its bills, but this fact does not free the government from budget formation, and, consequently, from indicating its sources [18]. P. Krugman believes that MMT does not solve the problem of government debt, and the introduction of the proposed innovations makes debt a potentially more severe problem than all economic agents now admit [26].

The deeper problem is that any policy that allows for unlimited spending without facing the unpopularity of the high taxes needed to fund that spending is abuse. At the same time, a number of economists resolutely deny the possibility of providing the economy with long money and at the same time federal money through additional emissions. Is this justified by doubts that individuals and legal entities will return money to the banking system [27]. Other economists argue that for a monetarily sovereign country, the real default risks lie in corporate debt, private individual debt, and foreign debt, but never national debt denominated in its own currency [17].

MMT proponents also point out that the government cannot continuously print money and directly pump it into the economy without any attention to the level of employment [28], i.e. it is the level of employment is a budget constraint.

Summarizing the above, we note that until MMT includes in its toolkit a legally approved set of rules for MMT performers, ideal goals will not be able to keep performers endowed with unlimited powers from possible distortions in the practical implementation of the new theory.

Are fiscal policy changes acceptable?

MMT does not directly address the choice or quality of fiscal policy, which can affect the real economy, inequality, and asset markets [20], but offers solutions that reflect changing attitudes towards taxes in society [14]. Taxes are traditionally presented as the main mechanism for financing government spending, and social policy as something determined and limited by tax revenues. Tax changes cannot happen quickly, as P. Krugman [26] says, and must be agreed with all state actors.

The essential proposal of MMT is to change the tax-spend cycle to the spend-tax one [11]. In other words, changing the tax cycle will absolve policymakers of the misappropriation of collected revenues and provide a practical tool for evaluating the performance of governments. But the duration of the spend-tax cycle completely destroys the above advantages.

The impact of MMT results on the sequencing of the tax-spend cycle, when taxes are not limited to revenue generation and tax receipts do not precede government spending, deserves closer examination. To accept the new theory, it is necessary to consider and understand tax policy more fully than is presented in MMT, in terms of a wider range of macroeconomic and social functions.

Therefore, T. Palley, as part of the proof of the inconsistency of the theory, made a calculation of the consequences when using MMT. According to his results, if spending on healthcare, education and climate is planned by the authors of MMT in the amount of 12% of GDP, then when the level of full employment is reached, there should be an excess aggregate demand of 13% of GDP. However, to eliminate it, it will be necessary to raise taxes by 78%. The economist argues that MMT is a simplified Keynesian economics with a tendency to underestimate political risks in implementation [29]. T. Palley points out that the old theses are understandable in MMT, but the new ones are not. In our opinion, despite the fact that the data of such calculations should be rechecked on other models that take into account the employment under the employment guarantee program and the level of inflation at the time MMT is adopted, T. Palley ended the discussion on the applicability of the new theory from the tax point of view.

Full employment as a replacement for unemployment benefits, or the Great Economic Idea

The creators of MMT and their supporters put forward the program of full employment as the main achievement of MMT, calling it universal. It is undeniable that unemployment increases the power of employers over workers, but MMT critics point out that a new approach is unnecessary given that many advanced economies have achieved full-employment growth without adopting MMT, such as Sweden, Canada, Singapore, and Chile [7].

The basic employment model in MMT is that the government must announce a fixed salary package for anyone who is looking for a job but cannot find one that suits their preferences. MMT recommends that jobs be oriented towards the service sector (healthcare, elderly care, green economy, etc.). Since the market price of the unemployed is zero, the government can create a market for these potential workers by setting the price it is willing to pay to hire them. The federal government will squeeze out a significant part of the labor market and destroy low-paid sectors of the economy. At the same time, MMT does not take into account the professional composition of the unemployed (the training level) and does not take into account the need for additional training. The question arises (which remains unanswered until the proposed measures are put into practice): when potential zero-pay workers are exhausted, will there be a shift from low-wage private-sector jobs to government-guaranteed jobs?

Obviously, this model is not only not universal, but also does not answer the question: will many unemployed people

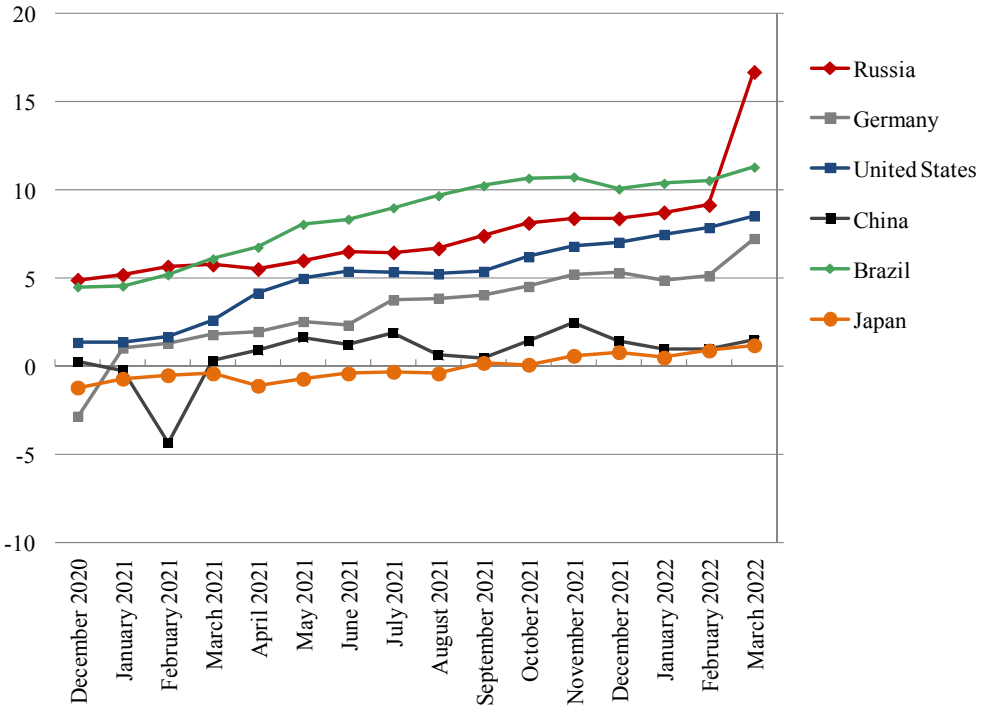


Fig. 2. **Inflation dynamics in different countries, %**

Source: compiled by the authors based on the data: Triami Media BV. CPI inflation – current international consumer price index inflation. URL: https://www.global-rates.com/en/economic-indicators/inflation/consumer-prices/cpi/cpi.aspx

agree to the minimum conditions that will be offered? The answer to this question has not yet been found and is not presented in MMT publications, and, in our opinion, it will be essential in the light of the coming automation of jobs.

Public work is a dream for everyone, but will a busy worker be paid fairly without a market valuation? And won't that lead to even more exploitation? These issues may be identified in future empirical studies rather than theoretical models.

Advantages and disadvantages of inflation in modern conditions

The component of government spending in MMT predetermines that inflation will occur only when the domestic economy exceeds the level of full employment. Is this true?

Monetarism suggests a similar inflationary origin but stipulates that inflation is primarily a monetary phenomenon that exists when the money supply grows at a rate that outpaces the growth in output. Over the past year, there has been a clear upward trend in inflation in most countries, regardless of their level of development (*Fig. 2*), and when comparing Russian indicators and indicators of a country using MMT in its policy (USA), there is a clear discrepancy between monetary and macroeconomic indicators (*Fig. 3*).

Proponents of MMT put forward a different approach to the nature of inflation. Thus, MMT considers the problem of inflation as an integral part of the power relations between workers and capital, that is, class conflict mediated by the state within the capitalist system. Inflation is spiraling out of control as workers and business owners step up their efforts to claim a bigger share of the national income. MMT offers state control over wages and prices as a kind of arbitrage in the ongoing struggle. Therefore, according to B. Putnam [9], the fusion of monetary and fiscal policy naturally leads to the need to study the applicability of MMT and its implications for future inflation forecasting. In our opinion, such a view gives an oversimplified understanding of inflation, especially in modern conditions, when

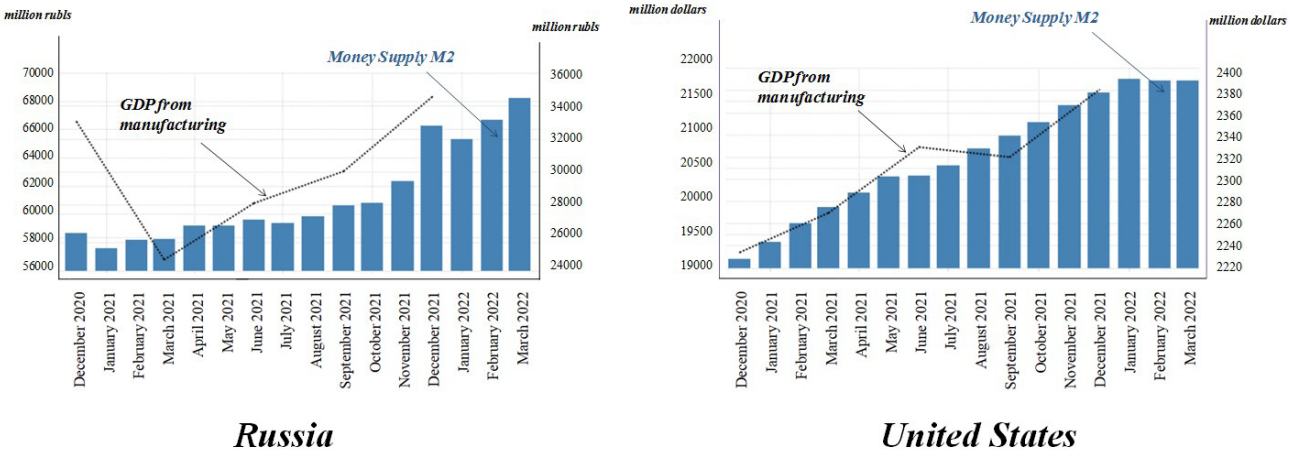


Fig. 3. **Monetary and macroeconomic indicators of countries**

Source: compiled by the authors based on the data: Trading Economics. Russia — Economic indicators. URL: https://ru.tradingeconomics.com/russia/indicators; Trading Economics. United States Indicators. URL: https://tradingeconomics.com/united-states/indicators.

inflation growth is required and will take a long time to achieve.

Macroeconomic indicators as an assessment of the effectiveness of public policy have become less popular, but hyperinflation can lead to the fact that public goals will be disavowed. Policymakers should worry about inflation [18], but this does not mean that the government can improve the social welfare by actively participating in the price-setting process.

CONCLUSIONS

The world is waiting for an economic recipe

What do we do when the economy is in a recession? Society and policymakers are again waiting for a salvation recipe from economists. Will MMT become such a recipe? There is a gap between the need for action and the recognition of this need by the government. As a result, when an error is recognized and corrective actions are taken, damage can be done, and the corrective actions can turn into an additional error [30].

MMT assumes a renaissance of a number of well-known theories and actualizes this renaissance at a time when the need for new theories is high. Our conclusion is that the most important part of MMT is the political package, not the economic package. Therefore, it is impossible to recommend MMT as a ready-made universal tool for solving the problems of economic development of states, including Russia. Only upon appropriate political and/or public demand, the proposed MMT tools should not be initially rejected, and some of them may be studied in order to apply them in practice.

Considering the possibility of practical implementation, we note that Greece has formed a negative attitude towards MMT, as it has already tested some of its measures, which has led to significant inflation, higher long-term interest rates, higher risk premiums, and financial crises [19].

We also note that researchers of the Chinese economy argue that MMT elements are partially used in projects that promote industrial modernization (for example, the "Made in China 2025" project), solve the problem of environmental degradation, and create goods production and services that improve the quality of life [17]. Thus, Chinese job security is seen as a political proposal in addition to the new infrastructure and allows to achieve a more complete list of public goals [17].

On the other hand, for example, in Croatia, they believe that it is possible to use the MMT conclusions for small countries [15], but it is necessary to convert all external debt into the national currency, for which there is neither political will, nor leaders, nor money. Therefore, for small countries, this path is closed,

unless the European Bank decides to start experiments. The International Organizations of the Supreme Audit Institutions can be involved in the implementation of this task, contributing to a more objective distribution of World Bank funds to countries where unemployment is critical.

But in any case, the idea of public goals financed indefinitely by the state will remain attractive for a long time to come. The crisis does not give time for necessary reforms, and MMT can increase the risks by unknown measures and inconsistency in the actions of various government structures, which leads to the conclusion that "modern monetary theory" rather only reveals problem areas of modern economic policy, rather than sets a new agenda.

REFERENCES

1. Wray L.R. Modern monetary theory: A primer on macroeconomics for sovereign money systems. London: Palgrave Macmillan; 2012. 289 p.
2. Wray L.R., Dantas F., Fullwiler S., Tcherneva P., Kelton S. Public service employment: A path to full employment. Annandale-on-Hudson, NY: Levy Economics Institute of Bard College; 2018. 52 p.
3. Mitchell W., Wray L.R., Watts M. Macroeconomics. London: Red Globe Press; 2019. 604 p.
4. Kelton S. The deficit myth: Modern monetary theory and the birth of the people's economy. New York: Public Affairs; 2020. 336 p.
5. Andryushin S.A. Is money the creation of the state or the market? (On the "modern monetary theory" as described in the textbook by W. Mitchell, L.R. Wray and M. Watts "Macroeconomics"). *Voprosy ekonomiki*. 2020;(6):121–134. (In Russ.). DOI: 10.32609/0042–8736–2020–6–121–134
6. Afontsev S.A. Methodological foundations of modern economic and political analysis. *Istoki*. 1998;(3):72–122. URL: https://ecsocman.hse.ru/data/402/700/1219/004Afontsev.pdf (In Russ.).
7. Skousen M. There's much ruin in a nation: An analysis of modern monetary theory. *Atlantic Economic Journal*. 2020;48(1):11–21. DOI: 10.1007/s11293–020–09651–9
8. Chohan U.W. The lives or livelihoods? The coronavirus dilemma for administrations. *CASS Working Papers on Economics & National Affairs*. 2020: EC 021UC. DOI: 10.2139/ssrn.3584726
9. Putnam B.H. From phase transitions to Modern Monetary Theory: A framework for analyzing the pandemic of 2020. *Review of Financial Economics*. 2021;39(1):3–19. DOI: 10.1002/rfe.1122
10. Adil M.H., Hatekar N.R., Ghosh T. Role of money in the monetary policy: A new Keynesian and new monetarist perspective. In: Barnett W.A., Sergi B.S., eds. Environmental, social, and governance perspectives on economic development in Asia. Bingley: Emerald Publishing Ltd; 2021:37–67. (International Symposia in Economic Theory and Econometrics. Vol. 29A). DOI: 10.1108/S 1571–03862021000029A018
11. Baker A., Murphy R. Modern monetary theory and the changing role of tax in society. *Social Policy and Society*. 2020;19(3):454–469. DOI: 10.1017/S 1474746420000056
12. Baronian L. The time-spaces of capitalism: Suzanne De Brunhoff and monetary thought after Marx. *Journal of the History of Economic Thought*. 2021;43(3):420–432. DOI:10.1017/S 1053837220000218
13. Brady G.L. Modern monetary theory: Some additional dimensions. *Atlantic Economic Journal*. 2020;48(1):1–9. DOI: 10.1007/s11293–020–09654–6
14. Coronado J. Giving MMT the credit it is due. *Business Economics*. 2020;55(1):23–25. DOI: 10.1057/s11369–019–00158-z
15. Grdešić M. What does modern monetary theory offer? *Croatian Political Science Review*. 2020;57(3):99–118. DOI: 10.20901/pm.57.3.04
16. Epstein G.A. What's wrong with modern money theory? A policy critique. Cham: Palgrave Pivot; 2019. 111 p. DOI: 10.1007/978–3–030–26504–5
17. Huang Y. Monetary sovereignty for public purpose sustaining and complementing China's "new infrastructure". *Global Institute for Sustainable Prosperity Policy Report*. 2020;(104). URL: http://www.global-isp.org/wp-content/uploads/PolicyReport-104.pdf
18. Mankiw N.G. A skeptic's guide to modern monetary theory. *AEA Papers and Proceedings*. 2020;110:141–144. DOI: 10.1257/pandp.20201102

19. Tavlas G.S. Modern monetary theory meets Greece and Chicago. *Cato Journal*. 2021;41(1):1–23. URL: https://www.cato.org/sites/cato.org/files/2021–01/cj-v41n1–1.pdf
20. Mann C.L. Some observations on MMT: What's right, not right, and what's too simplistic. *Business Economics*. 2020;55(1):21–22. DOI: 10.1057/s11369–019–00153–4
21. Kirdina-Chandler S.G. Mechanism of money circulation as a subject of mesoeconomic analysis. *Zhurnal institutsional'nykh issledovanii = Journal of Institutional Studies*. 2019;11(3):7–20. (In Russ.). DOI: 10.17835/2076–6297.2019.11.3.007–020
22. Mau V.A. Economics and politics in 2019–2020: Global challenges and national answers. *Voprosy ekonomiki*. 2020;(3):5–27. (In Russ.). DOI: 10.32609/0042–8736–2020–3–5–27
23. Afontsev S.A. Political markets and economic policy. Moscow: Lenand; 2015. 384 p. (In Russ.).
24. Palley T.I. Money, fiscal policy, and interest rates: A critique of Modern Monetary Theory. IMK Working Paper. 2013;(109). URL: https://www.econstor.eu/bitstream/10419/105973/1/imk-wp_109_2013.pdf
25. Baranovsky V.G. New international order: Overcoming or transforming the existing pattern? *Mirovaya ekonomika i mezhdunarodnye otnosheniya = World Economy and International Relations*. 2019;63(5):7–23. (In Russ.). DOI: 10.20542/0131–2227–2019–63–5–7–23
26. Krugman P. What's wrong with functional finance? (Wonkish). The New York Times. Feb. 12, 2019. URL: https://www.nytimes.com/2019/02/12/opinion/whats-wrong-with-functional-finance-wonkish.html
27. Demidenko M., Bukreev K., Romanyuk V. Emission of the Central Bank: Myth and reality. *Bankauski vesnik = Bank Bulletin Journal*. 2020;(10):21–36. (In Russ.).
28. Rabbitt T. Modern monetary theory. Economics Honors Papers. 2020;(31). URL: https://digitalcommons.conncoll.edu/cgi/viewcontent.cgi?article=1035&context=econhp
29. Palley T.I. The critics of modern money theory (MMT) are right. *Review of Political Economy*. 2015;27(1):45–61. DOI: 10.1080/09538259.2014.957473
30. Friedman M. Lerner on the economics of control. *Journal of Political Economy*. 1947;55(5):405–416. DOI: 10.1086/256579

ABOUT THE AUTHORS

Igor M. Stepnov — Dr. Sci. (Econ.), Professor, Head of Asset Management Department, Moscow State Institute of International Relations (MGIMO University), Moscow, Russia; Professor of the Department of Corporate Finance and Corporate Governance, Financial University, Moscow, Russia
http://orcid.org/0000-0003-4107-6397
Corresponding author
stepnoff@inbox.ru

Julia A. Kovalchuk — Dr. Sci. (Econ.), Professor, Professor of Asset Management Department, Moscow State Institute of International Relations (MGIMO University), Moscow, Russia; Professor of Department of Energy Service and Energy Saving Management, Moscow Aviation Institute, Moscow, Russia
http://orcid.org/0000-0002-9959-3090
fm-science@inbox.ru

Margarita V. Melnik — Dr. Sci. (Econ.), Professor, Professor of Department of Audit and Corporate Reporting, Financial University, Moscow, Russia
http://orcid.org/0000-0002-7428-9535
mmargarita@fa.ru

Tamara Petrovic — Cand. Sci. (Econ.), PhD in Economics, CEO, Innovation Project Management Center — IPM Center Ltd, Pancevo, Serbia
http://orcid.org/0000-0001-6842-3032
obradovictamara@yahoo.com

***Authors' declared contributions*:**
I. M. Stepnov — formulation of the problem, preparation of the initial draft of the study, development of the article's conceptual framework, description of the results and formation of conclusions of the study.
J. A. Kovalchuk — references selection, analysis of theoretical provisions on the article's subject.
M. V. Melnik — research conceptualization, review and editing, verification of the research results' validity.
T. Petrovic — analysis of the international practice of public finance.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 25.10.2021; revised on 25.11.2021 and accepted for publication on 27.12.2021.
The authors read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-19-32
UDC 336.6(045)
JEL 016

Trends and Factors of Development of the Financial System Elements of the Modern Russian Economy

G.N. Ismoilov
Saint Petersburg State University, Saint Petersburg, Russia

ABSTRACT

The subject of the research is the segments of the financial system of the Russian Federation: the budget system, the banking sector, the stock and insurance markets, and the currency policy of the state. **The purpose** of the study is to determine the trends and factors in the development of the main elements of the financial system at the present stage. **The relevance** of scientific research is due to the fact that the financial system is a key element of the strategy of socio-economic development of any state, providing economic processes with financial resources and capital. The author uses the following **methods**: analysis, synthesis, generalization, and the logical method. The study highlights promising directions, ways and mechanisms for the development of the Russian financial system that are relevant in the 2020s. The key factors influencing their trends and threats that create barriers are analyzed. The main directions, ways and mechanisms for stimulating the further development of the elements of the financial system are described. The author **concludes** that due to the spread of the coronavirus pandemic and economic sanctions imposed on Russia, the stability of the Russian financial system has been violated, which requires the adoption of state regulation mechanisms to improve the activities of financial institutions. The prospect of further research on this topic may be related to the development of areas for improving individual elements of the Russian financial system.

***Keywords*:** financial system; financial markets; banking system; financial capital; financial resources; finance; financial system management; stocks and bonds market; financial institutions; budgetary system; monetary policy; insurance market

For citation: **Ismoilov G.N. Trends and factors of development of the financial system elements of the modern Russian economy. *Finance: Theory and Practice.* 2022;26(3):19-32. DOI: 10.26794/2587-5671-2022-26-3-19-32**

INTRODUCTION

The financial system of modern Russia is the main element in ensuring national economic security and the functioning of the state [1].

Considering the trends and factors in the development of the elements of the financial system of the modern Russian economy, it is necessary to analyze information that reveals the state of the budget and banking systems, the stock and insurance market, and the monetary policy of the state.

Today, the financial system is one of the key elements of the socio-economic development of Russia. At the same time, within the framework of sustainable development, important issues are:

financial regulation;

distribution of state financial resources.

Yu. S. Lankina and L. O. Serdyukova, as part of their research, consider a comparative description of the financial systems of different countries, analyzing indicators that characterize the state and structure of the financial systems of Russia, France, the USA, and Japan [2].

V. V. Balyan analyzes the critical threats to the Russian financial system, reveals their content and main characteristics, and substantiates the need to optimize state policy to ensure the country's financial security [3].

E. S. Ibragimova and M.U Khadzhimuratov in their study reveal the essence of the financial system, and its structure, consider the budgetary system as a link in the Russian financial system, and also reveal the problems of the Russian financial system and ways to solve them [4].

L. A. Mylnikova explored the main problems of the development of the financial system of Russia, determined its structure, and proposed measures to address them and increase the economic growth of Russia [5].

However, despite the large base of scientific literature and sources on the Russian financial system problems, the factors and mechanisms that have a direct impact on the trends and prospects for the development of individual elements of the financial system, such as the budgetary system, stock market, banking and insurance industries, monetary policy remain unexplored.

The relevance of this scientific research is due to the following factors:

firstly, the financial system is a key element of the strategy of socio-economic development of any state, providing economic processes with financial resources and capital;

secondly, the financial system of modern Russia is still in the process of formation, however, in the context of the COVID-19 pandemic and economic sanctions, it has faced a number of urgent problems that create barriers to further development.

For this reason, the purpose of this scientific paper is to analyze the factors influencing the trends and prospects for the development of the main elements of the Russian financial system at the present stage.

The methodological basis of the study consists of general scientific methods for understanding economic phenomena, which make it possible to consider the development of elements of the Russian financial system, to explore the most pressing problems and barriers.

The theoretical basis of the scientific article was the results of fundamental and applied research in economic theory, state and municipal government, budget regulation, banking regulation and finance.

TRENDS AND FACTORS IN THE DEVELOPMENT OF THE BUDGET SYSTEM OF MODERN RUSSIA

In the current conditions of socio-economic instability, an important element of the economic policy and development strategy of the Russian financial system is the formation of an effective budgetary system. In the Russian Federation there are three levels of budget: federal, regional and municipal.

The budgetary system of Russia is based on economic relations and the state structure, regulated by the legislation of Russia, the aggregate of the federal budget, regional budgets, local budgets and budgets of state extra-budgetary [6, 7].

The main features of the formation of revenues and expenditures of the state budget of Russia are due to the influence of the oil and gas sector, the policy of socio-economic alignment of the regions and the general vector of development of the state. The characteristic features of the budget policy of Russia are the impact of current problems and barriers, the elimination of which is the main condition for the further development of the budgetary system of the Russian Federation.

The stability of the budgetary system is the basis of the socio-economic development of Russia. Due to the negative impact of external and internal economic factors, public authorities are faced with the need to make quick decisions in order to achieve fiscal sustainability. The state implements budget policy, and adopts legal acts regulating budgetary issues. The solution to budgetary problems occurs at all levels of government: federal, regional and municipal.

In our opinion, the main factors affecting the deficit/surplus of the state budget of the Russian Federation in a dynamic aspect are:

1. The price environment of the world oil and natural gas market, since oil and gas revenues play an important role in shaping the stability of the budgetary system of Russia.

2. The economic situation in the world and the Russian economy, since the macroeconomic indicators of the markets reflect the general state of the financial system, which includes budget policy.

Instability in the economic and financial activities of the state is reflected in the budget deficit. The reasons for the budget deficit can be different. Incomes reflect the state of the economy; thus, an inefficient economy necessarily creates difficulties in balancing the budget. In addition, the deficit may also arise as a result of miscalculations in the management of the national economy, and subsequently deliberate overstatement of expenditures and a reduction in the tax burden.

The reasons for the budget deficit also include [8]:

- economic crisis;
- inability of the state to manage the financial situation;
- high level of corruption schemes and crimes in the sphere of public administration;
- growth of social programs not provided with state revenues;
- inefficient tax policy, leading to a high share of the shadow economy;
- fiscal federalism.

One of the negative factors of the budget deficit in Russia is the misuse of budget funds. It is also important to note that the current stage of development of the budgetary system of the Russian Federation is associated with the negative impact of the COVID-19 pandemic and economic sanctions. The receipt of revenues to the state budget of the country was significantly disrupted due to the fall in production and consumption, which led to a decrease in revenues to the tax system of the country.

With respect to the structure of budget expenditures changed dramatically, since the areas of medicine, social policy and the national economy required much larger financial injections than planned. As a result, a budget deficit is formed, the elimination of which has become one of the main tasks facing the state.

Let us consider the situation until the 2020s. In 2000–2008, the Russian economy functioned exclusively under favorable economic conditions — oil prices rose from about 17.1 to 94.4 US dollars per barrel, i.e. 5.5 times. The country had a budget surplus. By 2009, in the context of the global financial crisis, revenues to the country's tax system began to decline, and, as a result, a state budget deficit arose.

Fig. 1 shows that in recent years, from 2015 to 2017 there was a budget deficit. This was caused by the collision of the Russian economy with the negative impact of factors in the context of a collapse in prices on the world oil and natural gas market and the regime of economic and political sanctions of Western countries, where the main trading and investment partners of Russian transnational companies and banks are concentrated.

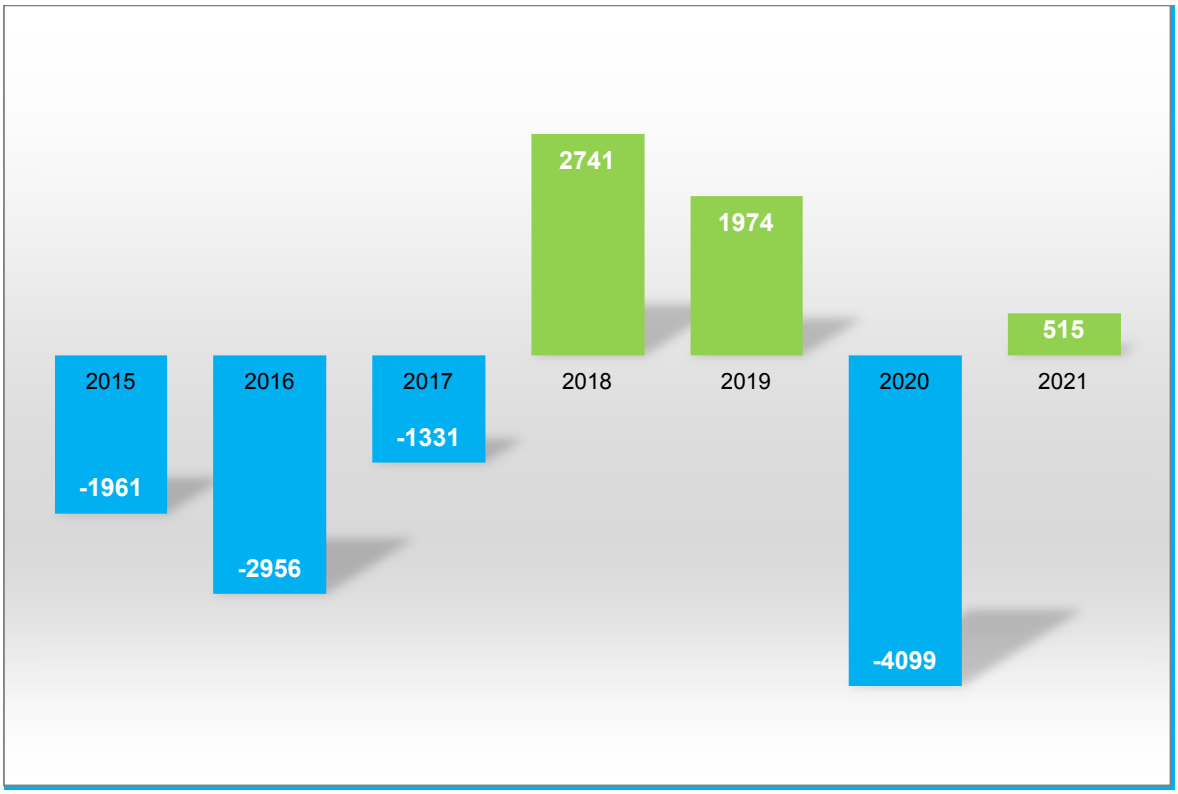


Fig. 1. **Dynamics of the surplus/deficit of the state budget of the Russian Federation in 2015–2021, billion rubles**

Source: compiled by the author based on data from the Ministry of Finance of the Russian Federation. URL: https://www.minfin.ru/ru/statistics/fedbud/ (accessed on 30.04.2022).

In 2018–2019, the dynamics of the state budget deficit changed in favor of a surplus, which can be explained by the following reasons:

- adaptation of the Russian economic system to the sanctions regime;
- withdrawal of exports of Russian companies from the West to the East;
- price growth recovery in the oil and natural gas market.

At the end of 2020, the state budget, due to the crisis of the COVID-19 pandemic, ended up with a record deficit of –4.099 billion rubles. And in 2021, a surplus of 514.76 billion rubles was recorded, thanks to rising oil prices and the rapid recovery of the Russian economy. Nevertheless, the problem of the budget deficit for Russia is still relevant.

The task of the Russian government is to form a mechanism for self-reproduction of the budget deficit. It is necessary to create conditions in which income is higher than expenses.

Fig. 2 shows the detailed structure of the revenues of the state budget of the Russian Federation in 2021.

Oil and gas revenues accounted for 36% of all revenues in 2021. These are mainly taxes on value added (5,479.53 billion rubles), import value added (3,732.99 billion rubles), and personal income (1,552.43 billion rubles). Each of these taxes plays an important role in shaping the revenue of the Russian budget.

The reason for the increased attention to the structure of state budget revenues in 2021 is that it is precisely such deductions as VAT, import VAT and personal income tax that may show their decline due to the impact of the coronavirus pandemic and sanctions against Russia. In the last 2020–2021, a pattern was observed: the decline in business activity in the national economy of the state led to a decrease in the volume of tax payments received from economic entities.

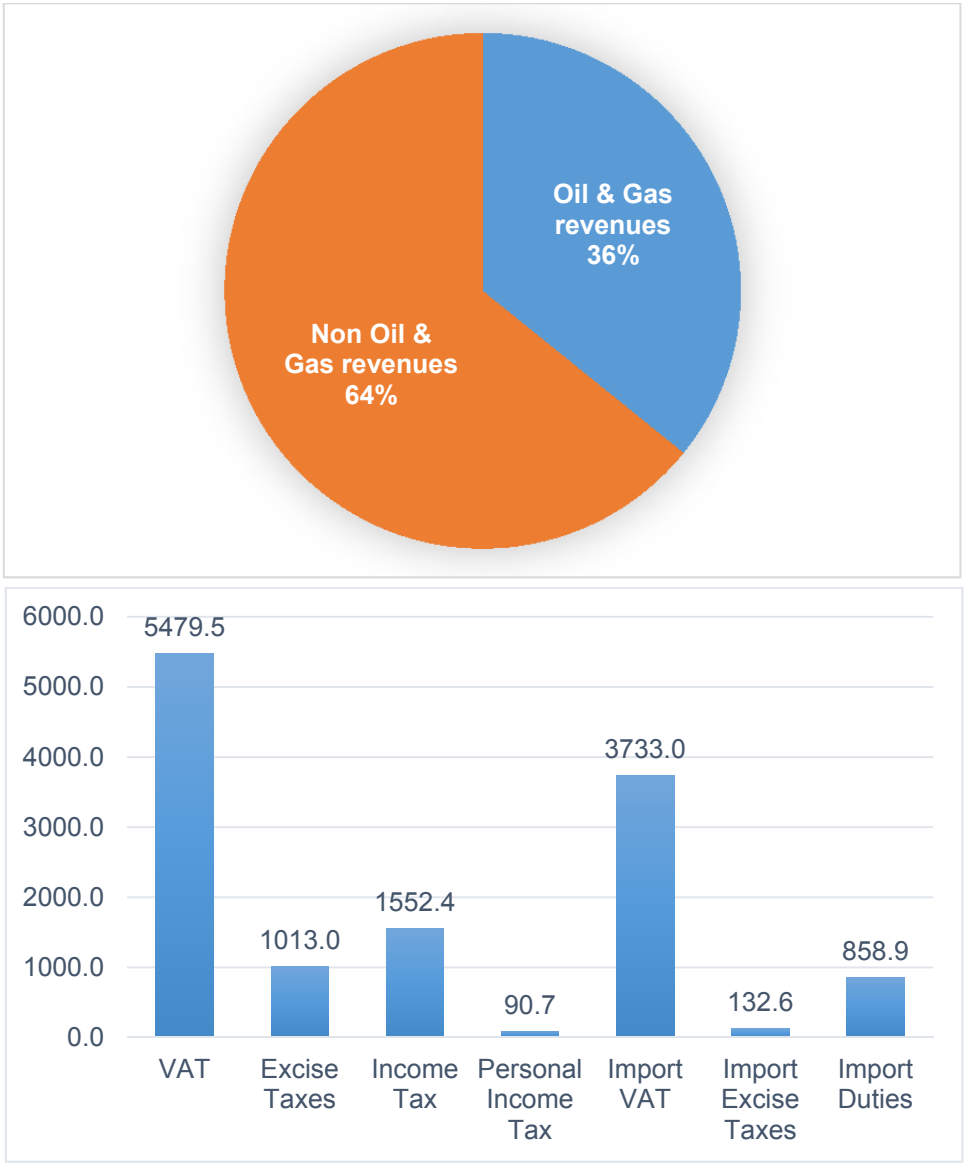


Fig. 2. **Structure of revenues of the state budget of the Russian Federation in 2021, in %, billion rubles**

Source: compiled by the author based on data from the Ministry of Finance of the Russian Federation. URL: https://www.minfin.ru/ru/statistics/fedbud/ (accessed on 30.04.2022).

Based on this, an important task of budget regulation of the national economy of the Russian Federation is to improve the state budget, aimed at reducing the size of the deficit and creating a budget surplus.

The main tools for improving the budget policy can be [9]:

- optimization of expenditure items of the budget, where there is the greatest shortage of funds and growth in funding;
- creation of free economic zones and programs to stimulate investments in fixed assets to support economic entities whose activities increase tax revenues to the budget;
- active use of domestic municipal debt market instruments to cover the budget deficit;
- application of tools of public-private, regional-private and municipal-private partnerships in the implementation of strategic infrastructure facilities to optimize budget expenditures.

Fig. 3 shows the most optimal and efficient algorithm for the procedure for optimizing the structure of revenues and expenditures of the budget of the Russian Federation.

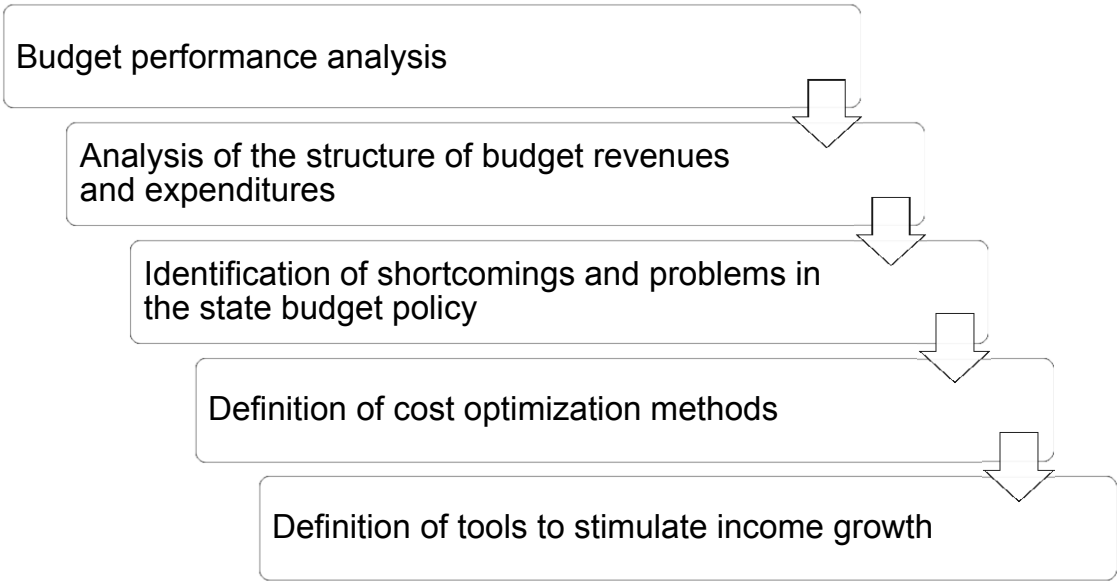


Fig. 3. **Algorithm for optimizing the structure of revenues and expenditures of the state budget of the Russian Federation**

Source: compiled by the author.

Based on *Fig. 3* algorithm for optimizing the structure of revenues and expenditures of the state budget consists of 5 main stages. Such an algorithm makes it possible to form a mechanism for certain purposeful actions of the authorities to optimize the structure of budget revenues and expenditures in order to comply with the principle of balance and efficiency in the use of state budget funds.

TRENDS AND FACTORS IN THE DEVELOPMENT OF THE BANKING SYSTEM IN MODERN RUSSIA

Today, the banking system is a system-forming element in the development of the market economy of the Russian Federation. The key task of credit institutions is the accumulation and distribution of funds between various forms of economic entities, each of which consumes or produces [10].

Analyzing the current state of the Russian banking sector, it is necessary to refer to *Fig. 4*, which shows the dynamics of changes in the size of domestic banks' assets.

In the period from 2014 to 2021, the total assets of the Russian banking system changed from 57.4 to 112.5 trillion rubles — almost doubled, despite the negative impact of the sanctions against Russia and the COVID-19 pandemic. Thus, the development of the Russian banking system is characterized by the positive results of its activities. Credit institutions have a sufficient level of financial stability.

At the same time, there is a trend toward a reduction in the number of operating credit institutions, which may seem like a negative signal when assessing the stability of the Russian banking system (*Fig. 5*).

Back in 2014, there were 923 credit institutions operating in the Russian Federation, and as of January 2022, there were already 370, which is 2.5 times less. At the same time, the bulk of such credit institutions is commercial banks with a universal license (232) and with a basic license (103). Only 35 companies operate among non-bank credit institutions.

However, such a decrease in the number of operating credit institutions is explained by the strategy of the state and the leadership of the Central Bank of the Russian Federation, aimed at ensuring the economic security of the banking system, which has a positive effect on the national economic security of the state [11]. There is a revocation of licenses from commercial banks with a low level of capital provision and the quality of portfolio credit risk management.

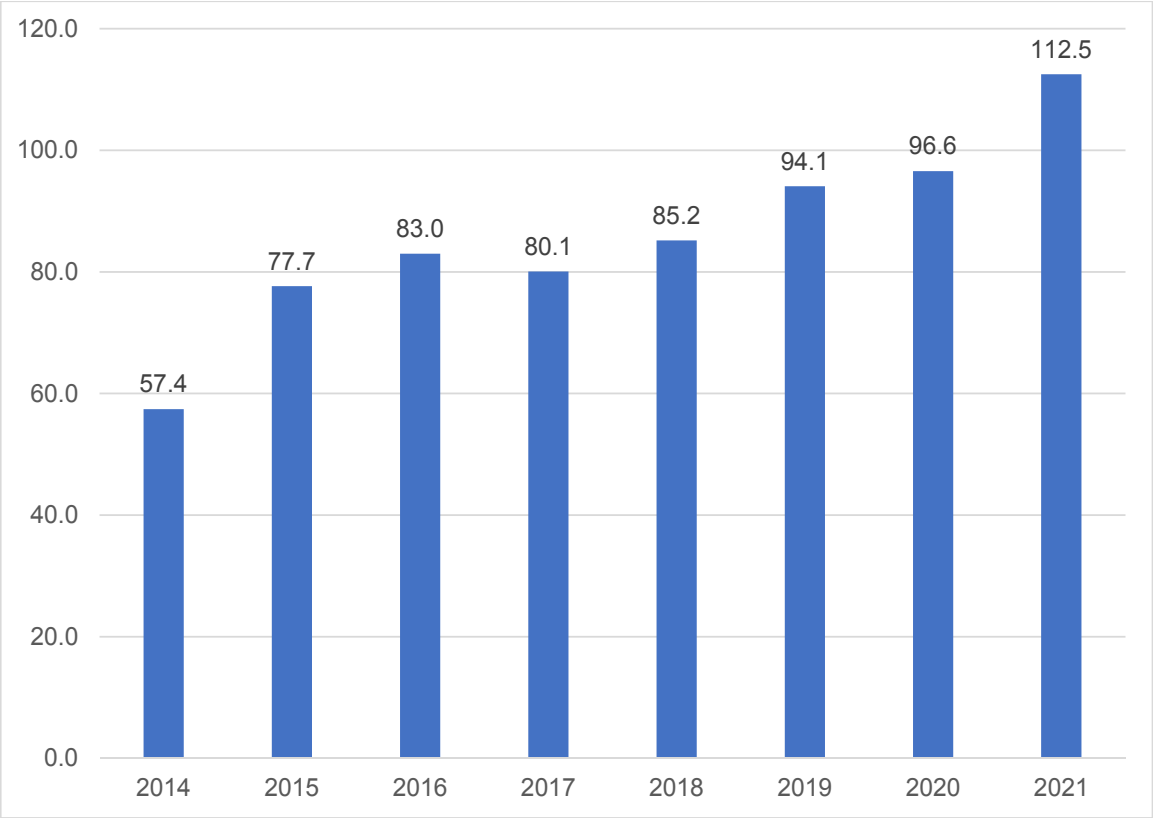


Fig. 4. **Dynamics of the total assets of the Russian banking system in 2014–2021, trillion rubles**

Source: compiled by the author based on data from the Central Bank of the Russian Federation. URL: https://www.cbr.ru/statistics/bank_sector/review/ (accessed on 30.04.2022).

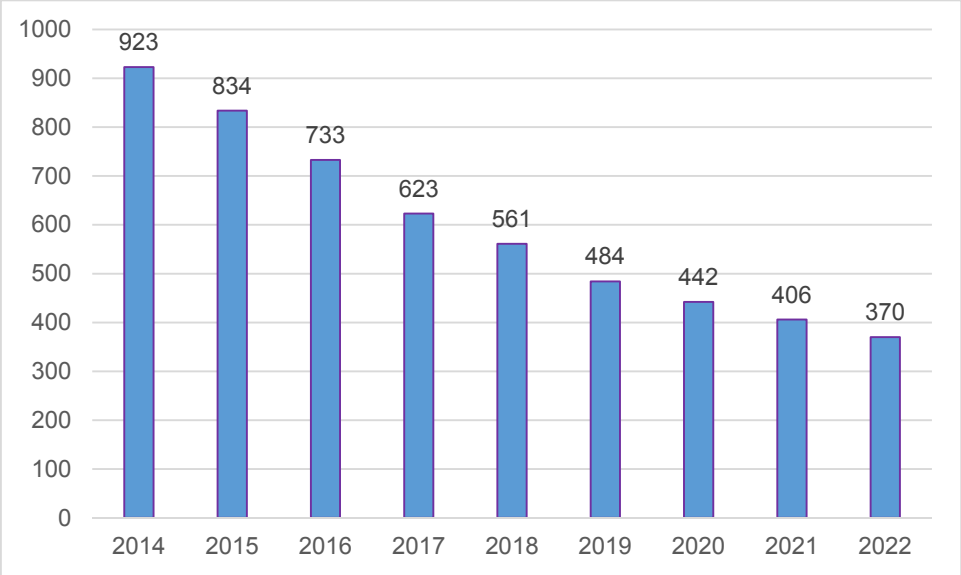


Fig. 5. **Dynamics of operating credit institutions in Russia in 2014–2022**

Source: compiled by the author based on data from the Central Bank of the Russian Federation. URL: https://www.cbr.ru/statistics/bank_sector/review/ (accessed on 30.04.2022).

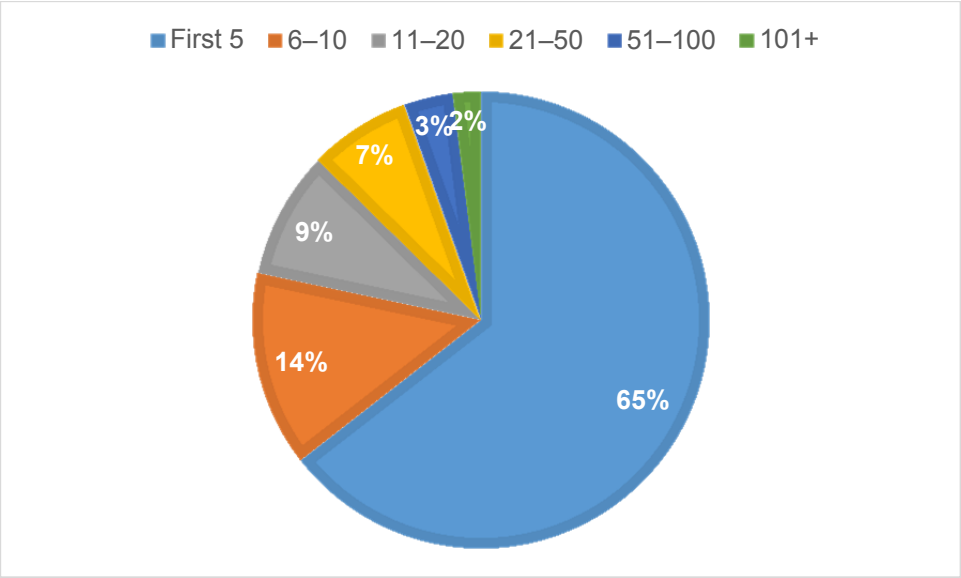


Fig. 6. **Asset concentration in the Russian banking sector, February 2022**

Source: compiled by the author based on data from the Central Bank of the Russian Federation. URL: https://www.cbr.ru/statistics/bank_sector/review/ (accessed on 30.04.2022).

Another indicator of the current state of the banking system of the Russian Federation is the concentration of assets in operating credit institutions, ranked by assets into different groups (*Fig. 6*).

Thus, 65% of the total assets of the banking system are concentrated in the five largest commercial banks in Russia which indicates a high level of consolidation in the banking sector. At the same time, many of the largest credit organizations are commercial banks with the participation of state capital. In particular, the market leader is Sberbank, which actually concentrates a third of the entire banking market of Russia [12].

This characteristic of the Russian banking system is rather negative than positive in our opinion. On the other hand, the largest Russian commercial banks, created with the participation of the state, have a high level of trust and assistance from regulatory bodies, which ensures their financial stability, and hence the prospects for the development of the financial system [13].

At present, it is necessary to pay attention to the fact that the unstable environmental situation associated with the economic sanctions against Russia and the COVID-19 pandemic persists, as a result of which various threats and risks may arise that can negatively affect the stability of the banking system of modern Russia and lead to the following negative consequences such as:

- decrease in the volume of retail sales of banking products, including a decrease in the volume of consumer lending;
- increase in the risks of the banks' loan portfolio, including the growth of overdue consumer debt and the portfolio of problem loans.

TRENDS AND DEVELOPMENT FACTORS OF THE STOCK MARKET IN MODERN RUSSIA

The events of 2020 will be remembered for an unexpected shock, the Black Swan, which turned out to be the COVID-19 pandemic. Its impact through the fear factor and the adoption of quarantine measures and restrictions led to the fact that the economic growth of Russia and other countries became negative. Accordingly, the negative impact of the coronavirus pandemic was also expected on the dynamics of financial markets, where one of the main participants in trading is mutual funds, which are divided into three main types:

- open-end investment funds;

- closed-end investment funds;
- interval investment funds.

It is also necessary to highlight a list of the main problems that impede the development of investment funds in Russia, which is why the stock market faces a significant barrier to its formation.

1. The low level of information transparency of mutual investment funds about their profitability and risks in the implementation of their main commercial activities. The essence of this problem lies in the fact that sometimes there is no information on such indicators as on the official websites of mutual funds:

- profitability of a mutual fund for a certain period of time;
- risk indicators of the investment portfolio of a unit investment fund;
- the ratio of profitability and risk of the investment portfolio.

2. The absence of a mechanism that guarantees the protection of the financial capital of investors and depositors of mutual investment funds. The essence of this problem lies in the fact that investors who purchase mutual funds for the purpose of investing in mutual funds risk their own capital and property. However, there are no mechanisms that guarantee the protection of the financial capital of investors from such cases as:

- unprofessional competence of investment portfolio managers;
- dishonest activity of the management company;
- bankruptcy of the management company that managed this mutual investment fund.

3. The low level of development of exchange traded investment funds, which can become an alternative to open-end investment funds and, at least, take second place in NAV after open-end investment funds, exceeding the performance of closed-end and interval investment funds [14].

The essence of this problem lies in the fact that today in the domestic practice of Russia there are certain legislative restrictions that do not allow the operation of exchange traded funds (ETFs), the main advantage of which is that their price is determined by the results of trading on the stock exchange in real time with a direct the dependence of the share price on the dynamics of the index reproduced by it.

In order to solve these problems, which will stimulate the prospects for further development of the stock market through the mechanism for the development of mutual funds in the Russian Federation, the following solutions can be proposed.

To solve the first problem, it is necessary to pass a legislative decision obliging managers of mutual funds to publish certain information in the media accessible to all readers, users and investors.

To solve the second problem, it is necessary to create a special fund for guaranteeing investment capital, which Russian citizens invest in mutual funds. The mechanism for creating a guarantee fund can be approximately the same as the fund used in the Russian banking sector. Thus, management companies that create mutual funds will be required to set aside a certain percentage of assets on a quarterly basis to ensure reserves, the funds of which will be used to provide a guarantee mechanism.

To solve the third problem, it is necessary to make significant amendments to the legal framework of Russia. For example, legislation could be passed that would allow ETFs to become a separate investment vehicle, similar to alternative investment funds (AIFs).

TRENDS AND DEVELOPMENT FACTORS OF THE INSURANCE MARKET IN MODERN RUSSIA

The insurance market is an important structural element in the development of the national financial system of the Russian Federation since it ensures the continuous socio-economic growth of the state through the accumulation and consolidation of financial resources.

Every year, the insurance industry of the Russian economy demonstrates an increase in market capitalization, which is reflected in the data in *Fig. 7*. In recent years, despite the sanctions against Russia and the COVID-19

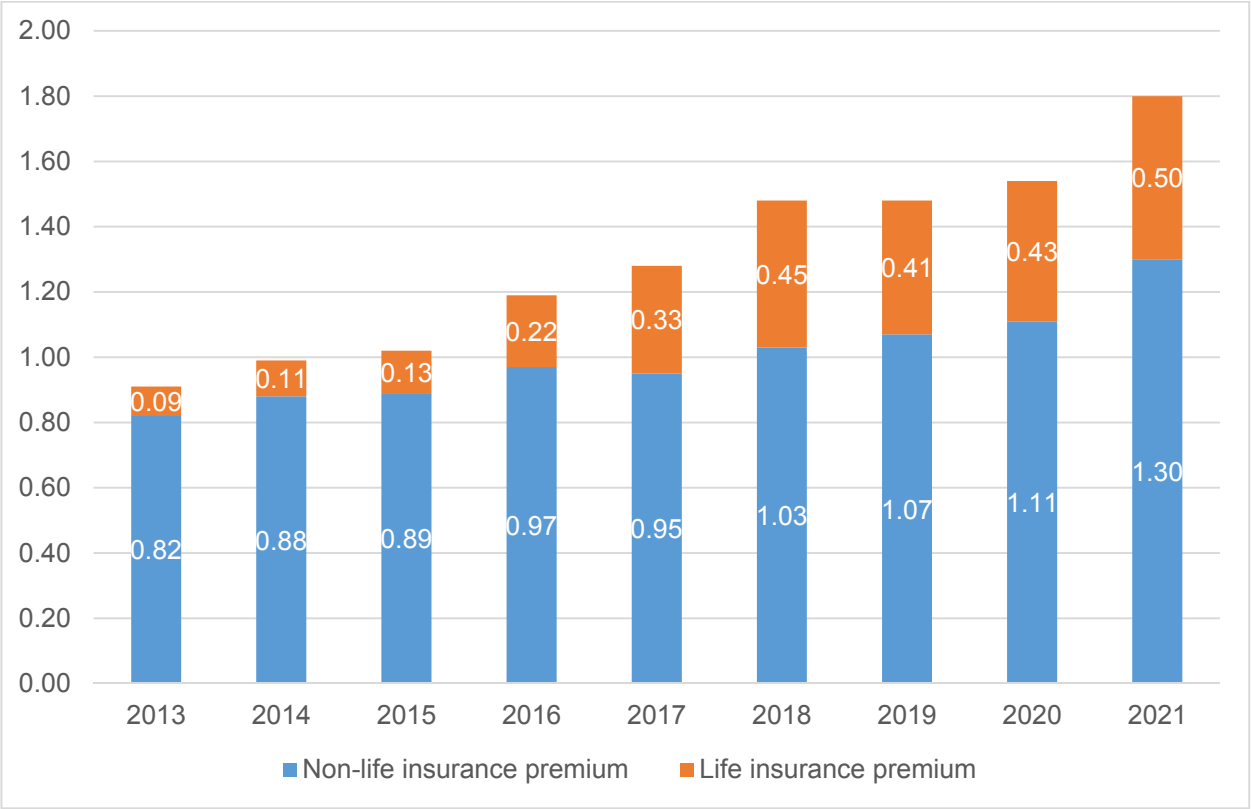


Fig. 7. **Dynamics of the market capitalization of the insurance industry of the Russian Federation in 2013–2021, trillion rubles**

Source: compiled by the author based on data from the Central Bank of the Russian Federation (accessed on 30.04.2022).

pandemic, this growth was 4.1% in 2020, and 17% in 2021.

Today, international financial risks are relevant for the insurance business in Russia, which include:

- instability of the volatility of the exchange rate of the Russian ruble against such foreign currencies as the US dollar, the Chinese yuan and the euro;
- instability of market volatility in the global stock market;
- instability of the market value of the main raw material assets, including oil and natural gas, as they affect the general background of the economic situation in the Russian state;
- volatility of vectors of monetary policy of central banks, especially the US (Fed), Eurozone (ECB) and China (People's Bank of China). This affects the change in trends in the market of interest rates and the cost of financial resources and funds, when attracting external financing or when placing free capital in investment projects, assets and securities.

Analyzing the internal factors in the development of the Russian insurance system, it is important to note that one of the areas/products of insurance companies is compulsory OSAGO insurance, with the following urgent problems:

1. Reducing the size of the compensation fund, which is regarded by clients as a negative signal, meaning the possible absence of material and financial compensation after the occurrence of an insured event.

2. Strengthening the factor of judicial practice in the investigation of insured events under OSAGO, which, firstly, increases the terms of insurance payments; secondly, it complicates the mechanism of material and financial compensation.

3. The growth of fraud and economic crimes in OSAGO leads to an increase in both the volume of insurance payments for fabricated

insured events and the number of refusals of insurance companies from payments for proven insured events. Such a problem acts as an additional trigger for strengthening the factor of judicial practice in the framework of the investigation of OSAGO-insured events.

M. A. Meda, as part of the scientific research, highlights other topical problems of civil liability insurance using the example of OSAGO [15]:

1. Refusal of insurance companies from payments upon the occurrence of an insured event in connection with the refusal to provide movable property for examination.

2. Examination of insured events by specialists and professionals who do not have a state license or accreditation, which increases the number of refusals of insurers when making insurance payments.

When analyzing the development of the insurance market in Russia, it is necessary to highlight the following factors that may become drivers of its further growth:

- optimization of the pricing policy of insurance products and the inclusion of a consumer loyalty program;
- optimization of cross-selling;
- enhancing the product penetration strategy per customer;
- automation of business processes within insurance companies, which will allow them to reduce personnel costs, office rent, etc.

TRENDS AND FACTORS IN THE DEVELOPMENT OF MONETARY POLICY IN MODERN RUSSIA

The current stage of development of the Russian financial system is also characterized by the digital transformation of the Russian economy, which leads to the introduction of new tools and products of the financial system, including cryptocurrencies, electronic and digital currencies.

The main stage in the further development of the market of cryptocurrencies and digital currencies can be considered the creation of national cryptocurrencies, which include the digital ruble. The advantages of national cryptocurrencies are in recognition by the state. This means that the national cryptocurrency will be absolutely legal tender and will offset the traditional fiat currency.

The digital ruble is money that the Bank of Russia plans to issue in digital form in addition to existing forms of money. The economically key innovation in the issuance of the digital ruble is not in a new monetary form, but in expanding the direct access of economic agents to the obligations of the Central Bank.

The digital ruble is money that the Bank of Russia plans to issue in digital form in addition to existing forms of money. The economically key innovation in the issuance of the digital ruble is not in a new monetary form, but in expanding the direct access of economic agents to the obligations of the Central Bank:

1. Increasing the level of distrust of money market participants in the central banking regulator, which will have more regulatory functions.

2. The creation of a digital ruble, its emission and financial control will require an increase in state budget expenditures by maintaining the regulatory functions of the Central Bank.

3. The creation of a digital ruble will lead to competitive pressure on innovative financial products, including banking innovation.

4. Due to the introduction of the digital ruble, an increase in the money supply in the country's economy is likely, which will allow borrowers to leave the bank lending market.

In our opinion, the introduction of the digital ruble by the Bank of Russia can be considered a significant factor that will affect the conditions of the state's fiscal policy. In particular, this is due to the fact that in connection with the introduction of the digital ruble, there will be fundamental changes in the calculation of the money supply in the country's economy.

In turn, the fiscal policy of any state is an important tool for implementing a unified state economic policy. The main goal is to ensure the stability of the economy in the country through the sustainable functioning of the national financial system.

Since the goal of the fiscal policy of the Bank of Russia is to regulate cash flows, then, accordingly, regulation of financial transactions, transfers and payment actions of market participants carried out using the digital ruble will be added to its duties.

One of the threats to the Russian financial system is that due to the emergence of the digital ruble and its active use by the population, part of the money can be transferred from bank accounts to electronic wallets. Accordingly, there will be a large outflow of financial resources from the country's banks, which will lead to a shortage of liquidity.

The Central Bank of the Russian Federation, as a banking regulator, will have to apply tools to ease its fiscal policy, which include:

- decreasing the level of the financial capital reserves by credit institutions;
- reducing the discount rate of refinancing;
- purchasing government securities in the financial markets by the banking regulator.

On the one hand, the Bank of Russia will be able to increase the level of liquidity of the banking sector's assets, which will protect it from the financial crisis. On the other hand, easing monetary policy could lead to such significant consequences for the Russian economy as:

- growing inflation;
- devaluation of the Russian ruble.

As a result, a spiral of financial crisis will occur, followed by a fall in GDP and a rise in unemployment.

In order for the Bank of Russia to secure the domestic national financial system from such consequences of the introduction of the digital ruble, it is possible to use the mechanism of restrictions on transactions with the digital ruble. From the point of view of those who are against such an event, an argument is made that due to such actions of the Central Bank, the development of infrastructure to support the digital currency system in Russia will be disrupted.

However, summing up the reasoning, it is important to draw the following conclusion: since the main goal of the regulatory activity of the Central Bank of the Russian Federation is to ensure the sustainability and stability of the national economic and financial system, measures to introduce limits on transactions with the digital ruble are an appropriate solution. This will eliminate the threat of loss of liquidity in the banking sector, which could lead to serious consequences — a stagflation trend in the Russian economy.

Analyzing foreign experience, it is worth noting that China turned out to be the most successful in issuing a digital national currency. Initially, the creation of a digital currency in this country was associated with the need to avoid settlements in US dollars. For the People's Bank of China, a digital yuan is a tool to fight for leadership in the financial sector [16].

In our opinion, the digital ruble can also become a tool for protecting the national financial system in this difficult period of Russia's international relations under the economic sanctions imposed by Western countries. Thus, in order to prevent the financial crisis of the Russian state, it seems possible to create and freely circulate a digital ruble.

CONCLUSIONS

Summing up the results of the study, the following conclusions can be drawn:

1. Considering the trends and factors in the development of the financial system of the national economy of modern Russia, it is necessary to analyze information that reveals the state of the budgetary, banking system, stock market and insurance industry, and the monetary policy of the state.

2. The main reasons for the Russian state budget deficit in recent years are the negative impact of the COVID-19 pandemic and economic sanctions. To eliminate the problem of the budget deficit, it is necessary to apply a mechanism for improving the budget policy, including the method of optimizing the revenues and expenditures of the state budget.

3. The development of the banking system of modern Russia is characterized by the positive results of its activities. Thanks

to credit organizations and commercial banks, a financial infrastructure is being created in the country that provides the population with consumer loans to maintain the size of aggregate demand, and business structures with commercial loans that allow them to finance the production activities of organizations and increase their production. Accordingly, the banking system is a key element contributing to the development of the financial system of the modern Russian economy.

4. In Russia, there are a number of urgent problems in the development of investment funds, on the activities of which the stock market depends:

low level of information transparency of mutual investment funds about their profitability and risks in the course of their main commercial activities;

lack of a mechanism that guarantees the protection of the financial capital of investors and depositors of mutual investment funds;

low level of development of exchange-traded investment funds, which can become an alternative to open-end investment funds and, at least, take second place in NAV after open-end investment funds, exceeding the performance of closed-end investment funds and interval investment funds.

5. The insurance industry of the modern Russian economy has a steady growth in market capitalization. It faces such urgent problems as a decrease in the volume of the compensation fund, an increase in the factor of judicial practice in the investigation of insurance claims, and an increase in the number of cases of fraud and economic crimes.

There are a number of factors stimulating the development of the insurance industry, including an increase in market demand for life insurance products among the population; the emergence of new health insurance products; technological development of the insurance services market; development of new sales channels; optimization of pricing policy for insurance products; enhancing the product penetration strategy per customer.

6. The digital transformation of the modern Russian economy leads to the introduction of new instruments and products of the financial system, including the digital ruble. It has both a positive and a negative impact on the trends and prospects for the development of the financial system of the Russian Federation. The introduction of the digital ruble will increase the money supply in the country's economy, there may be an outflow of financial resources from banks to electronic wallets, which will lead to a shortage of liquidity, etc.

The theoretical and practical significance of this paper lies in the study of factors affecting the efficiency of the functioning of the key elements of the national financial system of the modern Russian economy. The results obtained can be applied in the framework of further scientific research to develop recommendations for eliminating the main problems of the financial system. The practical value of the study lies in the proposal of specific measures aimed at improving the financial, budget and economic policy of the state authorities of Russia.

REFERENCES

1. Taranova I.V., Podkolzina I.M., Prokhorova V.V., Kolomyts O.N., Kobozeva E.M. Global financial and economic crisis in Russia: Trends and prospects. *Research Journal of Pharmaceutical, Biological and Chemical Sciences*. 2018;9(6):769–775. URL: https://www.rjpbcs.com/pdf/2018_9(6)/[106].pdf
2. Lankina Yu.S., Serdyukova L.O. Comparative analysis of foreign financial systems with the financial system of Russia. In: Socio-humanitarian and economic dimensions of modern society: Coll. sci. pap. Saratov: KUBiK; 2017;123–128. (In Russ.).
3. Balyan V.V. External and internal threats to the financial system of Russia. In: Integration processes in science in modern conditions: Proc. Int. sci.-pract. conference (in 3 pts.). Pt. 1. Ufa: Aeterna; 2018:20–22. URL: https://aeterna-ufa.ru/sbornik/NK-213–1.pdf (accessed on 30.04.2022). (In Russ.).

4. Ibragimova E.S., Khadzhimuratov M.U. The financial system of Russia, problems and solutions. In: Legal science in the 21st century: Coll. sci. pap. 2nd Int. round table. Moscow: Konvert; 2018:169–173. (In Russ.).
5. Mylnikova L.A. Problems of development of Russia's financial system. *International Journal of Advanced Studies*. 2018;8(3–2):138–143. (In Russ.).
6. Kashirina M.V. Prospects of development of budgetary system of the Russian Federation. *Azimut nauchnykh issledovanii: ekonomika i upravlenie = ASR: Economics and Management (Azimuth of Scientific Research)*. 2017;6(4):101–103. (In Russ.).
7. Kudrin A., Knobel A. Fiscal policy as a source of economic growth. *Voprosy ekonomiki*. 2017;(10):5–26. (In Russ.). DOI: 10.32609/0042–8736–2017–10–5–26
8. Gvarliani T.E., Oborin M.S., Saryan A.A. Problems of formation of budget expenditures in the conditions of deficit budgets. *Servis plus = Service Plus*. 2017;11(4):65–77. (In Russ.). DOI: 10.22412/1993–7768–11–4–7
9. Vasyunina M.L. Optimization of budget expenditures and financial support to state and municipal services. *Finansy i kredit = Finance and Credit*. 2016;(6):14–25. (In Russ.).
10. Drobyshevskaya M.N., Kulyakova M.N. Problems and prospects for the development of the modern Russian banking system. *Nauchno-metodicheskii elektronnyi zhurnal "Kontsept" = Scientific and Methodological Electronic Journal "Concept"*. 2017;(T18):62–65. URL: http://e-koncept.ru/2017/770385.htm (accessed on 30.04.2022). (In Russ.).
11. Lopatin E. Assessment of Russian banking system performance and sustainability. *Banks and Bank Systems*. 2019;14(3):202–211. DOI: 10.21511/bbs.14(3).2019.17
12. Belotelova N., Kukharenko O., Gizyatova A. Modernization of the Russian banking system as a necessary condition for the development of the Russian economy. Economic and social development: Proc. 25th Int. sci. conf. on economic and social development — 17th Int. social congr. (ISC-2017). (Moscow, 30–31 October 2017). Varazdin: Varazdin Development and Entrepreneurship Agency; 2017:413–421. URL: https://www.esd-conference.com/upload/book_of_proceedings/Book_of_Proceedings_esdMoscow_2017_Final_online.pdf (accessed on 30.04.2022).
13. Ono S. Financial development and economic growth nexus in Russia. *Russian Journal of Economics*. 2017;3(3):321–332. DOI: 10.1016/j.ruje.2017.09.006
14. Vishnever V. Ya., Kuskov V.M., Nedorezova E.S. Trends and prospects for the development of mutual investment funds in the collective investment system. *Ekonomicheskie nauki = Economic Sciences*. 2019;(178):19–24. (In Russ.). DOI: 10.14451/1.178.19
15. Meda M.A. Actual problems of civil liability insurance of vehicle owners. *Molodoi uchenyi*. 2019;(13):199–200. (In Russ.).
16. Vershinina O.V., Labusheva Ya.G., Sultaniev I.S. Analysis of the opportunities and risks of introducing digital currencies of central banks on the example of the "digital ruble". *Vestnik Rossiiskogo novogo universiteta. Seriya: Chelovek i obshchestvo = Vestnik of Russian New University. Series: Man and Society*. 2021;(1):51–60. (In Russ.). DOI: 10.25586/RNU.V9276.21.01.P.051

ABOUT THE AUTHOR

Gayratbek N. Ismoilov — 2nd year graduate student, St. Petersburg State University, St. Petersburg, Russia
http://orcid.org/0000-0002-5668-0119
ismoilov.gn@gmail.com

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 26.04.2021; revised on 12.06.2021 and accepted for publication on 27.12.2021.
The author read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-33-49
UDC 336.7(045)
JEL G21, G22

Russian Bancassurance Perspectives: From Path Dependence to Ecosystem

Yu.S. Evlakhova
Rostov State University of Economics, Rostov-on-Don, Russia;
Southern University (IMBL), Rostov-on-Don, Russia

ABSTRACT

The **aim** of the study is to identify the future of Russian bancassurance in the "digital" world based on the formation of a reasonable opinion on its prospects within the framework of path dependence and the ecosystem approach. The **relevance** of the article is determined by the significance of bancassurance for the development of the insurance sector and the Russian financial market as a whole. The author uses **methods** such as analysis and synthesis, analogy, induction, descriptive analysis, matrix construction, clustering using graphs, and index method. The study **is based** on the works of domestic and foreign authors, data from the Bank of Russia on the state of the banking sector of the Russian Federation and key indicators of the insurers; reports of the rating agency "RA Expert" on bancassurance; data of Professional Analysis Systems of Markets and Companies Interfax. Based on the results of quantitative analysis, the author **concludes** that the key factor determining the development of Russian bancassurance within the framework of path dependence is the role of banks as recipients of cash flows. The author **shows** that the digital transformation of bancassurance is mainly associated with the digitalization of intermediary relations between the bank, individual client and insurance company. To quantify the level of such transformation, the author proposes indices of digitalization of bancassurance. The author shows that bancassurance can be successfully built into ecosystems due to its peculiarities. However, the conditions for the development of ecosystems in the Russian market can lead to the dissolution of bancassurance in ecosystems and the loss of its significance for the development of the Russian financial market. The presented conclusions are to be discussed taking into account the study of **new issues** in this problem area, including a study of the mergers and acquisitions market to analyse changes in the ownership structure of banks and insurance companies; a study of corporate bancassurance; broader analysis of quantitative values of bancassurance due to new data; analysis of ecosystem development in the Russian financial market in order to move from predictive judgments to a discussion of real situations.

Keywords: the financial market; digitalization; integration of financial services; credit institutions; insurance companies; ownership structure; digitalization index; Insurtech; platform business model; competition

For citation: **Evlakhova Yu.S. Russian bancassurance perspectives: From path dependence to ecosystem.** ***Finance: Theory and Practice.*** **2022;26(3):33-49. DOI: 10.26794/2587-5671-2022-26-3-33-49**

INTRODUCTION

Today, bancassurance is playing a significant role in the development of the Russian insurance sector (accounting for 43.1% of its volume[1] in 2019) and in the income of credit institutions (accounting for 46.7% of banks' commission income and premiums as of 01.01.2019). Over the past ten years (2010–2019), the volume of the Russian bancassurance market has grown 6.8 times, reaching 638 billion rubles in 2019.[2] Although the rapid growth was interrupted for objective reasons (economic and social consequences of the pandemic), its prospects remain, at least in the short term, taking into account the path dependence factor. Based on the trajectory of the previous development, the future of Russian bancassurance lies in an increase in the volume of the bancassurance services market, the accompanying strengthening of the relationship between credit and insurance organizations, and even, according to some estimates [1], in the complete control of banks over the insurance market and the transformation of the latter from competitive to oligopolistic.

At the same time, digitalization has a powerful transformative effect on bancassurance, as well as other segments of the financial market, which is expressed in the emergence of digital financial innovations, such as virtual currencies and asset tokenization, the spread of end-to-end technologies (blockchain, open API, big data, etc.) in the formation of ecosystems. The fintech factor radically changes both the external environment for the functioning of credit and insurance organizations and their internal processes, products and technologies. Given that the genesis of bancassurance is due to natural inter-industry interaction in the financial market, and the scale of the client base is a key success factor for the interaction between banks and insurance companies, the phenomenon of bancassurance fits well into the ecosystem approach. At the same time, the digital transformation of the economy and especially the development of ecosystems can significantly affect the future of bancassurance, changing the trajectory of its development in such a way that it will lose its independence as an economic phenomenon and lose its significance for the development of the Russian financial market as a whole and its insurance and banking sectors in particular. Thus, determining the future of Russian bancassurance in the digital world and ecosystems is an urgent scientific task.

A review of the literature on the prospects for the development of bancassurance made it possible to note the key provisions of the theory and practice of the development of bancassurance, indicating that the development of this phenomenon in foreign and domestic studies is considered a consistent process, due to the previous specifics of development, the purpose of which is the integration of financial services.

In foreign studies, the prospects for the development of bancassurance are considered as due to national characteristics and the duration of its development period. Analysis of bancassurance practices in 28 advanced economies and emerging market economies in the 2000s showed that the factors influencing its development include: the size of a financial institution, its expenses and incomes, the size of the national banking sector, the level of financial deregulation in the country and the level of inflation [2].

From a chronological point of view, bancassurance has the longest history in European countries (since the 1970s), which allowed researchers to accumulate enough data for empirical analysis. In particular, options for the ownership structure of bancassurance companies in 1998–2012 were studied and the impact of financial crises on them. Based on this new information was obtained on the positive role of bancassurance in the development of the financial market due to the effect of economies of scale [3].

[1] Bank insurance market in 2019 and forecast for 2020: A premonition of a crisis. RA Expert Report. May 2020. URL: https://raexpert.ru/researches/insurance/bancassurance_2020/ (accessed on 05.04.2021).

[2] Ibidem.

In Asian countries, this concept appeared only in 2002 [4]. Bancassurance has been actively developed in China, India [5], South Korea and Taiwan [6]. The researchers explain the dynamism of the spread of bancassurance in Asian countries primarily by the cultural characteristics of the financial behavior of the population, due to the strong influence of collectivism [7]. In China, the prospects for the development of bancassurance are associated with ensuring social stability with its help. Thus, the purpose of forming various models of bancassurance companies with "Chinese" characteristics is to revitalize China's rural areas and develop small and micro enterprises [8].

In the United States, the peculiarities of financial market regulation led to the fact that the conditions for the emergence of bancassurance existed for a fairly short period of time: from 1999 after the adoption of the Gramm-Leach-Bliley Act [9] and until the adoption of the Dodd-Frank Act in 2010, which did not contribute to the rapid development of bancassurance.

In Russia, despite the legal ban on banks to engage in insurance activities, bancassurance has been actively developing since the 2000s. In 2004, the first scientific articles on bank insurance appeared, and in 2006, the first study by the rating agency "RA Expert" was published, devoted to the analysis of the interaction between banks and insurance companies.[3]

The general prospect for the development of bancassurance for different countries was the formation of universal financial markets, the spread of companies — financial conglomerates, the rejection of industry and functional division in the sale of financial products and the provision of financial services.

A new factor influencing the development of bancassurance is the digital transformation of the economy. Studies show that the digitalization of bancassurance depends both on the level of penetration of digital technologies into the business processes of banks and insurance companies [10], and on the availability and scale of the state program for digitalization of the national financial industry [11]. At the same time, in the scientific literature, the digital transformation of bancassurance is assessed in terms of the concept of path dependence, i.e. the significance of the factor is recognized, and the technological changes caused by it are taken into account, but in general, the development trajectory is assumed to be unchanged.

Unlike existing approaches, we believe that the fintech factor and the ecosystem approach are breakthrough innovations for bancassurance, i.e. innovations that can "destroy" the existing bancassurance and give impetus to the emergence of a new phenomenon.

The aim of the article is to predict the future of Russian bancassurance in the digital world and ecosystems based on the formation of an informed opinion about its prospects within the framework of the path dependence and ecosystem approach.

To achieve this goal, the following tasks are set and solved in the article:

1) to characterize the current state of Russian bancassurance based on the formalization of its structure, taking into account the diversity of types of relationships between banks and insurance companies;

2) to identify areas of influence of the "fintech" factor on bancassurance and develop digitalization indices for bancassurance;

3) to determine the positions of bancassurance within the framework of the ecosystem approach.

The paper uses various methods of analysis in accordance with their capabilities and adequacy to the subject of analysis. With the help of descriptive data analysis, the indicators of interaction between credit and insurance organizations were studied. The construction of matrices and clustering using graphs are used to analyze the ownership relations of banks and insurance companies.

[3] Rating Agency "RA Expert". URL: https://raex-a.ru/project/bankstrah/2006/conference (accessed on 15.04.2021).

The index method was used to develop the bancassurance digitalization indices.

In quantitative analysis of bancassurance, one should distinguish between data that is published at the macro level for the banking and insurance sectors as a whole, and data that is collected at the micro level, the level of individual banks and insurance companies. Macroeconomic, banking and insurance statistics provide a general picture of the interaction between banks and insurance companies, and microeconomic data allow a detailed analysis of the relationship between these financial institutions. At the same time, only macro data are available in scientific circulation, which turns the analysis of microeconomic data on bank insurance into a promising, but so far difficult to implement this study.

The article uses official data of the Bank of Russia on the state of the banking sector of the Russian Federation and key performance indicators of insurers as sources of information; reports of the rating agency "RA Expert" on bancassurance; data and tools of the Professional Analysis Systems of Markets and Companies Interfax.

RESULTS AND DISCUSSION

Formalized structure of Russian bancassurance

In our opinion, bancassurance is a system of economic relations between credit and insurance organizations, including various levels of interaction.

The interaction of credit and insurance organizations is carried out at various levels, which includes: the provision of services that are mandatory for participants; customer relations; intermediary (agency) relations; relations within financial and industrial groups, financial groups, or conglomerates [12]. In general, we support this system, considering it unnecessary to single out the level of interaction associated with the provision of services for compulsory insurance of deposits of individuals in banks. In our opinion, the structure of bancassurance includes three levels of relations and can be formalized as follows (*Table 1*).

The presented three levels (types) of relations between credit and insurance organizations can be disclosed through various economic indicators, the analysis of the values of which makes it possible to holistically characterize the current state of Russian bancassurance.

Quantitative characteristics of Russian bancassurance

Given the differences in the nature of the financial services provided, banks and insurance companies are each other's clients. *Fig. 1* shows the dynamics of individual indicators characterizing the client relations of Russian banks and insurance companies.

According to *Fig. 1,* the value of the contribution of insurance companies as clients of banks exceeds the contribution of banks as clients of insurers. Indeed, over the past 5 years, insurance organizations have increased the volume of bank deposits in absolute terms by 1.6 times — up to 640.2 billion rubles (at the same time, the share of deposits in the structure of insurers' assets slightly decreased from 24.7% in 2015 to 22% in 2019). At the same time, the volume of premiums from the insurance of banks' own risks did not change significantly in 2019 compared to 2015, remaining at the level of 11–12 billion rubles, which was due to the influence of various factors, including the desire of credit institutions to optimize non-production expenses.

The next level of bancassurance is intermediary relations, in which the bank acts as an intermediary between the client of the insurance company and the insurer. Theoretically, in such a system of relations, the interests of all participants are taken into account: insurance companies receive an additional channel for the sale of insurance products, banks receive the opportunity to receive additional (commission) income, and customers receive time and money savings. However, studies point to problems in the implementation of intermediary relationships. Insurance companies face conflicts between different sales channels for

Table 1

Structure of bancassurance by levels of interaction between banks and insurance companies

No.	Type of relationship	Characteristic
1	Client relations	*Insurance companies – clients of banks*: 1. Transactional services, including cash management services, centralized treasury solutions, card solutions and acquiring, data transmission channels, customer service and support. 2. Placement of funds on deposits, etc.
		Banks – clients of insurance companies: 1. Insurance of specific banking risks. 2. Bank property insurance (real estate, vehicle fleet, etc.). 3. Insurance of bank employees (voluntary health insurance, liability insurance, etc.)
2	Intermediary relations	1. *Bank as an intermediary between an individual and an insurance company (retail)*: 1.1. Insurance related to lending (life insurance of the borrower, property insurance – collateral, etc.). 1.2. Insurance not related to lending (investment, savings). 2. *Bank as an intermediary between a legal entity and an insurance company (corporate segment)*
3	Ownership relations	1. Banks directly or indirectly participate in the capital of insurance companies. 2. Insurance companies directly or indirectly participate in the capital of banks.

Source: compiled by the author.

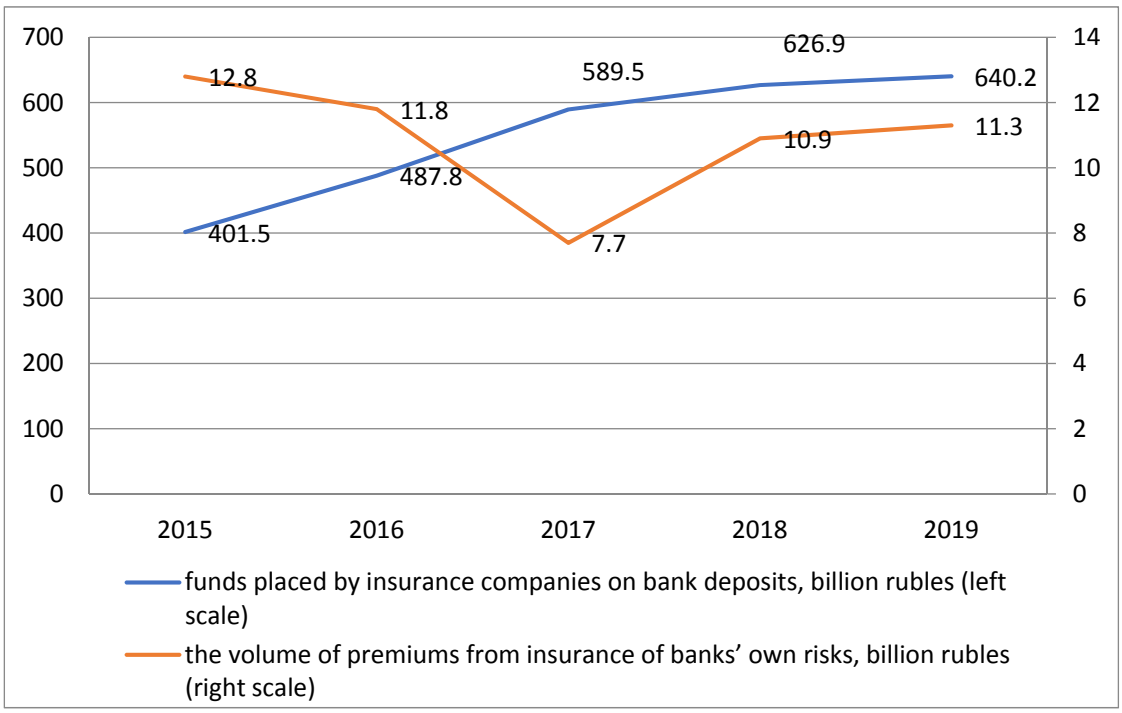


Fig. 1. **Quantitative characteristics of client relations between Russian banks and insurance companies in 2015–2019**

Source: calculated and compiled by the author.

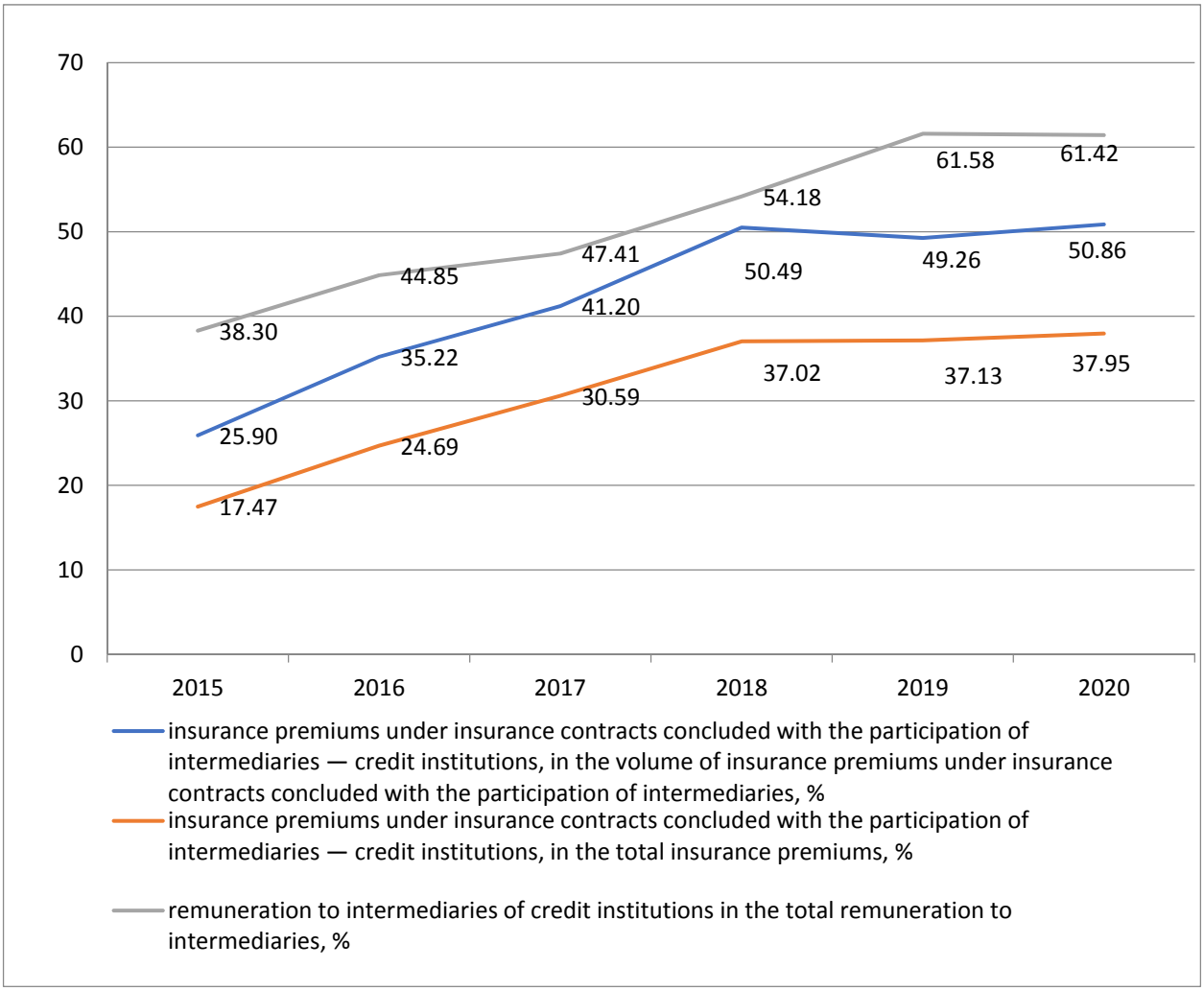


Fig. 2. **Dynamics of indicators characterizing the role of banks as intermediaries of insurance companies in 2015–2020**

Source: calculated and compiled by the author based on Statistics for the review of key performance indicators of insurers. URL: http://www.cbr.ru/insurance/reporting_stat

insurance products [13]. For banks, income from bancassurance is negligible compared to total income [14]. Despite the fact that customers receive financial services at a lower cost due to bancassurance [15], the question of their satisfaction with the quality of services provided is acute [16]. In Russia, the provision of insurance services to the population by banks under agency programs is accompanied by a distortion of the economic meaning of insurance, since financial organizations are focused not on providing insurance protection, but on generating income with ultra-low payments [17, 18].

To analyse intermediary relations between banks and insurance companies, based on the key performance indicators of insurers compiled by the Central Bank of the Russian Federation, indicators characterizing the role of banks as intermediaries in the sale of insurance products can be used. The indicators characterizing the role of banks in attracting insurance premiums, as well as the role of banks (in terms of insurance premiums) among all intermediaries involved by insurers, are presented in *Fig. 2*. These indicators can be supplemented with similar indicators for insurance payments that cannot be calculated at present due to the lack of detailed statistics on insurance payments.

Fig. 2 presents the dynamics of indicators characterizing intermediary relations between banks and insurance companies.

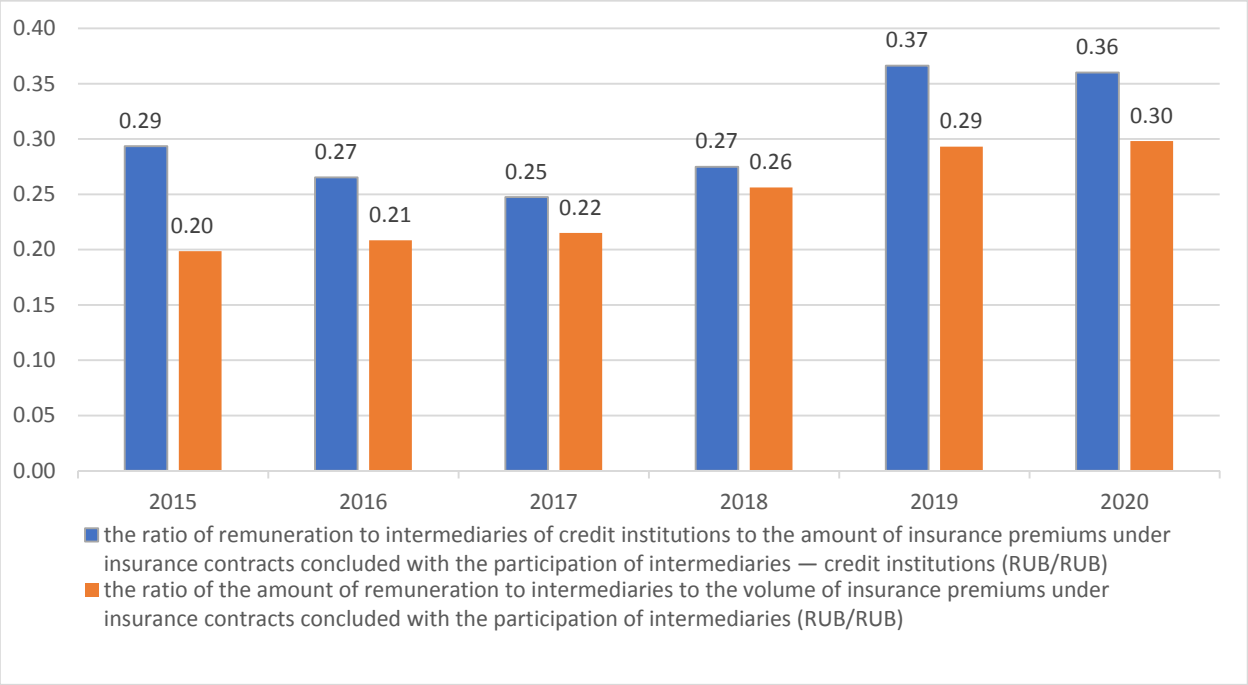


Fig. 3. **Dynamics of indicators characterizing the ratio of remuneration to intermediaries and the volume of insurance premiums collected by them in 2015–2020**

Source: calculated and compiled by the author based on Statistics for the review of key performance indicators of insurers. URL: http://www.cbr.ru/insurance/reporting_stat (accessed on 15.04.2021).

Fig. 2 indicates a steady increase in the values of indicators characterizing the role of banks as intermediaries in the provision of insurance services.

In 2020, credit institutions accounted for 50.86% of insurance premiums under insurance contracts concluded with the participation of intermediaries. Compared to 2015, the growth was 96%, i.e. the volume of insurance premiums collected with the participation of banks almost doubled. In the total volume of insurance premiums under insurance contracts concluded both by insurers independently and with the participation of intermediaries, the share of credit institutions in 2020 was 37.95%, compared to 2015, the growth was 117%. This confirms the fact that the banking sales channel of insurance products has not only become predominant among other intermediaries (such as insurance brokers, car dealers, individuals, etc.) but also competes with the independent conclusion of contracts by insurers, including online. Based on the difference in growth rates, it can be concluded that the second process is proceeding at a faster pace. In turn, this suggests that insurers have not fully used the potential of online interaction with their customers.

It is logical that following the increase in the share of insurance premiums under insurance contracts concluded with the participation of credit institutions, there is an increase in the share of their remuneration in the total amount of remuneration to intermediaries. The growth rate of remuneration to credit institutions in 2020 compared to 2015 was 60%. In more detail, the ratio of remuneration to banks and the amount of insurance premiums for insurance contracts concluded by them is shown in *Fig. 3*.

According to *Fig. 3,* in 2020 for 1 ruble of insurance premiums under insurance contracts concluded with the participation of credit institutions, the remuneration accounted for 36 kopecks, i.e. more than a third. A significant increase in remuneration to banks occurred in 2019 (0.37 in 2019 against 0.27 in 2018). Apparently, the record growth in remuneration was due to the fact

Table 2

Matrix of property structure relations between insurance companies and systemically important banks of the Russian Federation

	Gazprombank	VTB Bank	Sberbank	Otkritie FC Bank	Rosbank	Promsvyazbank	Russian Agricultural Bank	Credit Bank of Moscow
Sberbank Strahovanie Zhizni (Sberbank Life Insurance)			1					
Ingosstrah	2							3
RESO-Garantiya		3				3	3	
VSK				3				
Rosgosstrah				1				
Sberbank Strahovanie (Sberbank Insurance)			1					
Sos'ete Zheneral' Strahovanie Zhizni (Rosbank Insurance)					1			
Rosgosstrah Zhizn (Rosgosstrah Life)				1				

Source: compiled by the author.

Note: 1 — participation of the bank in the capital of the insurance company/insurance companies; 2 — participation of the insurance company in the capital of the bank; 3 — indirect relations by ownership structure.

that the types of insurance, which had a significant weight in the overall structure of bancassurance, showed steady growth in 2019, namely: the volume of insurance of borrowers on consumer loans, the share in the total structure is 26.2%, an increase of 16.1%, while the volume of mortgage insurance, which accounts for 13%, increased by 18.5%.[4] At the same time, the presented values are average indicators summarizing the values by types of insurance. According to a survey of the Federal Antimonopoly Service of Credit and Insurance Organizations, the range of commissions for banks is from 6 to 94% for personal insurance, and from 15 to 80% for property insurance.[5]

In general, for all types of intermediaries in 2020, for 1 ruble of the insurance premium under insurance contracts concluded with the participation of intermediaries the remuneration accounted for 30 kopecks (and unlike banks, the growth of this indicator in 2015–2019 is "smooth", without significant peaks). Thus, along with the fact that banks as intermediaries prevail among all sales channels of insurance products, for insurers the cost of their intermediation is above the average level (if we consider the level

[4] The bancassurance market in 2019 and the forecast for 2020: A premonition of a crisis. RA Expert Report. May 2020. URL: https://raexpert.ru/researches/insurance/bancassurance_2020/ (accessed on 05.04.2021).

[5] FAS Russia. Improvement of interaction between credit and insurance organizations. 19.12.2018. URL: https://fas.gov.ru/spheres/3?type=presentation (accessed on 13.10.2021).

of remuneration of all intermediaries as an average value). This conclusion does not contradict the provision that "banks, offering insurance products to retail borrowers, pursue the goal of obtaining excess commission income from insurance companies" [17].

Unlike other studies, for example [17, 19], in our article, we do not analyse data on investment life insurance and life insurance of borrowers, because these are special cases of intermediary relations that are usually taken into account in our indicators.

The interaction between banks and insurance companies is accompanied by a synergistic effect [20], which can be enhanced by choosing the optimal ownership structure [21]. Significant efficiency in the organization of ownership relations between credit and insurance organizations shows the model of a financial holding [4].

Let us consider the types of links by ownership structure between Russian banks and insurance companies. As a result of the analysis of ownership relations between systemically important banks and the largest insurance companies (in terms of insurance premiums), it was found that there are links between 8 systemically important banks and 8 largest insurers (as of 2019). Among the systematically important banks, UniCredit Bank, Raiffeisenbank, Alfa-Bank have no connection with the insurance companies in terms of ownership structure.

For credit and insurance organizations that have ties in terms of ownership structure, we have built a relationship matrix (*Table 2*), which reflects the three types of relationships identified in the course of the analysis: 1) bank participation in the capital of an insurance organization /insurance companies; 2) participation of the insurance company in the capital of the bank; 3) indirect connections by ownership structure (through a chain of companies or through a common nominal shareholder of a bank and an insurance company).

We consider it necessary to clarify the indirect links in the ownership structure. According to SPARK-Interfax, VSK is connected with SAFMAR Financial Investments on the principle of mutual ownership of each other's shares: VSK owns 1.71% of SAFMAR Financial Investments and, conversely, SAFMAR Financial Investments owns 49% of VSK shares. In addition, both companies share the same board members. In turn, 7.67% of the shares of SAFMAR Financial Investments are owned by Trust Bank, whose members of the Board of Directors are also members of the Board of Directors of FC Otkritie. Thus, it can be argued that, through a chain of companies, VSK has connections in terms of ownership structure with FC Otkritie.

As for the ownership structure of the insurance company RESO-Garantiya, according to SPARK-Interfax, its co-owner is Rossiya Airlines, 75% of which, in turn, belongs to Aeroflot. 51.17% of Aeroflot shares are owned by the Federal Property Management Agency, which also owns stakes in such systemically important banks as Russian Agricultural Bank (100%), Promsvyazbank PJSC (100%), VTB Bank (77.47%). Thus, RESO-Garantiya through a chain of companies is included in the pool of organizations, the main owner of which is the state structure.

As indirect links, we include the presence of a common nominal shareholder: Ingosstrah and Moscow Credit Bank have a common nominal shareholder — the National Settlement Depository (NSD).

To complete the picture, historical connections should be mentioned that are not reflected in *Table 2*. Until 2019, VTB Bank owned VTB Strahovanie (VTB Insurance), as well as VTB Strahovanie Zhizni (VTB Life Insurance). Until 2019, UNS-Holding, a subsidiary of Alfabank, owned Alfastrakhovanie (Alfabank Insurance).

For insurance companies that are subsidiaries of banks, as well as indirectly related by ownership structure, we carried out clustering. Despite the small number of objects that can be subjected to clustering, we consider it important to carry it out, since this is an additional, previously not presented in the study method for analyzing the

ownership relations of banks and insurance companies, which can be a promising research method with an increase in the number of insurance companies that are subsidiaries of Russian banks. Due to limited sampling, graph clustering was chosen (more complex clustering algorithms require a larger sample size).

As features that characterize an insurance company as an object of clustering, we identified: the share of a bank in the ownership of an insurance company (x_1), the share of deposits in credit institutions in the total assets of an insurance company (x_2). Distances between pairs of objects were calculated in the Euclidean metric (*Table 3*).

The clustering algorithm consists of the fact that an arbitrary number R is given as input, after which all edges in the graph whose metric is greater than R are removed. Accordingly, the regions (connectivity components) into which the graph is divided are clusters. In our case, R values from 0.2 to 0.9 were sequentially numbered in increments of 0.1. However, for any given value of R, the objects remained in the same cluster.

Thus, the result of clustering allows us to conclude that objects are similar — insurance organizations that are subsidiaries of banks, as well as indirectly related to each other in terms of ownership, despite differences in the ownership shares of banks and the share of bank deposits in the asset structure.

The results of the analysis of ownership relations between Russian banks and insurance companies confirm the conclusions of the study [1] about the strengthening of banks' control over the Russian insurance market. However, in our opinion, in order to obtain more accurate conclusions, the market for mergers and acquisitions of credit and insurance organizations should be further studied.

It should be noted that the formation of insurance groups in the Russian market, i.e. groups of insurance companies connected with each other by ownership relations. An example (the only one so far) is AO SOGAZ, which owns OOO SOGAZ Life and VTB Insurance (also owns VTB Life Insurance). We believe that the formation and development of insurance groups have a positive effect on the Russian insurance market, although it does not prevent its oligopolization trend. The fact is that the strong influence of the banking sector on the insurance market has its pros and cons, the latter can be attributed to the "domino effect" in the transfer of risks. In this sense, the presence of insurance groups independent of banks means greater resilience of the insurance market to financial

Table 3

Clustering insurance companies with graphs: initial data

Objects	Feature values	
	x_1	x_2
A	0.19	0.098
B	0.23	0.26
C	1	0.085
D	0.49	0.148
Edges	**Metrics (lengths) of edges**	
AB	0.167	
AC	0.81	
AD	0.3	
BC	0.79	
BD	0.283	
CD	0.514	

Source: calculated by the author according to the SPARK-Interfax system (x_1), according to the financial statements (x_2). URL: https://www.vsk.ru/upload/documents/4/352/doc/VSK__2019_AUDIT_RPRT.pdf; https://sberbank-nsurance.ru/upload/23/238f754d72475b6bb48c1374f242f794.pdf https://www.rgs.ru/upload/iblock/11e/osbu-_-rgs-_-2019_12.pdf; https://rosbankinsurance.ru/documents/AZ_Otchetnost_SZHSZH_2019.pdf (accessed on 13.10.2021).

Note: A — Sos'ete Zheneral' Strahovanie Zhizni (Rosbank Insurance); B — Rosgosstrah; C — Sberbank strahovanie (Sberbank Insurance); D — VSK.

and macroeconomic shocks that can be transmitted by credit institutions.

Thus, the current state of Russian bancassurance is characterized by the following. In client relations, the contribution of insurers as clients of credit institutions in value terms exceeds the contribution of banks as clients of insurance companies. In intermediary relations, the banking channel for the sale of insurance products is not only predominant among other intermediaries (such as insurance brokers, car dealers, individuals, etc.) but also competes with the independent conclusion of contracts by insurers, including online, while for insurers the cost their mediation is above average. In terms of ownership structure, 8 out of 12 systemically important banks and 8 out of 20 largest insurance companies have direct or indirect links, while clustering using graphs showed the similarity of objects — insurance companies that are subsidiary banks, as well as indirectly related to each other in terms of ownership, despite on differences in the shares of ownership of banks and the shares of bank deposits in the structure of assets. In general, if we evaluate the overall balance of cash flows between Russian credit and insurance companies, then the recipients of this balance are banks, since insurance companies share their financial resources with them by placing assets on deposits and provide them with opportunities to receive additional income. Thus, the role of banks as recipients in Russian bancassurance is, in our opinion, a key factor in path dependence.

Russian bancassurance and the fintech factor

A holistic and detailed characteristic of the influence of the fintech factor on Russian bancassurance is beyond the scope of the research tasks. We note the obvious: the main direction of digitalization of finance is the interaction of financial institutions with clients — individuals. Therefore, predominantly "digital" changes in bancassurance relate to intermediary relations.

Thus, the main directions of Insurtech (the process of introducing digital technologies into the insurance sector) are the development of insurance for risks and events occurring in the life cycle of an insurance object (person, property, business), as well as the use of digital technologies in the business processes of an insurance company. As reflected in [22], Russian insurers have achieved maximum digitalization in the organization of sales of insurance products and minimum in risk management of the insurer itself, while IIoT and big data are most in demand by the types of technologies used, and virtual and augmented reality, and blockchain — least of all.

The digital transformation of the banking sector is so diverse that there is no single term for it (for example, BankTech). Using the KPMG[6] global fintech market segmentation criteria, we note that the digitalization of the banking sector is actively taking place in the following areas: payment services; personal finance and wealth management (WealthTech); risk management and cybersecurity. CreditTech (digital technologies used in lending to individuals and legal entities) and OperTech (digital technologies for optimizing operational processes) can also be added to these areas [23].

It is important that it is the digital transformation of the banking sector that in many cases is carried out through partnership with fintech companies, and not only through competition with them, as, for example, indicated in [24]. Moreover, such cooperation was and remains necessary, since for fintech companies interaction with credit institutions is one of the main mechanisms for their involvement in the financial market.

The typology of interaction between banks and fintech companies includes the following. Firstly, the partnership of companies on various terms, which, as a rule, involves the sharing of risks and volume of investments in new technology or a new product (an example of a partnership is the interaction between the largest Swiss financial holding UBS Group

[6] Pulse of fintech: half 2 2020. KPMG. February 2021. URL: https://home.kpmg/xx/en/home/insights/2020/02/pulse-of-fintech-archive.html (accessed on 15.03.2021).

AG and the blockchain company Clearmatics). Secondly, the formation of funds by banks for investing in fintech companies, which makes it possible to separate the risks of traditional banking and the risks of investing in technological innovations. Thirdly, the so-called internal developments, i.e. the creation by banks of their own fintech divisions, which eventually develop into independent companies (for example, Goldman Sachs Group, one of the world's largest investment banks and bitcoin startup Circle).

As for Russia, according to Deloitte,[7] which considered data from 15 largest Russian banks, in terms of the level of digitalization of banks by stages of the customer journey, Russian banks, although they are not among the world leaders, exceed the world average in most cases (by 5 out of 6 assessed stages of interaction between the bank and the client, except the stage "the first steps of the client").

Thus, the existing estimates of the digitalization of the insurance and banking sectors of the Russian financial market indicate a high intensity of processes. At the same time, estimates of the digitalization of bancassurance are not presented in the scientific literature.

A quantitative assessment of the influence of the "fintech" factor can be given by the digitalization index, which is usually understood as some indicator that reflects the depth (level) of penetration of digital technologies into the object under study. The national economy, as well as the region, industry (field of activity, business) and a separate enterprise, can be considered as an object. Various information and empirical bases are used — macro data of financial organizations, data from sociological surveys of the population or surveys of financial organizations.

Based on the search results, we identified the following Russian digitalization indices related to the activities of credit and insurance organizations:

1) financial sector digitalization index,[8] which characterizes of use of broadband Internet, cloud services, RFID technologies, and ERP systems;

2) the level of digital maturity of banks,[9] which assesses the state of digital retail channels in three components (functionality, customer needs, user experience);

3) the level of digitalization of the insurance market [22], shows the ratio of business processes using digital technologies to the total number of business processes in an insurance company;

4) the digital insured index,[10] shows the level of digitalization of products and services of insurance companies for individual clients.

A comparative analysis of these indices based on publicly available and open information showed that the indices differ in the set of initial data, and their information base was formed mainly through surveys and questionnaires (of experts, consumers of financial services, and market participants).

In the development of the idea of a quantitative assessment of the influence of the fintech factor, we have developed indices of digitalization of bancassurance (in particular, intermediary relations between banks, individual clients and insurance companies). The indices are based on the use of some digital technologies, such as broadband Internet, cloud services and ERP systems, since this allows the use of official data as an information base, rather than survey data, and thus makes it possible to study changes in indices over time, which means — monitor the digitalization of bancassurance.

We offer the following bancassurance digitalization indices (similar to stock indices):

[7] Digital Banking Maturity. URL: https://www2.deloitte.com/content/dam/Deloitte/ru/Documents/research-center/DBM_2020_rus.pdf (accessed on 28.04.2021).

[8] Digital Economy Indicators: 2020: Statistical Collection. G.I. Abdrakhmanova, K.O. Vishnevsky, L.M. Gohberg, and others; National Research University "Higher School of Economics". Moscow: NRU HSE; 2020.

[9] Digital Banking Maturity. URL: https://www2.deloitte.com/content/dam/Deloitte/ru/Documents/research-center/DBM_2020_rus.pdf (accessed on 28.04.2021).

[10] The All-Russian Union of Insurers and AS&M present the Digital Insured Index. URL: https://www.insur-info.ru/press/161309/ (accessed on 28.04.2021).

1) the "broad market" index, which involves determining the average value of three indicators, each of which reflects the share of credit institutions using a certain technology (from among those indicated above) to provide insurance services to the population under agency programs, to the total number of credit institutions connected by intermediary relations with insurers;

2) an index that takes into account ownership relations, which involves determining the average value of three indicators, each of which reflects the share of credit institutions using a certain technology (from among those indicated above) to provide insurance services to the population under agency programs, to the total number of credit institutions associated by ownership relations with insurance organizations.

Thus, the proposed indices of digitalization of bancassurance will make it possible to determine the level of penetration of digital technologies into the interaction of credit and insurance organizations, and comparing them with each other will make it possible to study the impact of ownership relations on the digitalization of bancassurance. At the same time, it should be recognized that the proposed indices have certain limitations: 1) fragmentation (indices do not cover all levels of bancassurance, do not include all possible digital technologies); 2) disproportionate (did not reflect the indices of digitalization of insurance companies).

Nevertheless, at the conceptual level, the proposed indices demonstrate the possibility of analyzing the level of digitalization of bancassurance as a complex phenomenon, including the interaction of various organizations with each other. We believe that prospects of scientific research may be the analysis of the quantitative values of the indices of bancassurance digitalization, their dynamics, as well as the inclusion of the indices in the existing indices of the development of bancassurance, which will make it possible to form a broader and more accurate description of bancassurance in the digital world.

Prospects of bancassurance: An ecosystem approach

Bancassurance has long been considered within the framework of the concepts of "financial supermarket" and financial conglomerate [25]. In the context of the digital transformation of the economy, these concepts are giving way to an ecosystem approach, which involves the transition from the traditional to the "platform business model that uses ... both technological and behavioral changes".[11] Bancassurance, as an example of natural cross-sector interaction in the financial market and relying on the scale of the client base as a key success factor, successfully fits into the ecosystem approach. In our opinion, bancassurance can be adapted to different types of ecosystems. For ecosystems that develop services to meet a variety of human needs, the principles of everyday finance (lifestyle banking + everyday insurance), are applicable, i.e. restructuring of banking and insurance to meet the needs of the life cycle (person, property, business) based on digital technologies. Ecosystems built around one or more basic needs may include the provision of banking and insurance products that accompany a certain basic need, which is also one of the ways to develop bancassurance.

At the same time, the development of ecosystems is accompanied not only by a decrease in the transaction costs of participants [26] but also by a change in the methods and types of competition. The competition involves traditional and platform business models; platforms compete with each other and with offline partners, and there is internal competition between members of the same ecosystem [27].

In the existing global and Russian ecosystems, financial services operate according to a closed model, which implies the importance of having products or services in the ecosystem perimeter,

[11] Ecosystems: Approaches to regulation. Report of the Bank of Russia for public consultations. April 2021. URL: https://www.cbr.ru/Content/Document/File/119960/Consultation_Paper_02042021.pdf (accessed on 25.05.2021).

rather than a wide choice of them, and, as a result, leads to an insignificant level of internal competition. Therefore, in future, bancassurance, having become part of the ecosystem, will naturally be built into the line of financial services offered and will lose its independence. We believe this scenario is likely for both bank-based and technology-based ecosystems.

For the Russian financial market, ecosystems are seen as a disruptive innovation, i.e. an innovation with strong market power that will lead to a structural transformation of the entire market. The task of the Bank of Russia as a regulator is to regulate the process of such a transformation in order to obtain an optimal market structure, which is understood as "the presence of several large national ecosystems that compete with each other and with foreign players, niche suppliers and smaller platforms that satisfy the demand of customers outside the ecosystems and challenging ecosystem leaders.[12] In this respect, the regulator identifies, on the one hand, an increase in the risks of traditional companies due to increased competition with ecosystems (both price and related to the market niche), and on the other hand, the likelihood of a radical reduction in the number of credit and insurance organizations with traditional business models. We believe that in combination with the spread of a closed model for financial services in ecosystems, both the risk of reducing the number of financial institutions and the level of business riskiness of the remaining organizations will increase even more. Perhaps this may cause an increase in mergers and acquisitions of credit and insurance organizations in the financial market in order to increase market share and resist ecosystems. The question remains open: What can compete on equal terms with ecosystems in the financial market: ecosystems, financial (bank and insurance) groups, banking or insurance groups? The prospect of bancassurance in companies outside the perimeter of ecosystems depends on the answer to this question.

Thus, bancassurance, due to its characteristics, can be successfully integrated into ecosystems, both based on banks and technology companies. At the same time, an insignificant level of intra-platform competition, as well as a non-zero probability of a reduction in the number of credit and insurance organizations with traditional business models due to increased competition with ecosystems, may lead to the fact that bancassurance will lose its independence and significance for the development of the Russian financial market.

CONCLUSIONS

Thus, the characterization of the current state of Russian bancassurance, carried out by us on the basis of a formalization of its structure, taking into account the variety of types of relations between banks and insurance companies, made it possible to determine that the key factor determining its development within the previous trajectory is the role of banks as recipients of cash flows, despite the fact that bancassurance does not play such a significant role for the functioning and development of the banking sector as it does for the insurance sector.

The digital transformation of bancassurance is mainly associated with the digitalization of intermediary relations between a bank, an individual client, and an insurance company, primarily because the main area of digitalization of finance is the interaction of financial institutions with individual clients. A quantitative assessment of the digitalization of bancassurance can be obtained using the corresponding digitalization indices, proposals for the construction of which at the conceptual level are formulated in this paper.

Due to its characteristics, bancassurance can be successfully integrated into ecosystems, both based on banks and technology

[12] Ecosystems: Approaches to regulation. Report of the Bank of Russia for public consultations. April 2021. URL: https://www.cbr.ru/Content/Document/File/119960/Consultation_Paper_02042021.pdf (accessed on 25.05.2021).

companies. However, the conditions for the development of ecosystems in the Russian market (first of all, the functioning of financial services according to a closed model, the involvement of financial organizations with traditional and platform business models in unequal competition, in which the latter have clear advantages) can lead to the fact that bancassurance, having dissolved into ecosystems, will lose its independence and significance for the development of the Russian financial market.

At the same time, the presented conclusions about the prospects for the development of Russian bancassurance are subject to discussion and further adjustment, taking into account the study of new issues in this problem area. Among the further prospects for work on the topic, we believe, can be indicated:

1) research on the mergers and acquisitions market in order to analyze changes in the ownership structure of banks and insurance companies;

2) research into intermediary relationships in corporate bancassurance, which, we believe, is less prone to involvement in the platform business model and, therefore, can maintain its independence;

3) expanding the quantitative analysis of bancassurance through:

a) involvement of micro-level data, and not only macroeconomic statistics; b) collection of detailed statistics on insurance payments; c) analysis of the quantitative values of the bancassurance digitalization indices, and their dynamics, which together will make it possible to form a broader and more accurate description of bancassurance in the digital world;

4) studying the dynamics and indicators of the development of ecosystems in the Russian financial market in order to move from forecasting judgments to discussing real situations.

The theoretical significance of the study is that it develops approaches to the study of bancassurance and analyses the prospects for its development in the context of digitalization and the ecosystem approach. Using the example of bancassurance, the study shows how a cross-sectoral phenomenon that has emerged during the integration of various financial services can develop in the context of the integration of financial and non-financial services on a digital platform.

The practical significance of the results of the study lies in the possibility of using the proposed bancassurance digitalization indices to more accurately characterize its state in the digital world. The presented opinion on the prospects for the development of bancassurance in the digital world and ecosystems can be useful for managers and specialists of credit and insurance organizations in the framework of strategic planning and forecasting of their activities.

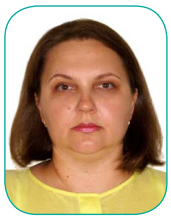
REFERENCES

1. Tarasova Yu.A., Khokhlova E.S. The degree of influence of the banking sector on the insurance market. *EKO: vserossiiskii ekonomicheskii zhurnal = ECO Journal*. 2019;(12):154–169. (In Russ.). DOI: 10.30680/ECO0131–7652–2019–12–154–169
2. Chen Z., Li D., Liao L., Moshirian F., Szablocs C. Expansion and consolidation of bancassurance in the 21st century. *Journal of International Financial Markets, Institutions and Money*. 2009;19(4):633–644. DOI: 10.1016/j.intfin.2008.10.002
3. Dreassi A., Schneider M.C. Bancassurance and scale economies: Evidence from Italy. *The Geneva Papers on Risk and Insurance — Issues and Practice*. 2015;40(1):89–107. DOI: 10.1057/gpp.2014.13
4. Wu C.-R., Lin C.-T., Lin Y.-F. Selecting the preferable bancassurance alliance strategic by using expert group decision technique. *Expert Systems with Applications*. 2009;36(2.Pt.2):3623–3629. DOI: 10.1016/j.eswa.2008.02.016
5. Harmanpreet K., Shikha M. Bancassurance: A step towards financial inclusion. *South Asian Journal of Marketing & Management Research*. 2017;7(6):43–52. DOI: 10.5958/2249–877X.2017.00036.4

6. Peng J.-L., Jeng V., Wang J. L., Chen Y.-C. The impact of bancassurance on efficiency and profitability of banks: Evidence from the banking industry in Taiwan. *Journal of Banking & Finance*. 2017;80:1–13. DOI: 10.1016/j.jbankfin.2017.03.013
7. Hong J., Lee Y-i. Bancassurance in East Asia: Cultural impact on customers' cross-buying behaviour. *Journal of Financial Services Marketing*. 2014;19(3):234–247. DOI: 10.1057/fsm.2014.16
8. Qian M., Gong X., Zheng Y., Wan F. Research on the social sustainability of China's bancassurance. *Revista de Cercetare şi Intervenţie Socială*. 2020;68:179–187. DOI: 10.33788/rcis.68.12
9. Ricci O. Studying the bancassurance phenomenon: A literature review. In: Fiordelisi F., Ricci O., eds. Bancassurance in Europe: Past, present and future. Basingstoke, New York: Palgrave Macmillan; 2012:139–178. (Palgrave Macmillan Studies in Banking and Financial Institutions). DOI: 10.1057/9780230358287_7
10. Rusin V. N. Features of digitalization in bank and insurance sectors of Russia. *Ekonomika. Biznes. Banki = Economy. Business. Banks.* 2019;10:37–49. (In Russ.).
11. Relan M. Digitalization in the bancassurance model of distribution — opportunities and challenges. *Bimaquest*. 2020;20(3):34–50. URL: http://bimaquest.niapune.org.in/index.php/bimaquest/article/view/79/58
12. Babenko I. V., Babenko A. I. Bancassurance as a financial category. Evaluation of the Russian bancassurance services market. *Finansy i kredit = Finance and Credit.* 2016;(4):11–23. (In Russ.).
13. Fan C. K., Lu W.-C. Analyze cyber-channel conflict while adopting bancassurance: A frequency perspective. *Journal of Applied Finance & Banking*. 2015;5(5):99–112.
14. Brar A. S., Singh S. Bancassurance: A study in relation to total income & branch network of private sector banks in India. *IOSR Journal of Business and Management*. 2016;2(2):43–47. DOI: 10.9790/487X-15010020243–47
15. Okeahalam C. C. Does bancassurance reduce the price of financial service products? *Journal of Financial Services Research*. 2008;33(3):147–162. DOI: 10.1007/s10693–008–0031-x
16. Choudhury M., Singh R. Identifying factors influencing customer experience in bancassurance: A literature review. *Journal of Commerce & Accounting Research*. 2021;10(2):10–22. URL: https://www.researchgate.net/publication/351049878_IDENTIFYING_FACTORS_INFLUENCING_CUSTOMER_EXPERIENCE_IN_BANCASSURANCE_A_LITERATURE_REVIEW
17. Ezrokh Yu. S. Russian bank insurance in the retail segment: Myths and reality. *Voprosy ekonomiki*. 2018;(5):92–109. (In Russ.). DOI: 10.32609/0042–8736–2018–5–92–109
18. Yusupova O. A. New algorithms of bank assurance transactions and the infringement of borrowers' rights. *Finansovaya analitika: problemy i resheniya = Financial Analytics: Science and Experience*. 2019;12(1):107–124. (In Russ.). DOI: 10.24891/fa.12.1.107
19. Dementiev N. P. Russian insurance companies: Questionable achievements of recent years, credibility issues. *EKO: vserossiiskii ekonomicheskii zhurnal = ECO Journal.* 2020;(5):77–98. (In Russ.). DOI: 10.30680/ECO0131–7652–2020–5–77–98
20. Fields L. P., Fraser D. R., Kolari J. W. Bidder returns in bancassurance mergers: Is there evidence of synergy? *Journal of Banking & Finance*. 2007;31(12):3646–3662. DOI: 10.1016/j.jbankfin.2007.01.014
21. Averchenko O. D. Bank insurance in ensuring the development of the financial sector of the economy. Moscow: Plekhanov Russian University of Economics; 2019. 212 p. (In Russ.).
22. Bryzgalov D. V., Gryzenkova J. V. Tsyganov A. A. Prospects for digitalization of the insurance business in Russia. *Finansovyi zhurnal =Financial Journal*. 2020;12(3):76–90. (In Russ.). DOI: 10.31107/2075–1990–2020–3–76–90
23. Khalimon E. A., Makeeva V. G., Kafiyatullina Yu.N., Kharchilava G. P. Identification and analysis of factors determining the readiness degree of the Russian Federation financial sector to the digitalization processes. *E-Management*. 2019;(4):74–84. (In Russ.). DOI: 10.26425/2658–3445–2019–4–74–84
24. Stulz R. M. FinTech, BigTech, and the future of banks. *Journal of Applied Corporate Finance*. 2022;34(1):106–117. DOI: 10.1111/jacf.12492

25. Alifanova E.N., Evlakhova Yu.S. On the convergence of financial market models. *Finansy i kredit = Finance and Credit.* 2007;(38):27–33. (In Russ.).
26. Kleiner G.B., Rybachuk M.A., Karpinskaya V.A. Development of ecosystems in the financial sector of Russia. *Upravlenets = The Manager.* 2020;11(4):2–15. (In Russ.). DOI: 10.29141/2218–5003–2020–11–4–1
27. Bauer V.P., Eremin V.V., Ryzhkova M.V. Digitalization of the financial activities of platform companies: Competitive potential and social impact. *Finance: Theory and Practice.* 2021;25(2):114–127. DOI: 10.26794/2587–5671–2021–25–2–114–127

ABOUT THE AUTHOR



Yuliya S. Evlakhova — Dr. Sci. (Econ.), Assoc. Prof., Prof., Financial Monitoring and Financial Markets Department, Rostov State University of Economics, Rostov-on-Don, Russia; Head of the Finance, Accounting and Taxation Department, Southern University (IMBL), Rostov-on-Don, Russia
https://orcid.org/0000-0002-2561-6165
evlahova@yandex.ru

Conflicts of Interest Statement: The author has no conflicts of interest to declare.

The article was submitted on 01.06.2021; revised on 14.06.2021 and accepted for publication on 27.12.2021.
The author read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-50-63
JEL G30, L25

Working Capital Management and Profitability Relationship: The Role of Macroeconomic Indicators

I. Yilmaz, G. Acar
Dhofar University, Salalah, Oman

ABSTRACT

The aim of this article is to search for the relationship between working capital management (WCM) and profitability under different macroeconomic conditions. The dataset includes 179 companies from Gulf Cooperation Council (GCC) countries for a six-year period. Results reveal a nonlinear relationship between CCC and profitability in a way that, shorter CCC periods have opposite direction with profitability yet as the period increases this relation turns to be positive. The partial effect of CCC has a more significant influence than those of its components. However, overall significances of the components are higher. Gross profit and operating profit margins have high responsiveness against WCM changes. Among macroeconomic indicators, inflation and unemployment are distinctive in models yet GDP per capita has a moderate influence on profitability. To the best of our knowledge, this is the first study to examine the WCM-profitability relationship with a focus on macroeconomic variables in the GCC region.

Keywords: cash conversion cycle (CCC); profitability; working capital management; macroeconomic indicators

For citation: **Yilmaz I., Acar G. Working capital management and profitability relationship: The role of macroeconomic indicators. *Finance: Theory and Practice.* 2022;26(3):50-63. DOI: 10.26794/2587-5671-2022-26-3-50-63**

INTRODUCTION

As the business environment becomes more competitive, the pressure on the firms to manage the resources as effectively and efficiently as possible increases. To gain and sustain a competitive advantage, the managers must focus on the margins and effective cost control. In addition to several operating cost items, financing costs may have a critical effect on the profitability and therefore long-term survival of the firm. In this context, the minimization of financing costs and making the best use of operating cash flows deserve a detailed examination and an appropriate policy development. Proper planning of collections from customers and payments for purchases and the synchronization between them may make a profound impact on the performance of the firm. Management of cash and cash equivalents, credit policy for the customers, and relationships with suppliers are all within this category. Even though long-term financial decisions are attached more importance, WCM is crucially important for the success or even the survival of the firms. Especially for the firms operating in an environment with limited sources of external finance or for the firms with their own credibility problems, the effective management of working capital becomes more important. The policies adopted for managing working affect and are affected by the operating activities of the company.

Mainly, it is aimed to investigate the relationship between WCM and firm performance. This relationship is reciprocal because an efficient WCM is expected to improve firm performance. Similarly, a firm with good financial performance is expected to have more flexibility in managing working capital accounts. Moreover, the liquidity-profitability trade-off makes them more interdependent. For instance, a firm aiming to boost sales needs to extend more trade credit to its customers by allowing a longer collection period. Due to the trade-off between profitability and liquidity, the firms may have an optimal working capital maximizing value [1]. Most of the previous studies analyzed and found a linear relationship between WCM and profitability [1–5]. However, due to the interdependence between working capital and performance, and several other factors which might potentially affect both of them, the relationship is expected to be

non-linear. There are some studies which found non-linear relationships [6–8]. The nature of the relationship between working capital and financial performance gives rise to potential endogeneity problems. For this reason, we use a generalized method of moments (GMM). By using the data for a sample of 179 companies from GCC countries for a six-year period of 2010–2015, we introduce new evidence on the working capital- financial performance relationship.

The article is structured as follows; the next Section makes a short review of related literature. Section 3 gives details about the data, sample and methodology used in the study. Section 4 provides analysis results and the findings. The last Section makes the conclusion.

LITERATURE REVIEW

In the literature of accounting and finance, one of the most studied topics is WCM, and a significant number of those studies aimed to figure out the relationship between WCM and profitability. Firms may adopt a conservative or an aggressive WCM policy, depending on the amount they invest in working capital accounts. The conservative policy includes the investment of increased amounts into inventories and receivables; providing some benefits such as lower risk of inventory shortages and increases in sales due to extended collection period for customers [1]. The aggressive policy requires the minimization of the amounts invested in those accounts. The lower investment in inventories decreases storage and insurance costs, the lower investment in receivables enables the firm to use those funds for other investment opportunities. [1, 4, 9]. However, both policies have some potential disadvantages and risks; an aggressive policy may cause interruptions in the operations, adversely affect sales due to shorter periods of trade credit for the customers, and a conservative policy may increase finance costs. Therefore, an aggressive policy may bring higher returns and higher risks, whereas a conservative policy may bring lower returns and lower risks.

The previous studies used several indicators to be the measure of WCM, one of the most used indicators is the cash conversion cycle (CCC). Richards and Laughlin [10] describe CCC as the net time interval between a firm's actual cash outflows for the purchase of productive resources and the cash inflows from sales of products. It is a financial metric used to evaluate the management's effectiveness and whether the firm needs external financing [11, 12]. It provides a better understanding of liquidity compared to static ratios calculated based on balance sheet data. Even though aggressive or conservative policies are dealt with the amounts of investment in working capital accounts, it is equally important to convert those invested amounts into cash. The inability to convert into cash may cause liquidity problems [13].

It is a commonly accepted argument that the aim of a firm's WCM is to minimize CCC, however, this should not be a blindly followed target and the related operating policies should not be affected negatively by the policy of WCM. In other words, while reducing the period of collections, relationships with customers must be considered and while extending the period of payments for purchases, relationships with suppliers must be considered. Those policies have an impact on the profitability of the firm. Previous studies reported different views on whether a shorter or a longer CCC is better. Some studies claim that a longer CCC affects profitability positively [14–17], while some other studies claim that a shorter CCC is better and it affects profitability positively [18–22]. The conflicting results might be attributed to several reasons such as country conditions, macroeconomic situation, and so on. The firms with a longer CCC would need more short-term, external financing in order to finance their operating cycle, such firms perform worse in financial crisis periods [23, 24]. Moreover, those results indicate that the relationship between WCM and firm profitability is affected by many factors and is a complex relationship. Even though many studies proposed a linear relationship, there are some studies based on the assumption that there is an optimal level of working capital, stating that the relationship is non-linear and there is an inverted U-shape [6–8].

DATA AND METHODOLOGY

The dataset includes 179 non-financial companies listed in GCC countries for the six-

year period of 2010–2015. *Table 1* presents the distribution of companies among countries. The majority of companies belong to Saudi Arabia and it is followed by Oman, UAE (United Arab Emirates), Kuwait, Bahrain and Qatar.

We constructed models for which profitability measures are taken as dependent variables and working capital measures as well as country specific and company specific control variables are considered as independent variables. Gross Profit Margin (GPM) and Operating Profit Margin (OPM) are selected as dependent variables. On the other side, Cash Conversion Cycle (CCC), and components of CCC which can be stated as Accounts Receivable Days (ADAY), Days Inventory Held (IDAY), and Accounts Payable Days (PDAY) are selected as independent variables. To control for company specific deviations, we added natural logarithm of total assets (A), sales growth (SG), Operating Cash Flows divided by Sales (OCF), Capital Expenditures divided by Total Assets (CPX), and gearing ratio — debts divided by total assets — (GEAR). To control for country specific deviations we added country dummies. These dummies are selected from macroeconomic indicators to control for country differences as well as macroeconomic changes in years. We selected unemployment rates (unmp), inflation rates (inf), and GDP per capita (gdppc). To avoid dummy variable trap, we added five country dummies and excluded Saudi Arabia dummy variables because Saudi Arabia has the biggest share of companies and it is worth being selected as a benchmark. We used Stata software to perform statistical functions.

We constructed 16 models including two dependent variables — GPM and OPM, which are shown as PR (Profitability Ratios) in models below. GPM stands for Gross Profit Margin and it is derived as the gross profit divided by sales, OPM represents Operating Profit Margin and is computed as Operating Profit divided by Sales. Each profitability measure has been regressed upon Cash Conversion Cycle (CCC) and components of CCC — Accounts Receivables Days (ADAY), Days Inventory Held (IDAY), and Account Payable Days (PDAY) — along with their second powers, separately. CCC and components are shown in the equations below as WCMI (Working — Capital Management indicators).

Table 1

Distribution of Companies among Countries

	Frequency	Percent
Bahrain	9	5.03
Kuwait	27	15.08
Oman	38	21.23
Qatar	7	3.91
Saudi Arabia	68	37.99
United Arab Emirates	30	16.76
Total	179	100%

Source: compiled by the authors based on sample data.

To control for company specific deviations, we added five different measures; Natural Logarithm of Total Assets (A) for controlling company size, Sales Growth (SG) for growth potential, gearing ratio (GEAR) for capital structure, Operating Cash Flows ratio (OCF) for cash inflow for sales, and Capital Expenditure ratio (CPX) for changes in fixed asset and investments. All these control variables are stated as COMPC in framework models.

A general framework of models is presented below:

Model 1:

$$PR_{i,t} = \beta_0 + \beta_1 PR_{i,t-1} + \beta_2 WCMI_{i,t} + \\ + \beta_3 WCMI_{i,t}^2 + \partial COMPC_{i,t} + \varepsilon_{i,t}$$

Model 2:

$$PR_{i,t} = \beta_0 + \beta_1 PR_{i,t-1} + \beta_2 WCMI_{i,t} + \\ + \beta_3 WCMI_{i,t}^2 + \partial COMPC_{i,t} + \gamma unmp_{i,t} + \varepsilon_{i,t}$$

Model 3:

$$PR_{i,t} = \beta_0 + \beta_1 PR_{i,t-1} + \beta_2 WCMI_{i,t} + \\ + \beta_3 WCMI_{i,t}^2 + \partial COMPC_{i,t} + \gamma \inf_{j,t} + \varepsilon_{i,t}$$

Model 4:

$$PR_{i,t} = \beta_0 + \beta_1 PR_{i,t-1} + \beta_2 WCMI_{i,t} + \\ + \beta_3 WCMI_{i,t}^2 + \partial COMPC_{i,t} + \gamma gdppc_{i,t} + \varepsilon_{i,t},$$

Table 2

Summary Statistics of Variables

Variable	Observation	Mean	Std. Dev.	Min	Max
GPM	1,074	0.240	0.237	−1.795	0.842
OPM	1,074	0.108	0.195	−2.041	0.640
CCC	1,074	113.8	101.3	−274.3	434.18
RDAY	1,074	84.2	61.6	0.6	444.04
IDAY	1,074	91.6	70.9	1.2	355.88
PDAY	1,074	62.0	47.5	0.7	409.15
A	1,074	19.54	1.64	15.55	25.23
GEAR	1,074	0.207	0.187	0.000	1.062
OCF	1,074	0.169	0.280	−4.580	1.913
CPX	1,074	0.053	0.069	0.000	0.747
SG	1,074	0.085	0.321	−0.808	5.940

Source: compiled by the authors based on statistical output.

where PR — Profitability Ratios (GPM, OPM); *WCMI* — Working Capital Management Indicators (CCC, or components of CCC — ADAY, IDAY, PDAY); *COMPC* — Company Specific Control Variables (A, Gear, OCF, CPX, SG); *unmp* — unemployment rate; *inf* — Inflation rate; *gdppc* — GDP per capital; *i* — company identifier; *t* — time.

We constructed four different sets of models in order to see the effects of various macroeconomic indicators (unemployment, inflation, GDP Per Capita, and plain model including none) on the profitability. We used dummy variables for these indicators in a way that each country's economic indicator changes in time. Hence, we not only control for country differences, but we have a chance to account for economic changes in time.

The dataset includes cross section and time dimensions so that we used panel data models. Because of the dynamic patterns of dependent variables, we added the first lag of each regressand to the model to control for endogeneity. Furthermore, we assumed a two-direction relation between working capital measures and profitability measures. Prior literature [7, 8] has suggested a bilateral relation between profitability and working capital. Hence, we used Arellano-Bond Dynamic Panel Model. Secondly, as documented in prior literature [8, 9, 21], we assume a non-linear relationship between working capital and profitability, and we tested whether the relationship between dependent and independent variables is non-linear and we added CCC 2 and squares of its components (ARD 2, DIH2 and, APD 2) to the models. We tested for homoskedasticity and our dataset shows a heteroskedastic pattern, therefore we used robust standard errors. We had tested for autocorrelation and they do not show any serial correlation in second and more order conditions. Multicollinearity has been tested and there is no evidence for multicollinearity among independent variables.

RESULTS AND DISCUSSION

In this section, we firstly present descriptive statistics and they are followed by the analysis of models, then we discuss the findings.

Summary statistic of variables in *Table 2* reveals that 24% of sales return as gross profit, around 11% as operating profit. Companies have an average of 114 days of CCC. Days sales outstanding is almost 22 days higher than account payable days and average inventory conversion takes around three months for companies. Besides, companies

Table 3

Correlation Coefficients among Variables

	GPM	OPM	CCC	RDAY	IDAY	PDAY	A	GEAR	OCF	CPX	SG
GPM	1										
OPM	*0.57	1									
CCC	*–0.15	0.03	1								
RDAY	*–0.23	*–0.12	*0.53	1							
IDAY	*0.08	*0.10	*0.78	*0.12	1						
PDAY	*0.14	*–0.08	*–0.28	*0.34	–0.02	1					
A	*0.07	*0.17	–0.06	*0.06	0.00	*0.20	1				
GEAR	–0.02	–0.03	0.04	*0.19	0.00	*0.16	*0.29	1			
OCF	*0.45	*0.55	*–0.06	*–0.15	0.02	–0.03	*0.07	*–0.07	1		
CPX	*0.10	*0.07	*–0.16	*–0.16	*–0.12	–0.06	–0.05	0.05	*0.21	1	
SG	0.05	*0.10	*–0.12	*–0.09	*–0.12	–0.04	0.03	*0.09	0.06	*0.07	1

Source: compiled by the author based on statistical output.

Note: * Significant at 5%.

finance 21% of their assets with long- and short-term debt as gear ratio is around 21%. OCF rate shows us 17% of sales return as net cash inflow to the company. The difference between OPM and OCF may come from the weight of non-cash expenses, such as depreciation and amortization, on operating profit. Companies invest around 5% of their assets as capital expenditures whereas, their growth in sales reaches to almost 9%.

Table 3 exhibits the correlations among variables. Starred numbers show a significant correlation at 5%. Results show that all dependent variables have a significant correlation with many of the explanatory variables. Secondly, we did not encounter any evidence for multicollinearity among independent variables because the coefficients are less than 0.9.

Firstly, we consider models with no economic indicators and country differences which are presented in Model 1. In this model, we assumed six regression models; two regressions dealing with profitability ratios and CCC, two regressions dealing with profitability ratios and components of CCC (ADAY, IDAY, PDAY).

The first two regressions between profitability and CCC are presented below:

1 —

$$GPM_{i,t} = \beta_0 + \beta_1 GPM_{i,t-1} + \beta_2 CCC_{i,t} + \beta_3 CCC_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \varepsilon_{i,t}$$

2 —

$$OPM_{i,t} = \beta_0 + \beta_1 OPM_{i,t-1} + \beta_2 CCC_{i,t} + \beta_3 CCC_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \varepsilon_{i,t}$$

The results reported in *Table 4* reveal that CCC has a strong impact on GPM and is followed by OPM. There is a negative significant relation between CCC and CCC^2 suggesting a parabolic relation between CCC and profitability. We can clearly see this in GPM and OPM. Overall significances are sufficient at 1% for both equations; however, equation 2 has higher explanatory power. Equation 1 has only one variable losing significance (GEAR).

Secondly, the relationship between components of CCC and profitability measures without any country differences are taken into account. The following two models are constructed and tested for this relationship.

3 —

$$GPM_{i,t} = \beta_0 + \beta_1 GPM_{i,t-1} + \beta_2 RDAY_{i,t} + \beta_3 RDAY_{i,t}^2 + \\ + \beta_4 IDAY_{i,t} + \beta_5 IDAY_{i,t}^2 + \beta_6 PDAY_{i,t} + \beta_7 PDAY_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \varepsilon_{i,t}$$

4 —

$$OPM_{i,t} = \beta_0 + \beta_1 OPM_{i,t-1} + \beta_2 RDAY_{i,t} + \beta_3 RDAY_{i,t}^2 + \\ + \beta_4 IDAY_{i,t} + \beta_5 IDAY_{i,t}^2 + \beta_6 PDAY_{i,t} + \beta_7 PDAY_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \varepsilon_{i,t}$$

Table 4

Regression Results of Equations 1–2

	Equation 1: GPM–CCC		Equation 2: OPM–CCC	
Values	**Coefficients**	**Std. Error**	**Coefficients**	**Std. Error**
Constant (β_0)	–1.0096	*0.5663	0.4059	0.6559
LAG PR	0.3419	**0.1424	0.2256	*0.1192
CCC	–0.0011	**0.0005	–0.0012	*0.0007
CCC^2	0.0000	***0.0000	0.0000	*0.0000
a	0.0648	**0.0292	–0.0132	0.0321
gear	–0.0667	0.0733	–0.0911	0.0698
ocf	0.0624	***0.0189	0.1034	***0.0270
cpx	–0.3111	***0.0767	–0.0240	0.0503
sg	0.0616	*0.0354	0.0643	0.0614
Significance	***48.79		***77.59	

Source: compiled by the authors based on statistical analyses.

Note: *** Significant at 1%, ** Significant at 5%, *Significant at 10%.

In the proceeding part of the analysis for the model with no country dummies, presented in *Table 5*, we did not observe any significant effect of individual components on profitability measures. Furthermore, overall significances show a non-systematic pattern compared to the first two equations. The significance of Equation 4 (164.43) is considerably higher than that of Equation 2 (77.59), yet for Equation 3, there is no evidence of a crucial difference among its counterpart (equations 1).

Unemployment is one of the main indicators of the economy of a country and as the companies' performance gets worse and capacity utilization needs to be decreased, they lay off employees. Therefore, the unemployment rate gives insight into the trend in the macroeconomic outlook. Hence, we constructed Model 2 for testing the effects of WCMI on profitability and for controlling country-wise differences using unemployment. We constructed four equations, including 2 profitability ratios as dependent variables and CCC with its components as independent variables. The following two equations are set for the relations between PR and CCC. The results are given in Table 6 for these equations.

5 —

$$GPM_{i,t} = \beta_0 + \beta_1 GPM_{i,t-1} + \beta_2 CCC_{i,t} + \beta_3 CCC_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \\ + \partial_5 sg_{i,t} + \gamma unmp_{j,t} + \varepsilon_{i,t}$$

6 —

$$OPM_{i,t} = \beta_0 + \beta_1 OPM_{i,t-1} + \beta_2 CCC_{i,t} + \beta_3 CCC_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \\ + \partial_5 sg_{i,t} + \gamma unmp_{j,t} + \varepsilon_{i,t}$$

The results for equation 5 show CCC and CCC^2 have significant effects on GPM with opposite signs. CCC has an inverse relation whereas its square is positively related. Among countries, Kuwait and UAE differ from others significantly. Unemployment has a significant negative effect on profitability in Kuwait and UAE compared to other countries. CCC and CCC^2 lack a significant effect on OPM, which can be seen in *Table 6* (Equation 6). Unlike equation 5 results, only UAE has a distinctive pattern among countries in case of the effect of unemployment on OPM.

Overall significances of equations in model 2 are higher than those of model 1 which has no dummies. As we compare the overall significances

Table 5

Regression Results of Equations 3–4

	Equation 3: GPM–CCC Components		Equation 4: OPM–CCC Components	
Values	**Coefficients**	**Std. Error**	**Coefficients**	**Std. Error**
Constant (β_0)	–0.7053	0.6424	0.4436	0.9957
LAG PR	0.2615	**0.1324	0.2549	***0.0887
RDAY	0.0003	0.0004	0.0000	0.0006
RDAY^2	0.0000	0.0000	0.0000	0.0000
IDAY	–0.0001	0.0008	0.0007	0.0015
IDAY^2	0.0000	0.0000	0.0000	0.0000
PDAY	0.0002	0.0006	0.0009	0.0007
PDAY^2	0.0000	0.0000	0.0000	**0.0000
a	0.0456	0.0342	–0.0217	0.0518
gear	–0.0842	0.0768	–0.1229	0.0902
ocf	0.0774	***0.0281	0.1101	***0.0232
cpx	–0.2431	**0.1109	–0.0068	0.0655
sg	0.0618	**0.0310	0.0653	*0.0378
Significance	***46.59		***164.43	

Source: compiled by the authors based on statistical analyses.

Note: *** Significant at 1%, ** Significant at 5%, *Significant at 10%.

Table 6

Regression Results of Equations 5–6

	Equation 5: GPM–CCC		Equation 6: OPM–CCC	
Values	**Coefficients**	**Std. Error**	**Coefficients**	**Std. Error**
Constant (β_0)	–0.9638	*0.5840	0.2162	0.8716
LAG PR	0.3893	***0.1455	0.2164	*0.1141
CCC	–0.0011	**0.0005	–0.0011	0.0008
CCC^2	0.0000	*0.0000	0.0000	0.0000
a	0.0631	**0.0302	–0.0026	0.0424
gear	–0.0711	0.0726	–0.0959	0.0688
ocf	0.0607	***0.0178	0.0997	***0.0231
cpx	–0.2911	***0.0835	–0.0241	0.0534
sg	0.0677	*0.0356	0.0456	0.0535
BAH	0.0253	0.2173	–0.3801	0.2905
KWT	–0.0168	**0.0075	0.0207	0.0193
OMN	–0.0036	0.0142	0.0142	0.0193
QTR	–0.0416	0.1922	0.1360	0.1272
UAE	–0.0328	*0.0191	0.0445	*0.0229
Significance	***63.82		***107.79	

Source: compiled by the authors based on statistical analyses.

Note: *** Significant at 1%, ** Significant at 5%, *Significant at 10%.

Table 7

Regression Results of Equations 7–8

	Equation 7: GPM–CCC Components		Equation 8: OPM–CCC Components	
Values	**Coefficients**	**Std. Error**	**Coefficients**	**Std. Error**
Constant (β_0)	−0.5662	0.6228	0.4635	1.0926
LAG PR	0.2971	***0.1282	0.2337	***0.0895
RDAY	0.0002	0.0004	−0.0002	0.0006
RDAY^2	0.0000	0.0000	0.0000	0.0000
IDAY	0.0000	0.0008	0.0010	0.0016
IDAY^2	0.0000	0.0000	0.0000	0.0000
PDAY	0.0003	0.0006	0.0009	0.0008
PDAY^2	0.0000	0.0000	0.0000	*0.0000
a	0.0392	0.0327	−0.0211	0.0567
gear	−0.0779	0.0636	−0.0969	0.0844
ocf	0.0756	***0.0264	0.1117	***0.0241
cpx	−0.2334	**0.1156	−0.0157	0.0671
sg	0.0695	**0.0300	0.0490	0.0426
BAH	0.0085	0.1834	−0.5029	0.3652
KWT	−0.0132	0.0092	0.0029	0.0154
OMN	−0.0062	0.0130	0.0052	0.0145
QTR	−0.0815	0.1812	0.1291	0.1504
UAE	−0.0498	***0.0180	−0.0731	*0.0431
Significance	***71.85		***164.43	

Source: compiled by the authors based on statistical analyses.

Note: *** Significant at 1%, ** Significant at 5%, *Significant at 10%.

of Equations 1 and 5, Equation 5 (63.82) has a substantial superiority over Equation 1 (48.79). We see a similar pattern in the comparison of the overall significance of OPM equations (Equation 2, 77.59 and Equation 6, 107.79).

The second part of model 2 takes components of CCC with their squares and unemployment rates of countries and we generated the following equations.

7 —

$$GPM_{i,t} = \beta_0 + \beta_1 GPM_{i,t-1} + \beta_2 RDAY_{i,t} + \beta_3 RDAY_{i,t}^2 + \\ + \beta_4 IDAY_{i,t} + \beta_5 IDAY_{i,t}^2 + \beta_6 PDAY_{i,t} + \beta_7 PDAY_{i,t}^2 + \partial_1 a_{i,t} + \\ + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \gamma unmp_{j,t} + \varepsilon_{i,t}$$

8 —

$$OPM_{i,t} = \beta_0 + \beta_1 OPM_{i,t-1} + \beta_2 RDAY_{i,t} + \beta_3 RDAY_{i,t}^2 \\ + \beta_4 IDAY_{i,t} + \beta_5 IDAY_{i,t}^2 + \beta_6 PDAY_{i,t} + \beta_7 PDAY_{i,t}^2 + \partial_1 a_{i,t} + \\ + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \gamma unmp_{j,t} + \varepsilon_{i,t}$$

As shown in *Table 7*; partial significances of CCC components do not satisfy at least a 10% level for GPM and OPM. Another minor exception is the effect of PDAY 2 on OPM. Country differences show almost the same pattern, yet UAE disperses significantly in estimation of GPM and OPM from others.

Overall significances in these equations (7–8) do not differ from those of Model 1 except for

Table 8

Regression Results of Equations 9–10

	Equation 9: GPM–CCC		Equation 10: OPM–CCC	
Values	**Coefficients**	**Std. Error**	**Coefficients**	**Std. Error**
Constant (β_0)	–0.8967	*0.5327	0.2999	0.6482
LAG PR	0.3582	**0.1625	0.2779	**0.1203
CCC	–0.0010	*0.0006	–0.0009	0.0005
CCC²	0.0000	*0.0000	0.0000	0.0000
a	0.0587	**0.0277	–0.0085	0.0321
gear	–0.0784	0.0696	–0.0846	0.0672
ocf	0.0609	***0.0180	0.1004	***0.0203
cpx	–0.3009	***0.0832	–0.0434	0.0482
sg	0.0686	0.0431	0.0651	0.0656
BAH	–0.0071	0.0057	–0.0035	0.0364
KWT	0.0062	0.0064	–0.0096	0.0124
OMN	–0.0045	0.0055	0.0000	0.0050
QTR	0.0008	0.0262	–0.0016	0.0193
UAE	0.0055	0.0036	0.0045	0.0047
Significance	***66.41		***116.38	

Source: compiled by the authors based on statistical analyses.

Note: *** Significant at 1%, ** Significant at 5%, *Significant at 10%.

the difference between Equation 3 (46.59) and Equation 7 (71.85).

In the third model, we added inflation as the primary control variable to distinguish country differences. Inflation is another important indicator of a country's economic situation and the changes in the inflation rate show us whether a country has a downturn or boost or stability. We set the following equations to see the impact of CCC on different profitability ratios.

9 —

$$GPM_{i,t} = \beta_0 + \beta_1 GPM_{i,t-1} + \beta_2 CCC_{i,t} + \beta_3 CCC_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \gamma \inf_{j,t} + \varepsilon_{i,t}$$

10 —

$$OPM_{i,t} = \beta_0 + \beta_1 OPM_{i,t-1} + \beta_2 CCC_{i,t} + \beta_3 CCC_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \gamma \inf_{j,t} + \varepsilon_{i,t}$$

Results in *Table 8* show that CCC has a nonlinear influence on GPM. CCC and its square have shown no significant influence on OPM. In equations 9–10, we did not observe any difference resulting from the inflation rates of countries. Overall significances of equations are higher than those of no dummy model, they are slightly higher than equations reported in the unemployment model.

The next analysis takes CCC components as independent variables and inflation as a differentiating variable among countries. This part includes the last two equations of the third model. The regression equations followed are in *Table 9* which shows the results of these equations.

11 —

$$GPM_{i,t} = \beta_0 + \beta_1 GPM_{i,t-1} + \beta_2 RDAY_{i,t} + \beta_3 RDAY_{i,t}^2 + \\ + \beta_4 IDAY_{i,t} + \beta_5 IDAY_{i,t}^2 + \beta_6 PDAY_{i,t} + \beta_7 PDAY_{i,t}^2 + \partial_1 a_{i,t} + \\ + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \gamma \inf_{j,t} + \varepsilon_{i,t}$$

12 —

$$OPM_{i,t} = \beta_0 + \beta_1 OPM_{i,t-1} + \beta_2 RDAY_{i,t} + \beta_3 RDAY_{i,t}^2 + \\ + \beta_4 IDAY_{i,t} + \beta_5 IDAY_{i,t}^2 + \beta_6 PDAY_{i,t} + \beta_7 PDAY_{i,t}^2 + \partial_1 a_{i,t} + \\ + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \gamma \inf_{j,t} + \varepsilon_{i,t}$$

Table 9

Regression Results of Equations 11–12

	Equation 11: GPM–CCC Components		Equation 12: OPM–CCC Components	
Values	**Coefficients**	**Std. Error**	**Coefficients**	**Std. Error**
Constant (β_0)	–0.5009	0.5569	0.6701	0.8598
LAG PR	0.2886	**0.1423	0.2345	**0.0947
RDAY	0.0004	0.0005	–0.0002	0.0004
RDAY^2	0.0000	0.0000	0.0000	0.0000
IDAY	0.0000	0.0008	0.0010	0.0011
IDAY^2	0.0000	0.0000	0.0000	0.0000
PDAY	0.0003	0.0005	0.0010	0.0007
PDAY^2	0.0000	0.0000	0.0000	**0.0000
a	0.0332	0.0297	–0.0341	0.0459
gear	–0.0877	0.0635	–0.0839	0.0740
ocf	0.0739	***0.0260	0.1063	***0.0181
cpx	–0.2332	**0.1092	–0.0117	0.0667
sg	0.0709	**0.0320	0.0538	0.0399
BAH	–0.0039	0.0057	0.0428	0.0554
KWT	0.0042	0.0072	–0.0071	0.0103
OM	–0.0039	0.0038	–0.0017	0.0048
QTR	0.0101	0.0177	–0.0220	0.0288
UAE	0.0087	0.0032	0.0133	*0.0069
Significance	***78.8		***285.5	

Source: compiled by the authors based on statistical analyses.

Note: *** Significant at 1%, ** Significant at 5%, *Significant at 10%.

These equations fail to have a significant effect of CCC components on dependent variables. A minor exemption is noticed in PDAY 2 on OPM with 5% significance. Further, we could not witness any significant difference among countries in terms of inflation rates. Only UAE differs from others with a 10% significance level in explaining OPM. This effect implies that an increase in inflation will increase OPM by around 1.33%, which is an expected result. Because inflation shows an increase in general price levels and it will inevitably affect sales and operating income.

Overall significances show the highest scores in these two equations. Particularly, Equation 12 shows a very high significance (285.5) making it distinct from its counterparts.

Model 4 takes GDP per Capita (gdppc) as a distinguishing measure among countries. Instead of total GDP, we took GDP per capita since it reduces variance among countries with higher population and countries with lower population. Secondly, by taking GDP per capita, we standardize the data, and it is a more precise measurement of economic development compared to total GDP. We constructed six equations; the first two (13–14) take CCC as an independent variable and the proceeding two (15–16) take components as regressors. The analyses of CCC and profitability relationship are presented in the following equations.

Table 10

Regression Results of Equations 13–14

	Equation 13: GPM–CCC		Equation 14: OPM–CCC	
Values	**Coefficients**	**Std. Error**	**Coefficients**	**Std. Error**
Constant (β_0)	–0.9578	0.5966	0.0440	0.5766
LAG PR	0.3286	**0.1349	0.2691	**0.1079
CCC	–0.0010	**0.0005	–0.0010	0.0007
CCC^2	0.0000	*0.0000	0.0000	*0.0000
a	0.0628	**0.0309	0.0021	0.0281
gear	–0.0626	0.0732	–0.1058	0.0748
ocf	0.0633	***0.0198	0.1018	***0.0219
cpx	–0.3100	***0.0815	–0.0602	0.0427
sg	0.0608	*0.0361	0.0649	0.0635
BAH	0.0000	0.0000	0.0000	0.0000
KWT	0.0000	0.0000	0.0000	*0.0000
OMN	0.0000	0.0000	0.0000	0.0000
QTR	0.0000	0.0000	0.0000	0.0000
UAE	0.0000	0.0000	0.0000	0.0000
Significance	***52.07		***111.22	

Source: compiled by the authors based on statistical analyses.

Note: *** Significant at 1%, ** Significant at 5%, *Significant at 10%.

13 —

$$GPM_{i,t} = \beta_0 + \beta_1 GPM_{i,t-1} + \beta_2 CCC_{i,t} + \beta_3 CCC_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \\ + \partial_5 sg_{i,t} + \gamma gdppc_{j,t} + \varepsilon_{i,t}$$

14 —

$$OPM_{i,t} = \beta_0 + \beta_1 OPM_{i,t-1} + \beta_2 CCC_{i,t} + \beta_3 CCC_{i,t}^2 + \\ + \partial_1 a_{i,t} + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \\ + \partial_5 sg_{i,t} + \gamma gdppc_{j,t} + \varepsilon_{i,t}$$

Results show a significant effect of CCC and its square on GPM, however, it lacks impact on OPM, and instead CCC^2 is significant on this variable (*Table 10*). We observe a non-linear relation referred from opposite signs of CCC and CCC^2. We observed no difference among countries in GPM, however, in OPM estimation Kuwait differs from others. Both relations are positive, implying that an increase in per capita GDP would boost OPM in these countries. Overall significances are relatively higher than those of model 1 however they are lower than models 2 and 3.

The second part of model 4 is composed of two equations testing the potential effect of CCC components on PR under different per capita GDP. The equations and results follow.

15 —

$$GPM_{i,t} = \beta_0 + \beta_1 GPM_{i,t-1} + \beta_2 RDAY_{i,t} + \beta_3 RDAY_{i,t}^2 + \\ + \beta_4 IDAY_{i,t} + \beta_5 IDAY_{i,t}^2 + \beta_6 PDAY_{i,t} + \beta_7 PDAY_{i,t}^2 + \partial_1 a_{i,t} + \\ + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \gamma gdppc_{j,t} + \varepsilon_{i,t}$$

16 —

$$OPM_{i,t} = \beta_0 + \beta_1 OPM_{i,t-1} + \beta_2 RDAY_{i,t} + \beta_3 RDAY_{i,t}^2 + \\ + \beta_4 IDAY_{i,t} + \beta_5 IDAY_{i,t}^2 + \beta_6 PDAY_{i,t} + \beta_7 PDAY_{i,t}^2 + \partial_1 a_{i,t} + \\ + \partial_2 gear_{i,t} + \partial_3 ocf_{i,t} + \partial_4 cpx_{i,t} + \partial_5 sg_{i,t} + \gamma gdppc_{j,t} + \varepsilon_{i,t}$$

Table 11

Regression Results of Equations 15–16

	Eq. 15: GPM–CCC Comp.		Eq.16: OPM–CCC Comp.	
Values	**Coefficients**	**Std. Error**	**Coefficients**	**Std. Error**
Constant (β_0)	–0.6040	0.6077	0.2626	0.8831
LAG PR	0.2503	*0.1353	0.2239	**0.0936
RDAY	0.0003	0.0004	0.0000	0.0005
RDAY^2	0.0000	0.0000	0.0000	0.0000
IDA	0.0000	0.0008	0.0007	0.0011
IDAY^2	0.0000	0.0000	0.0000	0.0000
PDAY	0.0004	0.0006	0.0005	0.0006
PDAY^2	0.0000	0.0000	0.0000	**0.0000
a	0.0410	0.0322	–0.0138	0.0457
gear	–0.0775	0.0685	–0.0852	0.0884
ocf	0.0794	***0.0296	0.1091	***0.0254
cpx	–0.2417	**0.1127	–0.0277	0.0675
sg	0.0635	*0.0344	0.0301	0.0342
BAH	0.0000	0.0000	0.0000	0.0000
KWT	0.0000	0.0000	0.0000	0.0000
OMN	0.0000	0.0000	0.0000	0.0000
QTR	0.0000	0.0000	0.0000	0.0000
UAE	0.0000	*0.0000	0.0000	0.0000
Significance	***51.52		***183.98	

Source: compiled by the authors based on statistical analyses.

Note: *** Significant at 1%, ** Significant at 5%, *Significant at 10%.

Table 11 demonstrates to what extent components of CCC affect profitability with given GDP per capita differences of countries. The components have no significant impact on profitability ratios. Only PDAY 2 has an impact on OPM. GDPs per capita of countries except UAE are indifferent. UAE deviates from the group in GPM tests.

CONCLUSION

This study aimed to shed light on the relationship between profitability and WCM indicators in changing macroeconomic conditions across time and countries. Four models are constructed and tested whether two profitability ratios are related to CCC or its components. Additionally, we selected three macroeconomic indicators and added them to the model as control variables to include country specific as well as period-wise differences among observations. We added company specific control variables to have more accurate results which are aimed to be unbiased. We used a dynamic panel data model as there is endogeneity in the independent variables.

CCC has a significant partial effect on profitability itself. Components of CCC generally lack partial significance. The components of CCC have no significant influence on profitability but their combination makes aggregation and thus CCC itself has a much more meaningful impact on dependent variables. It can be concluded that the companies must adopt a comprehensive working capital approach covering all components of CCC.

The responsiveness of GPM is more significant to CCC than that of OPM. In most equations, the partial significance of CCC and its square on GPM is higher than OPM.

Firstly, equations taking OPM as regressand bear the highest overall significance and they are followed by those with GPM. Second, the equations with components of CCC provide considerably higher overall significances than those with CCC itself in explaining OPM. For the equations taking GPM as a dependent variable, though CCC components have relatively higher overall significances, these differences are not sharp as OPM. The conflict between partial significances and overall significances in regressions shows us control variables have more power in explaining variations in OPM rather than GPM since OPM has lower partial significance but higher overall significances. Company specific variables are selected to control for sources of variations for each company. Among the company specific variables, OCF disperses from others in significant effect. Almost all regressions show the same pattern, especially in all profitability ratios.

One of the models has no economic indicator and the other three include one indicator namely, unemployment, inflation, and GDP per capita. Firstly, all models with economic indicators provide higher overall significances compared to the plain model. Among economic indicators, unemployment and inflation provide the best results in explaining variations among countries and changes in economic conditions in time. GDP per capita has a moderately significant effect on profitability. Saudi Arabia has the biggest number of companies in the dataset; hence we took Saudi Arabia as the benchmark and tested how other companies deviate from companies in Saudi Arabia. We noticed that UAE disperses significantly from Saudi Arabia and in some models; we see Kuwait differs. Oman, Qatar and Bahrain companies follow the same pattern as Saudi Arabia Companies. We verified the relationship between cash conversion cycle and profitability and our models present a non-linear relationship between profitability and cash conversion cycle since the signs of CCC in models are opposite whereas signs of CCC^2 are positive.

REFERENCES

1. Deloof M. Does working capital management affect profitability of Belgian firms? *Journal of Business Finance & Accounting*. 2003;30(3–4):573–588. DOI: 10.1111/1468–5957.00008
2. Shin H.H., Soenen L.A. Efficiency of working capital management and corporate profitability. *Financial Practice and Education*. 1998;8(1):37–45.
3. Garcia-Teruel P.J., Martinez-Solano P. Effects of working capital management on SME profitability. *International Journal of Managerial Finance*. 2007;3(2):164–177. DOI: 10.1108/17439130710738718
4. Bhatia S., Srivastava A. Working capital management and firm performance in emerging economies: Evidence from India. *Management and Labour Studies*. 2016;41(2):71–87. DOI: 10.1177/0258042X16658733
5. Le B. Working capital management and firm's valuation, profitability and risk: Evidence from a developing market. *International Journal of Managerial Finance*. 2019;15(2):191–204. DOI: 10.1108/IJMF-01–2018–0012
6. Baños-Caballero S., García-Teruel P.J., Martínez-Solano P. How does working capital management affect the profitability of Spanish SMEs? *Small Business Economics*. 2012;39(2):517–529. DOI: 10.1007/s11187–011–9317–8
7. Singhania M., Mehta P. Working capital management and firms' profitability: Evidence from emerging Asian countries. *South Asian Journal of Business Studies*. 2017;6(1):80–97. DOI: 10.1108/SAJBS-09–2015–0060
8. Altaf N., Shah F.A. How does working capital management affect the profitability of Indian companies? *Journal of Advances in Management Research*. 2018;15(3):347–366. DOI: 10.1108/JAMR-06–2017–0076
9. Nazir M.S., Afza T. Working capital requirements and the determining factors in Pakistan. *The IUP Journal of Applied Finance*. 2009;15(4):28–38.
10. Richards V.D., Laughlin E.J. A cash conversion cycle approach to liquidity analysis. *Financial Management*. 1980;9(1):32–38. DOI: 10.2307/3665310
11. Raddatz C. Liquidity needs and vulnerability to financial underdevelopment. *Journal of Financial Economics*. 2006;80(3):677–722. DOI: 10.1016/j.jfineco.2005.03.012

12. Tong H., Wei S.-J. The composition matters: Capital inflows and liquidity crunch during a global economic crisis. *The Review of Financial Studies*. 2011;24(6):2023–2052. DOI: 10.3386/w15207
13. Dhole S., Mishra S., Pal A.M. Efficient working capital management, financial constraints and firm value: A text-based analysis. *Pacific-Basin Finance Journal*. 2019;58:101212. DOI: 10.1016/j.pacfin.2019.101212
14. Wilner B.S. The exploitation of relationships in financial distress: The case of trade credit. *The Journal of Finance*. 2000;55(1):153–178. DOI: 10.1111/0022–1082.00203
15. Padachi K. Trends in working capital management and its impact on firms' performance: An analysis of Mauritian small manufacturing firms. *International Review of Business Research Papers*. 2006;2(2):45–58.
16. Sharma A.K., Kumar S. Effect of working capital management on firm profitability: Empirical evidence from India. *Global Business Review*. 2011;12(1):159–173. DOI: 10.1177/097215091001200110
17. Bhuni A., Das A. Underlying relationship between working capital management and profitability of pharmaceutical companies in India. *American Journal of Theoretical and Applied Business*. 2015;1(1):27–36. DOI: 10.11648/j.ajtab.20150101.15
18. Nobanee H., Abdullatif M., AlHajjar M. Cash conversion cycle and firm's performance of Japanese firms. *Asian Review of Accounting*. 2011;19(2):147–156. DOI: 10.1108/13217341111181078
19. Baños-Caballero S., García-Teruel P.J., Martínez-Solano P. The speed of adjustment in working capital requirement. *The European Journal of Finance*. 2013;19(10):978–992. DOI: 10.1080/1351847X.2012.691889
20. Saravanan P., Sivasankaran N., Maram S., Shaw T.S. Enhancing shareholder value through efficient working capital management: An empirical evidence from India. *Finance India*. 2017;31(3):851–871.
21. Chang C.-C. Cash conversion cycle and corporate performance: Global evidence. *International Review of Economics and Finance*. 2018;56:568–581. DOI: 10.1016/j.iref.2017.12.014
22. Gonçalves T., Gaio C., Robles F. The impact of working capital management on firm profitability in different economic cycles: Evidence from the United Kingdom. *Economics and Business Letters*. 2018;7(2):70–75. DOI: 10.17811/ebl.7.2.2018.70–75
23. Duchin R., Ozbas O., Sensoy B.A. Costly external finance, corporate investment, and the subprime mortgage credit crisis. *Journal of Financial Economics*. 2010;97(3):418–435. DOI: 10.1016/j.jfineco.2009.12.008
24. Almeida H., Campello M., Laranjeira B., Weisbenner S. Corporate debt maturity and the real effects of the 2007 credit crisis. *Critical Finance Review*. 2012;1(1):3–58. DOI: 10.1561/104.00000001

ABOUT THE AUTHORS

Ilker Yilmaz — Assistant Professor, Department of Accounting, Dhofar University, Salalah, Oman
https://orcid.org/0000-0001-8501-9664
Corresponding author
iyilmaz@du.edu.om

Goksel Acar — Assistant Professor, Department of Finance and Economics, Dhofar University, Salalah, Oman
https://orcid.org/0000-0001-6907-2430
gacar@du.edu.om

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 23.10.2021; revised on 10.11.2021 and accepted for publication on 17.12.2021.
The authors read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-64-84
UDC 336.63,519.237.5(045)
JEL C23, C26, G30

Endogeneity Problem in Corporate Finance: Theory and Practice

Z.V. Selezneva, M.S. Evdokimova
National Research University Higher School of Economics, Moscow, Russia

ABSTRACT

Endogeneity can cause a significant bias in the coefficient estimation, up to the change in sign. It leads to controversial research results, which also makes it difficult to adequately test individual hypotheses and theories in corporate finance (CF). For practitioners, such as company valuation consultants, these model problems interrupt obtaining the most reliable estimates in the interests of the customer. The **aim** of this study is to review an endogeneity problem in CF and ways to solve a problem of endogeneity. We will illustrate the methods found in the systematic review with an empirical example. The paper provides the reasons for this problem from an econometric point of view and with examples from the CF and econometric methods of dealing with it. As a **result** of a systematic literature review, we have shown that dynamics panel models, in particular the Blundell-Bond **method**, are mostly used to consider endogeneity in CF studies. We have verified empirically the conclusion made in the framework of the literature review. To detect the endogeneity, we used the Hausman test, the endogeneity test, and the analysis of the correlation matrix, including the saved regression residuals. Eliminating step-by-step endogeneity, we **concluded** that the Blundell-Bond method is not always the optimal one for dealing with endogeneity in CF, as well as regression with a fixed effect. It was revealed that the two-stage least squares method (IV 2SLS) is the most appropriate method for the cost of capital model estimation eliminating endogeneity. In addition, the estimates of the cost of capital model, which analyzes the impact of non-financial reporting, have been improved.

Keywords: corporate finance; endogeneity; instrumental variables; panel data regression analysis; cost of capital; non-financial reporting

For citation: **Selezneva Z.V., Evdokimova M.S. Endogeneity problem in corporate finance: Theory and practice.** ***Finance: Theory and Practice.*** **2022;26(3):64-84. DOI: 10.26794/2587-5671-2022-26-3-64-84**

INTRODUCTION

Corporate finance (hereinafter referred to as "CF") as an independent empirical science has been developing for seventy years [1]. It is difficult to find theoretical papers in this area over the past two decades without elements of empirical research. At the same time, the intensive use of various econometric models in the CF.

In this area, hypotheses have traditionally been tested using linear regression. However, the complexity of the CF study object of a company suggests the potential mutual influence between several factors [2]. And the limited data collected suggests that linear regression may be misspecified, missing important factors. In other words, studies in CF are accompanied by endogeneity.

There has been an ongoing discussion about the presence of endogeneity in the CF models for about 9 years. Wintoki and Linck [2] note that researchers mainly rely on two possible sources of endogeneity and use models for panel data, in particular, the fixed effects model, to overcome it. Other scientists [3] have shown that most CF research is conducted without regard to potential endogeneity. Several studies [4, 5] use simulated data to show which methods are superior in accuracy to estimates of fixed effects (FE) models in the presence of endogeneity. They show that FE models may not be sufficient to deal with endogeneity. Although, as Banik notes [6], this approach is common in modern scientific research.

The presence of endogeneity in models is a violation of the premise of the Gauss-Markov theorem on the independence of regression residuals from explanatory variables, which causes bias and inconsistency in estimates. From a practical point of view, this has a negative effect:

- applicability of estimation models: results obtained with 'naive' estimating, such as the least squares method, do not provide unbiased

estimates in the presence of endogeneity. As a result, the actual influence of the independent variables may be underestimated or overestimated. For example, in the paper of Molina [7], taking into account endogeneity leads to an increase in the negative impact of financial leverage on the credit rating by 3 times, and in the article by Chen et al. [8] after evaluating the model taking into account endogeneity, the influence of the CEO Duality became insignificant for the firm's performance;

• interpretability of models. One of the manifestations of endogeneity is that it is impossible to establish causality, which can be caused by several reasons:

– reverse causality. Initially, the explanatory variable (X) is assumed to influence or precede the explanatory variable (Y). But the analysis reveals otherwise. A similar question is raised in many studies in various areas of CF [9–11];

– simultaneity. For example, Harada and Ngyen [12] note that the ownership structure and financial policy of the company are determined simultaneously;

– dependence of the explained and explanatory variables on the same factor not taken into account in the model (CMV). For example, Molina [7] proved that the financial leverage and credit rating of a company depend on the unobservable fundamental risk of the company;

• comparability of research results. Flannery and Hankins [5] note that regression estimates for the same models and similar data vary significantly from study to study. This is also observed in modern research. So, for example, the ambiguity of the impact of non-financial reporting on the performance of the company is noted by both Russian researchers: Zhukova and Melikova [13]; Martynova [14]; Polyakov et al. [15] and foreign ones: Zahid et al. [16];

• accuracy of predictive models, which is important, for example, for analysts when predicting the value of assets both in the financial market [17] and profitability in the real sector [18], and for representatives of the state [19];

• stability of models. Models become sensitive to the introduction of additional factors or changes in the sample, which indicates their instability. For example, Coles et al. [20] demonstrate that a 10% change in parameters leads to a change in the sign of the endogenous variable in the model without considering endogeneity;

• reliability of conclusions and adequacy of measures taken on the basis of the results of the regression analysis: decisions based on the conclusions of the model may not be optimal. Li [21] showed that the model without endogeneity evaluates the impact of CEO compensation on the firm's performance (Tobin's Q) as positive, but with it — negative.

Summarizing the above, the researcher runs the risk of encountering an incorrectly estimated model due to endogeneity. At the same time, the estimation bias in such models can vary from insignificant to huge. The estimates may even change sign or be insignificant.

In this article, we systematize information about potential sources of endogeneity and methods to overcome it in CF research. Like Flannery and Hankins [5], we are primarily interested in the structuring of information about endogeneity in CF studies, how researchers define it and how they deal with it.

Unlike previously cited articles, this paper provides descriptions of potential sources of endogeneity. Examples from different areas of CF research are also given. In addition, we will illustrate one way of overcoming endogeneity by examining the impact of companies' additional non-financial reporting about their activities on the firm's performance. This is a relatively new topic in the field of CF. The Journal of Corporate Finance published 56 articles on this topic, 33 of which were published in the last 5 years.[1] Within the framework of this topic, methods for determining endogeneity and ways to deal with it are demonstrated on a sample of 663 companies in the BRICS countries in 2007–2016. During model evaluation, it was shown that one of the most reliable methods

[1] The search was carried out using the keywords "non-financial report", "CSR", "ESG", "IR", "sustainability" in all publications of the journal in Scopus.

for estimating endogeneity regressions, the Blundell-Bond method, is not the default method for working with endogeneity, like the individual fixed effect model. From a statistical point of view, it was possible to overcome endogeneity in the model only by using the two-stage least squares method. This shows that researchers should choose the best estimation method using appropriate econometric instruments.

Further, this article is structured as follows: in the section "Endogeneity" this problem is considered as such, and the applicability of methods to deal with it in CF is analyzed; The section "Causes of endogeneity and ways to deal with it" is devoted to the systematization of potential sources of endogeneity in CF and methods to overcome it, depending on its source. The third section, "Endogeneity in panel data models", contains a table describing the advantages and disadvantages of the most popular methods for dealing with endogeneity using panel data. The fourth section "Testing and dealing with endogeneity using the example of a model for assessing the impact of non-financial information disclosure on the cost of capital in the BRICS countries" is devoted to an empirical test of the applicability of the above methods. The last section contains conclusions from the theoretical and empirical parts of the paper.

ENDOGENEITY

Before moving on to the real reasons for the inconsistent and/or biased results of model building with endogeneity, it is worth noting that it may not have been present in many research papers. Wintoki et al. [2] cited articles in which endogeneity was present, but there were no ways to overcome it. However, they might have gotten this impression because of the style of presenting the methodology in the articles. As in all branches of economics, econometric models serve as a tool in CF, a detailed description of which is quite often neglected. This is because the main research result usually consists in the refutation or confirmation of the hypotheses posed by the researcher in the field of CF, and not in the application of the econometric model.[2] In such cases, the authors omit "extra" details, because they are forced to present the main material in a text of a limited size. The description of important econometric details depends on technical factors (requirements for the text length in journals) and accepted norms among scientists from different fields of economics. The issue of detailing the description of methodologies deserves to be noted in this article and a separate discussion. Nevertheless, the insufficiently detailed description of the methodology prompted the authors to focus on one of the important problems in CF — endogeneity.

By definition, endogeneity is the presence of a non-zero conditional expectation of the regression residual for a particular explanatory factor. The second more familiar definition is the presence of non-zero covariance between the explanatory factor and the regression error.

How does this occur in CF? The company solves complex problems. Decision-making on each of them is influenced by both external and internal corporate factors. As a result, there is a simultaneous influence of internal factors on each other, which is one of the reasons for endogeneity in CF models — simultaneity. We consider this and other reasons for a specific example. Often the issue of endogeneity is raised in capital structure studies, which is one of the main topics in CF. The CEO power can influence the capital structure. This is a relatively new explanatory factor that endogeneity may entail.

The reasons for this may be the following:

1) it is a hard-to-measure quantitative value, which may remain an omitted variable;

2) a proxy can be invented for it; however, it can be inaccurate and will reflect the CEO power with a big error. Or the proxy will take into account not only the influence of the CEO in the company (measurement error);

3) if researchers decide to use peer review to measure CEO power and other indicators, then a specific measurement error may arise — common-method variance (CMV);

[2] An exception is the papers in which the econometric model is applied for the first time within the framework of the task. In such cases, the authors usually describe each step in detail.

4) during the study, simultaneity may take place. In CF, it can arise, for example, when studying the factors affecting the absolute liquidity ratio. Trade-off theory states that financial leverage and liquidity ratio are determined simultaneously.

The presence of one of the four sources of endogeneity makes the estimates of the coefficients biased and inconsistent, which also makes it difficult to interpret the model as a whole. Moreover, the inconsistency gets worse as the number of observations increases. And large samples (more than a thousand observations) are characteristic of studies in the field of CF published during the last decade.

In turn, Grisser and Hadlock [22] emphasize that a large number of studies in the field of CF do not test the model for endogeneity (in a strong or weak form), as well as the application of methods to deal with this problem. Gippel et al. [3, p. 144] note that only three out of 30 studies use tests to detect endogeneity and methods for leveling it. Most of the studies simply mention its possible presence. Although the volume of articles on CP with the mention of the word "endogeneity" has increased, a similar phenomenon and incorrect ways to deal with endogeneity continue to be observed not only in Australian articles but also in articles around the world [23–25].

However, in the last 5–7 years, CF researchers have come across the problem of endogeneity much more often. This can be caused by several reasons:

- researchers [5, 16] recognize that endogeneity may be one of the possible causes of the problem of inconsistency and incomparability of results;
- the number of articles devoted to the problem of endogeneity has increased;
- the growing interest in regressions to evaluate panel data, which have become widely used to work with CF endogeneity.

If we look at methodological econometric papers published around the same time as Gippel et al. [3], they are rarely cited in empirical studies on CF[3] (*Fig.*).

A small number of citations may be due to the fact that they were published in journals not devoted to the topic of CF: over the past three years, methodological articles [22, 26–29] from CF journals have been cited more often than articles from econometric journals [30–34].

This article contributes to methodological research on endogeneity in CF and adds to the existing articles as follows:

1. This article lists possible sources of endogeneity in CF studies.
2. This article highlights tests for endogeneity and ways to deal with it.
3. Like Flannery, Hankins [5] and Campbell, Nagel [35], we illustrate with real data the possible bias of the study results that does not take into account endogeneity. In contrast to these works, this illustration was made as part of solving the problem of assessing the impact of issuing non-financial reporting on the cost of equity of companies based on data of 663 companies from the BRICS countries for 2007–2016.

CAUSES OF ENDOGENEITY AND WAYS TO DEAL WITH IT

As mentioned above, there are 4 causes for endogeneity: measurement error variables, common method variance, omitted variables, and simultaneity. Endogeneity due to any of these causes can be detected using the Hausman test [36], also called the Durbin-Wu-Hausman (DWH) test, and the Wooldridge test [37].

The Hausman test [36] consists in comparing the reliability of estimates. In the case of endogeneity, the least squares method (LSM) estimates and instrumental variables (IV) are compared with the endogenous variable. If there is no difference between them in terms of consistency, then the null hypothesis is not rejected, the variable is exogenous. The DWH test is presented differently by the authors,

[3] We refer to citations for papers specifically on the method, and not to the result of applying the method. For example, Dang et al. (2015) [26] are cited in 67 papers. Most of the citations are related to the degree of influence of SOA (speed of adjustment) on the level of leverage. Only in some papers, researchers refer to the effectiveness of applying one or another method, depending on the presence of endogeneity and other problems in the data.

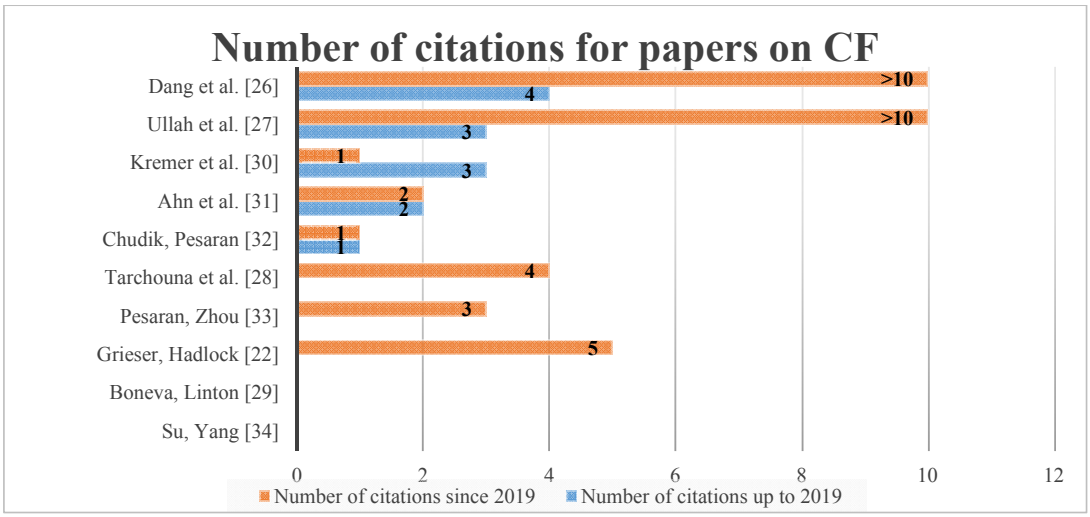


Fig. 1. **Citation rate of methodological publications on the use of panel models with endogeneity**

Source: authors' analysis of publications in Scopus.

however, if the errors are homoscedastic, the options of the three investigators are equivalent. In turn, the Wooldridge test [37] is the right instrument to test for endogeneity in the presence of heteroscedasticity. The null hypothesis is similar to the null hypothesis of the Hausman test.

Measurement errors

Measurement errors can be defined as the use of error-measured regressors instead of real regressors. In CF, measurement errors can be associated with:

- source data errors (for example, incorrectly loaded data from financial statements);
- errors in data aggregation (for example, irrelevant peer companies were selected to calculate the discount factor; an error in the formula when calculating the average capitalization for the industry);
- errors related to measurement methods (for example, CEO skills are poorly measured by their total work experience).

Measurement errors can be avoided by some common preventive measures to check the obtained data. This is controlled at the stage of uploading data and performing calculations to obtain exogenous factors. The control of the factor at the stage of unloading is carried out by checking the methodology for calculating the variable in the data source. Checking one's own calculations is also a method of dealing with measurement errors. However, such preventive measures do not guarantee the absence of measurement error. Several methods for dealing with data containing measurement errors have been proposed by Zhang et al. [38] and Qin et al. [39]. They are grouped as follows:

- methods that directly correct the bias of naive estimates: the naive estimator correction method, simulation extrapolation (SIMEX);
- correction methods based on the likelihood function, generalized linear models;
- methods based on unbiased evaluation functions;
- method of instrumental variables.

It is worth remembering that these methods have limitations in application. Instrumental variables may not be available. When implementing SIMEX, it may be difficult to select an extrapolation function. Correction methods based on the likelihood function require assumptions about the true distribution of factors and measurement errors. The construction of unbiased estimating functions under given assumptions about the distribution of factors is also a non-trivial problem. In addition to these methods, there are other robust methods for estimating regression equations [20, 21], which are improved methods for estimating the parameters of generalized linear models.

In financial economics, there is extensive practice in evaluating models with errors.

Aggregate information on this practice can be found in Chen et al. [40]. In the financial economy, measurement errors are reduced with the help of:

- calculation of estimates of the generalized method of moments (GMM);
- Bayesian estimates;
- three stage least squares method (3SLS);
- linear structural relations equations (LISREL);
- a particular case of LISREL is the Multiple Indicator, Multiple Cause (MIMIC) model.

Common method variance

The common method variance, CMV, is detailed by Wall et al. [41]. In fact, CMV is a pseudo-connection between variables, which is due to the presence of a factor — the source of the values of these variables. If a researcher were to conduct a survey among experts regarding the company performance on various parameters, the pessimistic experts would underestimate the estimates for all parameters. This problem is mainly related to personal data. CF researchers potentially rarely encounter such a source of endogeneity. If it is necessary to eliminate such bias in models, one should turn to articles in psychology or sociology [41–43]. Approaches to the compilation of survey materials and their processing are described in various sources, the most cited among which is the article by Podsakoff et al. [42]. Methods for checking CMV and some ways to deal with it are described in articles [41, 43].

Omitted variables

Omitted variables are one variation of the model specification error. The main ways to deal with the omitted variable error are to change the specification of the model and apply the IV method. A specification change can be expressed both in the transition to a completely alternative specification and in the transition to structural equation models (SEM).

Radically changing the model specification is an expensive way to solve the problem of omitted variables. This method will require a lot of time and resources. However, the main obstacle in its implementation is usually the lack of theory of using alternative model specifications. As for instrumental variables, these are variables that should have two properties: strongly correlated with the endogenous explanatory factor and weakly correlated with the model error. Although finding the right instrument is not a trivial task, this method is quite popular in articles. Over the past 5 years, the Journal of Corporate Finance has published about 270 articles on CF using the IV method.[4] This is partly due to the fact that there are some universal tools: endogenous variable lags or high-order moments. Instrumental variables can also be searched in new databases: specific industry and regional indicators can be found in the Observatory of Economic Complexity database,[5] CEO information can be found in the BoardEx of Management Diagnostics Limited,[6] Standard & Poor's ExecuComp database, etc.

Estimation methods with instrumental variables are usually instrumental variable method (IV), which includes Jackknife corrections (Jackknife IV estimates, described in the paper by Carrasco et al. [44]), two-stage least squares (2SLS, described in the paper by Schaffer [45]) and the generalized method of moments (GMM, described in the paper by Schaffer [45]), which is recommended for potentially overdetermined models. There is also a three-stage least squares method (3SLS, described in the paper by Qian, Schmidt [46]), the application of which will change the specification of the model to a system of externally uncoupled equations (SUR). Fixed effect models have become popular as a method of solving the omitted variable problem associated with each company. They clearly include the individual effect for each firm. However, this method is not a panacea, as described in more detail by Campbell, Nagel [35].

[4] Search results for articles with the keyword "instrumental variable" in the Journal of Corporate Finance. URL: https://www.sciencedirect.com/search?qs=%22instrumental%20variable%22&pub=Journal%20of%20Corporate%20Finance&cid=271687&years=2021%2C2020%2C2019%2C2018%2C2017&lastSelectedFacet=years (accessed on 06.07.2021).

[5] URL: https://oec.world/en/ (accessed on 06.07.2021).

[6] URL: https://corp.boardex.com/ (accessed on 06.07.2021).

Simultaneity

This problem is related to the fact that not one endogenous variable appears simultaneously in the model, but two or more. In other words, in one unit of time, both the explained variable and some explanatory factors are determined. This is also manifested in an unobvious causality between endogenous factors. In CF, simultaneity arises when considering the following questions:

- non-financial reports add value to the firm or high-value firms issue non-financial reports;
- the owners of the company set the level of credit risk the company accepts. Or can a certain level of the company's credit risk attract new owner investors?

As with the previous problem, lag variable models and systems of simultaneous equations such as vector error correction models (VECM) or vector autoregression (VAR) can be used to work with simultaneity).

The inclusion of an endogenous explanatory variable in the lag model is a fairly popular and effective option for dealing with endogeneity. However, some of its applications do not eliminate the bias caused by simultaneity but exacerbate it. Reed [47] showed that in the process of overcoming endogeneity, researchers evaluated models simultaneously, although there was no evidence for this. He also demonstrated that endogenous factor lags are better used as instrumental variables. They showed that lags are an effective instrument if they are sufficiently correlated with the variable being explained, and lag values are not included in the initially estimated equation. In other words, if the equation was originally estimated:

$$Y_t = \alpha + bX_t + cX_{t-1} + \varepsilon_t, \quad (1)$$

where Y_t, X_t — time series, $\varepsilon_t \sim N(0;\sigma_y)$, and the coefficients *b* and *c* can be equal to zero. Many researchers have replaced[7] specification with:

$$Y_t = \alpha + \beta X_{t-1} + \text{error term}. \quad (2)$$

But there is no reason to replace with, because X may not precede Y [29, p. 898–902]. And one should evaluate (3):

$$Y_t = \alpha + \beta X_t + \text{error term}, \quad (3)$$

using the IV method. In this case, the lag of the endogenous explanatory factor (X_{t-1}). will be a strong instrument. These results are also confirmed for panel models.

ENDOGENEITY IN PANEL DATA MODELS

In the course of developing models for panel data, methods for their evaluation were developed that make it possible to successfully deal with endogeneity, in particular, a model with a fixed effect (FE). Also, the method of instrumental variables was adapted for panel data. At the end of the 20th century, Arellano Bond and Blundell Bond [48] developed methods of the same name for estimating dynamic panel models (Arellano-Bond — AB and Blundell-Bond — BB). In addition to these methods, methodologies proposed in other studies can work successfully with endogenous variables [49, 50]. These 4 methods (FE, IV, AB, BB) were comprehensively compared by Flannery, Hankins [5]. They tested which of the dynamic panel regression estimates gave the smallest root mean square error (RMSE). The study was conducted on real and simulated data on CF, while the simulated data was created with the following imperfections: unbalanced panel, heterogeneity, and various types of endogeneity. The results of the study showed that the BB methodology in most cases gave the lowest RMSE values (*Table 1*).

In addition, BB proved to be effective in heterogeneity as well. Flannery, Hankins [5] and Dang et al. [26] showed that better estimates are obtained only when using the bootstrap technique for a fixed effect model. Nevertheless, the BB method is currently in demand among researchers [6, 52–54].

However, it is always worth sticking to common sense and not relying on the most popular methods. First, researchers must be sure that they are using the best model available today. Just as in assessing

[7] Some of them are listed in Reed, 2015, p. 897–898.

Table 1

Comparison of methodologies for panel models' estimation used in CF

Methods	Description
OLS	Not applicable
FE	The coefficient accuracy is low when working with wide (short) panels. Exogenous regressors are estimated quite accurately in the presence of regression lags, and the coefficients for lags are estimated biased. However, in the presence of second-level lags, it shows more reliable results than the Arellano-Bond and Blundell-Bond methods. Ineffective in the presence of weakly changing explanatory factors
LSDVC [51]	Despite the semantic identity of the coefficients in the case of FE models and the consistency of estimates with an unbalanced panel, this option is not suitable for estimating panel models with endogenous variables
IV (2SLS)	Two-stage least squares method is the traditional way to evaluate regressions to deal with endogeneity. Required prerequisites: homoscedasticity, including conditional, and does not imply work in systems of equations with dependent errors (simultaneous systems). Potentially can be used to deal with omitted variables, simultaneity, and measurement errors
Arellano-Bond	Refers to GMM estimates, which, together with the specification, allow model evaluation with endogenous variables. Perfect for evaluating models with instruments — lags. If the lag values of the explanatory factors are weak instruments, the estimates are biased
Blundell-Bond	Refers to GMM estimates, which, together with the specification, allow model evaluation with endogenous variables. Similar to the AB approach, it gives inconsistent estimates in the presence of second-order autocorrelation. Shows more effective results in the presence of weakly variable explanatory factors. Other things being equal, it is invariant to the degree of endogeneity, in contrast to AB, and is the preferred method of assessment, because it gives the lowest RMSE in the presence of different sources of endogeneity
Long difference [49]	Unlike the previously mentioned approaches, lags of the second and higher order can be used as instruments. The authors suggest choosing the largest of the available lags. The method showed average results in the presence of endogeneity and other issues
Four period differencing [50]	Unlike Hahn et al. [49], it is recommended to use not the largest available lag. The method showed average results in the presence of endogeneity and other issues

Source: authors' analysis [5].

it, is it reasonable to use a very sophisticated instrument to deal with endogeneity when a relatively simple problem is being solved? Should one immediately take an advanced method that has already been tested by someone, or should one first check the alternatives? The question is open whether at least it is worth showing in the study if the problem of endogeneity was discovered, if it really exists, and what methods were used to overcome it. And what if one managed to overcome it, and the tests say that everything is in order? Maybe not entirely. It is possible that there was more than one source of endogeneity. Based on this principle, in the next section, we will show that the Blundell-Bond method is not always the preferred method for estimating panel regressions in CF. Although often used without comparison with alternative methods.

TESTING AND DEALING WITH ENDOGENEITY USING THE EXAMPLE OF A MODEL FOR ASSESSING THE IMPACT OF NON-FINANCIAL INFORMATION DISCLOSURE ON THE COST OF CAPITAL IN THE BRICS COUNTRIES

As an example of testing and dealing with endogeneity, we use a model for assessing the impact of non-financial information disclosure on the cost of equity. To do this, we use a sample of 663 companies from the BRICS countries from 2007–2016 (balanced

panel). The topic of assessing the impact of non-financial information disclosure, including corporate social responsibility (CSR) and environmental, social, and corporate governance (ESG), began to gain popularity in the previous decade and is still relevant today. This is evidenced by numerous articles on this subject [55–64]. Non-financial information is disclosed in non-financial reports (hereinafter referred to as NFR). The sample includes the release of NFR according to the most popular standards: the Global Reporting Initiative (GRI)[8] and the International Integrated Reporting Council,[9] which recently included the International Integrated Reporting Committee (IIRC)[10] and the Sustainability Reporting Standards Board (SASB).[11]

This topic was chosen as an illustrative example, as many researchers in this field [64, 65] drew attention to the presence of an endogeneity problem in models. A meaningful interpretation of the endogeneity problem is: could the companies that issued NFR raise capital at a lower rate, or did companies that had access to cheaper equity issue NFR?

Model

Based on a number of studies [66–70], a model for assessing the impact of information disclosure in non-financial reporting on the cost of capital was formed, the description of the variables of which is presented in *Table 2*.

$$COE_{i,t} = \beta_0 + \beta_1 NFR_{i,t} + \beta_2 Size_{i,t} + \beta_3 M / B_{i,t} +$$
$$+ \beta_4 Lev_{i,t} + \beta_5 LTG_{i,t} + \beta_6 ROA_{i,t} + \sum_{j-1} Industry_j +$$
$$+ \sum_{t-1} Year_t + \sum_{k-1} Country_k + \varepsilon_{i,t}.$$

Sample

The sample contains 6630 observations of 663 companies over 10 years: 2007–2016. The data was collected from the Bloomberg database and a GRI database.[12] Descriptive statistics of variables are presented in *Table 3*.

The number of observations for the variable long-term growth rate (LTG) is 1326 fewer (663 observations each for 2007 and 2008), since data for 3 years were used in the calculation. The correlation matrix is presented in the *Appendix (Table 1*). *Tables 2* and *3* of the *Appendix* show the distribution of companies by country and industry, respectively. Financial sector companies were excluded from the sample due to their excellent balance sheet structure and additional government regulation.

Model evaluation and endogeneity testing

The models were evaluated using the STATA14 data analysis software package. The first step in the econometric analysis is the choice of a model estimation method. Let us compare the end-to-end model (pool) estimated by the least squares method without taking into account the individual effect with the estimates obtained by the generalized least squares method (GLM) with random (RE) and individual fixed (FE) effects (*Table 4*).

In the FE model, the inclusion of variables that do not change over time is impossible, since they are taken into account in the individual fixed error.

Based on the results of Breusch-Pagan [73], Hausman [36], and the F-test of the FE model [74], the FE model was chosen. The test results are presented in *Table 5*.

The Breusch-Pagan test demonstrates that the null hypothesis is rejected and the inclusion of individual random error makes the estimates more accurate, in other words, the RE model is preferred to the pooled model. According to the results of the Hausman test, the main hypothesis is rejected (p-value < 0.01): the FE model describes the data better than the RE model. The chosen model is significant because the null hypothesis F of the test is rejected. Note that fixed effects

[8] Global Reporting Initiative. URL: https://www.globalreporting.org/ (accessed on 22.07.2021).

[9] Value Reporting Foundation. URL: https://www.valuereportingfoundation.org/ (accessed on 22.07.2021).

[10] International Integrated Reporting Council. URL: https://integratedreporting.org/ (accessed on 22.07.2021).

[11] Accounting Standards Board. URL: https://www.sasb.org/ (accessed on 22.07.2021).

[12] URL: https://www.globalreporting.org/reporting-support/reporting-tools/sustainability-disclosure-database/ (accessed on 22.07.2021).

Table 2

Description of the variables used in the model

Variable	Description	Impact
$COE_{i,t}$	A company's cost of equity is calculated using the capital asset pricing model (CAPM [71]): $COE = r_f + \beta^*(r_m - r_f)$, where r_f – the risk-free rate measured as a yield on 10-year government bonds, r_m — market return; $\beta = \frac{cov(COE,\ r_m)}{var(r_m)}$ – a measure of systematic risk	
$NFR_{i,t}$	The binary variable NFR is 1 if the company issued an NFR in the current year, and 0 otherwise	–
$Size_{i,t}$	Natural logarithm of a company's total assets at the end of the year	–
$M\ /\ B_{i,t}$	The market-to-book ratio of the company at the end of the year	–
$Lev_{i,t}$	Financial leverage, measured as the debt-to-equity ratio at the end of the year	+
$LTG_{i,t}$	Long-term growth rate calculated as the average of 3-year sales growth rates	–
$ROA_{i,t}$	Return on assets, calculated as the company's net profit-to-assets ratio at the end of the year	–
$Industry_j$	Binary variable for industries	
$Country_k$	Binary variable for countries	
$Year_t$	Binary variable for years	

Source: Evdokimova M.S., Kuzubov S.A. [72].

models are a widely used way of estimating panel data in CF.

The obtained results indirectly show the presence of endogeneity: the FE model turned out to be the best, i.e. the usual error ($\varepsilon_{i,t}$) is not enough to correctly estimate the model. Further analysis using a modified Wald test revealed heteroscedasticity, which we will correct in the further analysis. Multicollinearity and spatial autocorrelation were not found. However, according to the Wooldridge test [75], there is serial autocorrelation. The feasible generalized least squares (FGLS) method can deal with this problem.

To test the hypothesis that the cost of capital decreases *after* the release of NFR the second model (*Table 6*) includes the NFR lag, which turned out to be significant. Therefore, the issuance of NFR leads to a reduction in the cost of equity, and not vice versa. Lags of higher order are insignificant. Note that financial leverage is significant only at the 15% significance level (p-value = 0.12). In the course of clarifying the reasons for the insignificance of this factor, a correlation was found between the residuals of the model and NFR equal to 0.34, which indicates the endogeneity of NFR. We also performed an endogeneity test, the null hypothesis of which is that the variable is exogenous. According to the results of this test, the endogeneity of NFR was again confirmed (*Table 7*).

Table 3

Descriptive statistics

Variables	Obser.	Average	Stand. error	Min.	Max.
COE	6.630	11.88	2.95	1.73	21.87
NFR	6.630	0.15	0.35	0	1
Size	6.630	6.75	1.38	2.86	11.48
ROA	6.630	5.03	6.47	−89.16	65.38
Leverage	6.630	0.35	0.55	0	11.95
LTG	5.304	7.83	11.24	−38.01	138.76
M/B	6.630	2.61	1.75	0.05	9.96

Source: authors' calculations.

Table 4

Estimates of pool model and GLS models with random (RE) and individual fixed effects (FE)

VARIABLES	pool	RE	FE
	Cost of equity		
NFR	−0.144	−0.452***	−0.730***
	(0.0917)	(0.103)	(0.117)
Size	0.0216	−0.0676	−0.551***
	(0.0263)	(0.0420)	(0.0950)
ROA	−0.0458***	−0.0245***	−0.0146***
	(0.00525)	(0.00526)	(0.00559)
Leverage	0.0658***	0.0845***	0.118***
	(0.0232)	(0.0303)	(0.0392)
LTG	−0.0155***	−0.00921***	−0.00555**
	(0.00280)	(0.00260)	(0.00268)
M/B	0.0871***	−0.139***	−0.266***
	(0.0200)	(0.0218)	(0.0240)
Dummies for years	Yes	Yes	Yes
Dummies for countries	Yes	Yes	No
Dummies for industries	Yes	Yes	No
	(0.117)	(0.0945)	(0.101)
Constant	11.47***	12.27***	14.51***
	(0.199)	(0.303)	(0.593)
Observations	5.304	5.304	5.304
R-squared	0.366		0.450
Number of Id		663	663

Note: Standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1.

Source: authors' calculations.

Table 5

Results of Breusch-Pagan, Hausman tests and F-test of FE model

Test	(1) Breusch-Pagan	(2) Hausman	(3) F-test of FE model
H0	No individual random effect	Complex error model with $cov(u_i, X_{it}) = 0$ – RE correctly specified	No individual fixed effect
Statistics	$\overline{\chi}^2(1)$ = 2222.33	$\chi^2(6)$ = 175.95	F (13,4628) = 290.77
P-value	0.0000	0.0000	0.0000

Source: authors' calculations.

Table 6

FGLS FE estimation taking into account serial autocorrelation and heteroscedasticity

VARIABLES	FGLS,	
	(1) NFR	(2) lag NFR
	Cost of equity	
NFR	−0.202**	
	(0.0952)	
L.NFR		−0.253**
		(0.103)
Size	0.109***	0.111***
	(0.0264)	(0.0260)
ROA	−0.0698***	−0.0699***
	(0.00531)	(0.00530)
Leverage	0.0368	0.0364
	(0.0237)	(0.0237)
LTG	−0.0194***	−0.0194***
	(0.00291)	(0.00291)
M/B	0.101***	0.101***
	(0.0199)	(0.0199)
Dummies for years	Yes	Yes
Constant	10.23***	10.21***
	(0.191)	(0.190)
Observations	5.304	5.304
Number of Id	663	663

Note: Standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1.

Source: authors' calculations.

Table 7

NFR endogeneity test

NFR endogeneity test statistics	10.634
P-value	0.0011

Source: authors' calculations.

DEALING WITH ENDOGENEITY

To deal with the NFR endogeneity, we turn to the method of instrumental variables (IV), the essence of which is to select instruments that are correlated with the endogenous variable, but not correlated with the residuals of the model. The search for instruments was carried out using the analysis of the extended correlation matrix, including the saved residuals of the FGLS model. The results of model estimation by the method of instrumental variables with NFR lag and financial leverage lag used as instruments are presented in *Table 8*.

It is worth noting that the IV method on panel data, when using a model with individual fixed effects, does not allow adding a constant to the model. Because of this, the value of the coefficients for binary variables for years has increased. To assess the quality of instruments, we have to perform several tests:

1. Underidentification test, its null hypothesis: instruments have insufficient explanatory power to predict endogenous variables [76].

2. Weak identification test checks whether instrumental variables explain the endogenous variable. Instruments are considered strong if the Cragg-Donald Wald F statistic exceeds all critical Stock-Yogo values [77].

3. Hansen-Sargan's J overidentification test [76]. Null hypothesis: instruments are valid (not correlated with errors).

In our case, all three tests were successfully passed, therefore, the instruments are strong (*Table 9*).

However, we established a second-order autocorrelation (AR (2)), to overcome this we included the first and second lags of the cost of equity in the instrumental model, which solved the problem of endogeneity. The results of the evaluation of the final model are presented in *Table 10*.

The instruments for the NFR are the NFR lag, the financial leverage lag, and the cost of equity lag used in the first stage of the 2SLS valuation. All three tests for this specification were successfully passed (*Table 11*), the instruments are strong, and the problem of endogeneity is solved (*Table 12*).

The correlation between the lags of the cost of capital and the residuals of the final model does not exceed 0.01 (*Table 13*).

In the presence of serial autocorrelation and endogeneity, dynamic models of AB and BB or LSDVC estimates and the method of long differences (LD) can be used. However, the necessity of their application is justified by the

Table 8

IV FE estimation taking into account heteroscedasticity and endogeneity

VARIABLES	IV, Cost of equity
NFR	−1.303***
	(0.223)
Size	−0.455***
	(0.0886)
ROA	−0.0160**
	(0.00671)
Leverage	0.110**
	(0.0440)
LTG	−0.00656***
	(0.00254)
M/B	−0.259***
	(0.0244)
Dummies for years	Yes
Observations	5.304
Number of Id	663
R-squared	0.446

Note: Robust standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1.

Source: authors' calculations.

Table 9

Test results for evaluating the quality of instruments

Test	Underidentification test	Weak identification test	Overidentification test
Statistics	Kleibergen-Paap rk LM statistic = 290.956	Cragg-Donald Wald F statistic = 1090.522	Hansen J statistic = 0.363
P-value	0.000		0.5467
Stock-Yogo critical values:			
10% maximal IV size		19.53	
15% maximal IV size		11.59	
20% maximal IV size		8.75	
25% maximal IV size		7.25	

Source: authors' calculations.

presence of a non-zero correlation between the lags of the explained variable and the residuals of the model, which is not observed in our model. In other words, in the case of our application of AB, BB or LD at once, without gradually solving all the problems that violate the Gauss-Markov assumptions, we would never get unbiased coefficient estimates. Going to any of these models without testing the assumptions and characteristics of the simpler models would show that they are irrelevant. The BB estimates bias can be seen in *Table 14*.

Considering that the average value of the cost of capital in the analyzed sample is 11.88 percentage points, its reduction from NFR issuance ranges from 0.73 and 0.78 percentage points (6% of the average) in FE and final models, respectively, up to 1.3 percentage points (11% of the average) in the IV model without taking into account autocorrelation. The estimates differ almost by a factor of 2, which is essential for interpreting the results obtained and using them in practice.

Since no other endogenous variables were identified, the analysis of the illustrative example is completed on the final model, the results of which are given in *Table 10*.

CONCLUSIONS

Due to the complexity of company organization, CF models can potentially always contain

Table 10

Final model: IV FE estimation taking into account heteroscedasticity, endogeneity and AR (2)

VARIABLES	IV with COE lags, Cost of equity
NFR	−0.784***
	(0.264)
L.COE	0.182***
	(0.0197)
L2.COE	−0.0970***
	(0.0145)
Size	−0.479***
	(0.131)
ROA	−0.0157**
	(0.00734)
Leverage	0.107**
	(0.0494)
LTG	−0.00337
	(0.00281)
M/B	−0.250***
	(0.0272)
Dummies for years	Yes
Observations	4.641
Number of Id	663
R-squared	0.446

Source: authors' calculations.

Note: Robust standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1.

Table 11

Test results for evaluating the quality of instruments

Test	Underidentification test	Weak identification test	Overidentification test
Statistics	Kleibergen-Paap rk LM statistic = 209.207	Cragg-Donald Wald F statistic = 710.499	Hansen J statistic = 0.672
P-value	0.000		0.4123
Stock-Yogo critical values:			
10% maximal IV size		19.53	
15% maximal IV size		11.59	
20% maximal IV size		8.75	
25% maximal IV size		7.25	

Source: authors' calculations.

endogeneity. This paper continues the discussion about the presence of endogeneity in these models and how to deal with it. In contrast to previous works on this topic, our article provides examples of potential sources in different areas of research in the field of CF and considers various ways to deal with endogeneity.

In the course of a systematic review of the literature, we found that many researchers recommend using regressions to evaluate dynamic panel models, as they are the most successful in leveling the effects of various kinds of endogeneity. In particular, according to the results of the review, the Blundell-Bond method is the most appropriate.

However, when conducting an empirical study using methods to detect and overcome endogeneity, we found that this method may not be sufficient to eliminate endogeneity. As part of an empirical study, we tested the impact of NFR on the cost of capital of 633 companies from the BRICS countries from 2007–2016. Our result is consistent with previously published research on this topic: NFR has a negative impact on the cost of capital. However, this result was obtained using the method of instrumental variables on panel data (IV 2SLS). And the generalized least squares method with an individual fixed effect (FGLS FE), as well as BB, was not enough to overcome endogeneity.

According to the estimates obtained in this study, we found a non-critical change in the coefficient for the endogenous variable between FE and the final IV models, however, if we settled on the 2SLS IV estimate without taking into account autocorrelation, the coefficient would be 2 times larger. According

Table 12

NFR endogeneity test

NFR endogeneity test statistics	2.148
P-value	0.1428

Source: authors' calculations.

Table 13

Correlation of potentially endogenous variables with the final model's residuals

Variables	COE	L.COE	L2.COE	NFR	Residuals
COE	1.00				
L.COE	0.49***	1.00			
L2.COE	0.21***	0.47***	1.00		
NFR	−0.09**	−0.01***	0.09***	1.00	
Residuals	0.58***	0.00	0.00	0.02	1.00

Source: authors' calculations.

Note: *** p < 0.01, ** p < 0.05, * p < 0.1.

Table 14

Beta coefficients of NFR variable in the estimated models

Model	FE	FGLS	IV	BB	IV final
NFR beta	−0.730*** (0.117)	−0.202** (0.0952)	−1.303*** (0.223)	−1.018*** (0.217)	−0.784*** (0.223)

Source: authors' calculations.

Note: Standard errors and robust standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1.

to the review of the literature, this is not the worst manifestation of endogeneity, since the absence of a correction for it can lead to an even greater coefficients bias, insignificance, and even a change in sign.

The results of our study will be primarily useful both for experienced and novice researchers of the cost of capital, and for CF researchers in general. Also, our study may be of interest to practitioners in the field of business valuation, and investors to evaluate and apply models that have a high predictive ability due to objective coefficient estimates. Also, this article can be useful when conducting seminars on CF or econometrics.

Research on the problem of endogeneity has great potential for development both from the point of view of researchers in the field of econometrics involved in the development of new models that take into account endogeneity and from the side of CF practitioners, many of whom are faced with the problem of choosing the best way to evaluate the model with endogeneity.

ACKNOWLEDGEMENTS

This paper is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University). National Research University Higher School of Economics, Moscow, Russia.

REFERENCES

1. Tucker G.M. On problems of corporate finance. *The American Journal of Economics and Sociology*. 1948;7(2):235–236. DOI: 10.1111/J.1536–7150.1948.TB 00679.x
2. Wintoki M.B., Linck J.S., Netter J.M. Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics*. 2012;105(3):581–606. DOI: 10.1016/j.jfineco.2012.03.005
3. Gippel J., Smith T., Zhu Y. Endogeneity in accounting and finance research: Natural experiments as a state-of-the-art solution. *Abacus*. 2015;51(2):143–168. DOI: 10.1111/abac.12048
4. Barros L.A.B.C., Bergmann D.R., Henrique Castro F., da Silveira A.D.M. Endogeneity in panel data regressions: Methodological guidance for corporate finance researchers. *Revista Brasileira de Gestao de Negocios*. 2020;22:437–461. DOI: 10.7819/rbgn.v22i0.4059
5. Flannery M.J., Hankins K.W. Estimating dynamic panel models in corporate finance. *Journal of Corporate Finance*. 2013;19:1–19. DOI: 10.1016/j.jcorpfin.2012.09.004
6. Banik A., Chatterjee C. Ownership pattern and governance-performance relation: Evidence from an emerging economy. *Global Business Review*. 2021;22(2):422–441. DOI: 10.1177/0972150920966699
7. Molina C.A. Are firms underleveraged? An examination of the effect of leverage on default probabilities. *The Journal of Finance*. 2005;60(3):1427–1459. DOI: 10.1111/j.1540–6261.2005.00766.x
8. Chen C.-W., Lin J.B., Yi B. CEO duality and firm performance: An endogenous issue. *Corporate Ownership and Control*. 2008;6(1):58–65. DOI: 10.22495/cocv6i1p6
9. Poletti Hughes J. R&D and dividend payments as determinants of corporate value in the UK: Empirical evidence after controlling for endogeneity. *International Journal of Managerial Finance*. 2008;4(1):76–91. DOI: 10.1108/17439130810837393
10. Zhou T., Li W.-A. Board governance and managerial risk taking: Dynamic analysis. *The Chinese Economy*. 2016;49(2):60–80. DOI: 10.1080/10971475.2016.1142823
11. Malik Q.A., Hussain S., Ullah N., Waheed A., Naeem M., Mansoor M. Simultaneous equations and endogeneity in corporate finance: The linkage between institutional ownership and corporate financial performance. *The Journal of Asian Finance, Economics and Business*. 2021;8(3):69–77. DOI: 10.13106/jafeb.2021.vol8.no3.0069
12. Harada K., Nguyen P. Ownership concentration and dividend policy in Japan. *Managerial Finance*. 2011;37(4):362–379. DOI: 10.1108/03074351111115313
13. Zhukova N. Yu., Melikova A.E. Corporate social responsibility: Strengthening brand value and affecting company's financial performance. *Finance: Theory and Practice*. 2021;25(1):84–102. DOI: 10.26794/2587–5671–2021–25–1–84–102

14. Martynova M. Climate disclosure as a factor of investment attractiveness of Russian companies. *Moskovskii economicheskii zhurnal = Moscow Economic Journal*. 2021(5):365–379. (In Russ.). DOI: 10.24411/2413–046X-2021–10277
15. Polyakov K.L., Polyakova M.V., Samoylenko S.V. Modeling the impact of the debt load on efficiency of business entities. *Voprosy statistiki*. 2016;(9):17–29. (In Russ.). DOI: 10.34023/2313–6383–2016–0–9–17–29
16. Zahid M., Rahman H.U., Khan M., Ali W., Shad F. Addressing endogeneity by proposing novel instrumental variables in the nexus of sustainability reporting and firm financial performance: A step-by-step procedure for non-experts. *Business Strategy and the Environment*. 2020;29(8):3086–3103. DOI: 10.1002/bse.2559
17. Cheng Q., Lo K. Insider trading and voluntary disclosures. *Journal of Accounting Research*. 2006;44(5):815–848. DOI: 10.1111/j.1475–679X.2006.00222.x
18. Gul F.A., Hutchinson M., Lai K.M.Y. Gender-diverse boards and properties of analyst earnings forecasts. *Accounting Horizons*. 2013;27(3):511–538. DOI: 10.2308/acch-50486
19. Kim C.-J., Nelson C.R. Estimation of a forward-looking monetary policy rule: A time-varying parameter model using ex post data. *Journal of Monetary Economics*. 2006;53(8):1949–1966. DOI: 10.1016/j.jmoneco.2005.10.017
20. Coles J.L., Lemmon M.L., Meschke J.F. Structural models and endogeneity in corporate finance: The link between managerial ownership and corporate performance. *Journal of Financial Economics*. 2012;103(1):149–168. DOI: 10.1016/j.jfineco.2011.04.002
21. Li F. Endogeneity in CEO power: A survey and experiment. *Investment Analysts Journal*. 2016;45(3):149–162. DOI: 10.1080/10293523.2016.1151985
22. Grieser W.D., Hadlock C.J. Panel-data estimation in finance: Testable assumptions and parameter (in) consistency. *Journal of Financial and Quantitative Analysis*. 2019;54(1):1–29. DOI: 10.1017/S 0022109018000996
23. Kirshin I.A. An empirical analysis of determinant factors of the company's capital structure. *Finansy: teoriya i praktika = Finance: Theory and Practice*. 2017;21(2):106–112. (In Russ.). DOI: 10.26794/2587–5671–2017–21–2–106–112
24. Mohammad W.M.W., Wasiuzzaman S. Environmental, social and governance (ESG) disclosure, competitive advantage and performance of firms in Malaysia. *Cleaner Environmental Systems*. 2021;2:100015. DOI: 10.1016/J.CESYS.2021.100015
25. Alareeni B.A., Hamdan A. ESG impact on performance of US S&P 500-listed firms. *Corporate Governance*. 2020;20(7):1409–1428. DOI: 10.1108/CG-06–2020–0258
26. Dang V.A., Kim M., Shin Y. In search of robust methods for dynamic panel data models in empirical corporate finance. *Journal of Banking & Finance*. 2015;53:84–98. DOI: 10.1016/j.jbankfin.2014.12.009
27. Ullah S., Akhtar P., Zaefarian G. Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. *Industrial Marketing Management*. 2018;71:69–78. DOI: 10.1016/j.indmarman.2017.11.010
28. Tarchouna A., Jarraya B., Bouri A. How to explain non-performing loans by many corporate governance variables simultaneously? A corporate governance index is built to US commercial banks. *Research in International Business and Finance*. 2017;42:645–657. DOI: 10.1016/j.ribaf.2017.07.008
29. Boneva L., Linton O. A discrete-choice model for large heterogeneous panels with interactive fixed effects with an application to the determinants of corporate bond issuance. *Journal of Applied Econometrics*. 2017;32(7):1226–1243. DOI: 10.1002/jae.2568
30. Kremer S., Bick A., Nautz D. Inflation and growth: New evidence from a dynamic panel threshold analysis. *Empirical Economics*. 2013;44(2):861–878. DOI: 10.1007/s00181–012–0553–9
31. Ahn S.C., Lee Y.H., Schmidt P. Panel data models with multiple time-varying individual effects. *Journal of Econometrics*. 2013;174(1):1–14. DOI: 10.1016/j.jeconom.2012.12.002
32. Chudik A., Pesaran M.H. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*. 2015;188(2):393–420. DOI: 10.1016/j.jeconom.2015.03.007
33. Pesaran M.H., Zhou Q. Estimation of time-invariant effects in static panel data models. *Econometric Reviews*. 2018;37(10):1137–1171. DOI: 10.1080/07474938.2016.1222225
34. Su L., Yang Z. QML estimation of dynamic panel data models with spatial errors. *Journal of Econometrics*. 2015;185(1):230–258. DOI: 10.1016/j.jeconom.2014.11.002

35. Campbell R.C., Nagel G.L. Private information and limitations of Heckman's estimator in banking and corporate finance research. *Journal of Empirical Finance*. 2016;37:186–195. DOI: 10.1016/j.jempfin.2016.03.007
36. Hausman J.A. Specification tests in econometrics. *Econometrica*. 1978;46(6):1251–1271. DOI: 10.2307/1913827
37. Wooldridge J.M. Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*. 1995;68(1):115–132. DOI: 10.1016/0304–4076(94)01645-G
38. Zhang Y., Qin G., Zhu Z., Zhang J. Robust estimation in linear regression models for longitudinal data with covariate measurement errors and outliers. *Journal of Multivariate Analysis*. 2018;168:261–275. DOI: 10.1016/j.jmva.2018.07.015
39. Qin G., Zhang J., Zhu Z. Simultaneous mean and covariance estimation of partially linear models for longitudinal data with missing responses and covariate measurement error. *Computational Statistics & Data Analysis*. 2016;96:24–39. DOI: 10.1016/j.csda.2015.11.001
40. Chen H.-Y., Lee A.C., Lee C.-F. Alternative errors-in-variables models and their applications in finance research. *Quarterly Review of Economics and Finance*. 2015;58:213–227. DOI: 10.1016/j.qref.2014.12.002
41. Wall T.D., Michie J., Patterson M., Wood S.J., Sheehan M., Clegg C.W., West M. On the validity of subjective measures of company performance. *Personnel Psychology*. 2004;57(1):95–118. DOI: 10.1111/j.1744–6570.2004.tb02485.x
42. Podsakoff P.M., MacKenzie S.B., Lee J.-Y., Podsakoff N.P. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*. 2003;88(5):879–903. DOI: 10.1037/0021–9010.88.5.879
43. Tehseen S., Ramayah T., Sajilan S. Testing and controlling for common method variance: A review of available methods. *Journal of Management Sciences*. 2017;4(2):146–175. DOI: 10.20547/jms.2014.1704202
44. Carrasco M., Doukali M. Efficient estimation using regularized jackknife IV estimator. *Annals of Economics and Statistics*. 2017;(128):109–149. DOI: 10.15609/annaeconstat2009.128.0109
45. Schaffer M.E. XTIVREG2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. 2020.
46. Qian H., Schmidt P. The asymptotic equivalence between the iterated improved 2sls estimator and the 3sls estimator. *Econometric Reviews*. 1997;16(4):441–457. DOI: 10.1080/07474939708800398
47. Reed W.R. On the practice of lagging variables to avoid simultaneity. *Oxford Bulletin of Economics and Statistics*. 2015;77(6):897–905. DOI: 10.1111/obes.12088
48. Roodman D.M. How to do Xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*. 2009;9(1):86–136. DOI: 10.1177/1536867X0900900106
49. Hahn J., Hausman J., Kuersteiner G. Long difference instrumental variables estimation for dynamic panel models with fixed effects. *Journal of Econometrics*. 2007;140(2):574–617. DOI: 10.1016/j.jeconom.2006.07.005
50. Huang R., Ritter J.R. Testing theories of capital structure and estimating the speed of adjustment. *Journal of Financial and Quantitative Analysis*. 2009;44(2):237–271. DOI: 10.1017/S 0022109009090152
51. Bruno G.S.F. Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models. *Economics Letters*. 2005;87(3):361–366. DOI: 10.1016/j.econlet.2005.01.005
52. Ezeani E., Salem R., Kwabi F., Boutaine K., Bilal, Komal B. Board monitoring and capital structure dynamics: evidence from bank-based economies. *Review of Quantitative Finance and Accounting.* 2022;58(2):473–498. DOI: 10.1007/S 11156–021–01000–4
53. Lartey T., Danso A., Boateng A. Co-opted boards and capital structure dynamics. *International Review of Financial Analysis*. 2021;77:101824. DOI: 10.1016/J.IRFA.2021.101824
54. Badayi S.A., Matemilola B.T., Bany-Ariffin A.N., Theng L.W. Does corporate social responsibility influence firm probability of default? *International Journal of Finance & Economics*. 2021;26(3):3377–3395. DOI: 10.1002/IJFE.1966
55. Schadewitz H., Niskala M. Communication via responsibility reporting and its effect on firm value in Finland. *Corporate Social Responsibility and Environmental Management*. 2010;17(2):96–106. DOI: 10.1002/csr.234
56. de Klerk M., de Villiers C. The value relevance of corporate responsibility reporting: South African evidence. *Meditari Accountancy Research*. 2012;20(1):21–38. DOI: 10.1108/10222521211234200

57. Iatridis G.E. Environmental disclosure quality: Evidence on environmental performance, corporate governance and value relevance. *Emerging Markets Review*. 2013;14:55–75. DOI: 10.1016/j.ememar.2012.11.003
58. Fatemi A., Fooladi I., Tehranian H. Valuation effects of corporate social responsibility. *Journal of Banking & Finance*. 2015;59:182–192. DOI: 10.1016/j.jbankfin.2015.04.028
59. Oshika T., Saka C. Sustainability KPIs for integrated reporting. *Social Responsibility Journal*. 2017;13(3):625–642. DOI: 10.1108/SRJ-07–2016–0122
60. de Villiers C., Venter E.R., Hsiao P.-C.K. Integrated reporting: background, measurement issues, approaches and an agenda for future research. *Accounting & Finance*. 2017;57(4):937–959. DOI: 10.1111/acfi.12246
61. García-Sánchez I.-M., Noguera-Gámez L. Integrated information and the cost of capital. *International Business Review*. 2017;26(5):959–975. DOI: 10.1016/j.ibusrev.2017.03.004
62. Albuquerque R., Koskinen Y., Zhang C. Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science*. 2019;65(10):4451–4469. DOI: 10.1287/mnsc.2018.3043
63. Bhuiyan M.B.U., Nguyen T.H.N. Impact of CSR on cost of debt and cost of capital: Australian evidence. *Social Responsibility Journal*. 2020;16(3):419–430. DOI: 10.1108/SRJ-08–2018–0208
64. Yeh C.-C., Lin F., Wang T.-S., Wu C.-M. Does corporate social responsibility affect cost of capital in China? *Asia Pacific Management Review*. 2020;25(1):1–12. DOI: 10.1016/j.apmrv.2019.04.001
65. Manchiraju H., Rajgopal S. Does corporate social responsibility (CSR) create shareholder value? Evidence from the Indian Companies Act 2013. *Journal of Accounting Research*. 2017;55(5):1257–1300. DOI: 10.1111/1475–679X.12174
66. Dhaliwal D.S., Li O.Z., Tsang A., Yang Y.G. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*. 2011;86(1):59–100. DOI: 10.2308/accr.00000005
67. El Ghoul S., Guedhami O., Kwok C.C.Y., Mishra D.R. Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance*. 2011;35(9):2388–2406. DOI: 10.1016/j.jbankfin.2011.02.007
68. Boujelbene M.A., Affes H. The impact of intellectual capital disclosure on cost of equity capital: A case of French firms. *Journal of Economics, Finance and Administrative Science*. 2013;18(34):45–53. DOI: 10.1016/S 2077–1886(13)70022–2
69. Zhou S., Simnett R., Green W. Does integrated reporting matter to the capital market? *Abacus*. 2017;53(1):94–132. DOI: 10.1111/abac.12104
70. Fama E.F., French K.R. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*. 1993;33(1):3–56. DOI: 10.1016/0304–405X(93)90023–5
71. Fama E.F., French K.R. The cross-section of expected stock returns. *The Journal of Finance*. 1992;47(2):427–465. DOI: 10.1111/j.1540–6261.1992.tb04398.x
72. Evdokimova M.S., Kuzubov S.A. Non-financial reporting and the cost of capital in BRICS countries. Higher School of Economics. Basic Research Program Working Paper. 2021;(83/FE). URL: https://wp.hse.ru/data/2021/06/25/1430000555/83FE 2021.pdf
73. Breush T.S., Pagan A.R. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*. 1979;47(5):1287–1294. DOI: 10.2307/1911963
74. Park H.M. Practical guides to panel data modeling: A step by step analysis using stata. Niigata: International University of Japan; 2011. 53 p. URL: https://www.iuj.ac.jp/faculty/kucc625/method/panel/panel_iuj.pdf
75. Drukker D.M. Testing for serial correlation in linear panel-data models. *The Stata Journal*. 2003;3(2):168–177. DOI: 10.1177/1536867X0300300206
76. Baum C.F., Schaffer M.E., Stillman S. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *The Stata Journal*. 2007;7(4):465–506. DOI: 10.1177/1536867X0800700402
77. Stock J.H., Yogo M. Testing for weak instruments in linear IV regression. In: Andrews D.W.K., Stock J.H., eds. Identification and inference for econometric models: Essays in honor of Thomas Rothenberg. Cambridge, New York: Cambridge University Press; 2005:80–108. DOI: 10.1017/CBO9780511614491.006

Appendix

Table 1

Correlation matrix

Variables	COE	NFR	Size	ROA	Leverage	LTG	M/B
COE	1.00						
NFR	−0.03***	1.00					
Size	0.04***	0.43***	1.00				
ROA	−0.15***	0.09***	−0.03**	1.00			
Leverage	0.06***	0.05***	0.29***	−0.28***	1.00		
LTG	−0.09***	0.03**	0.08***	0.32***	−0.01**	1.00	
M/B	−0.03***	−0.05***	−0.25***	0.24***	−0.043	0.12***	1.00

Note: *** p < 0.01, ** p < 0.05, * p < 0.1.

Source: authors' calculations.

Table 2

Companies' distribution across countries

Country	Number of companies
Brazil	15
China	437
India	163
Russia	11
South Africa	37
Total	663

Source: authors' calculations.

Table 3

Companies' distribution across industries

GICS code	Industry	Number of companies
10	Energy	17
15	Materials	133
20	Industry	177
25	Consumer Discretionary	115
30	Consumer Staples	58
35	Healthcare	54

Table 3 (continued)

GICS code	Industry	Number of companies
45	Information Technology	53
50	Communication Services	4
55	Utilities	39
	Diversified Corporations	13
	Total	663

Source: authors' calculations.

ABOUT THE AUTHORS

Zinaida V. Selezneva — Research Assistant, Financial Engineering and Risk Management Laboratory, PhD student, School of Finance, National Research University Higher School of Economics, Moscow, Russia
http://orcid.org/0000-0002-0487-4181
zseleznyova@hse.ru

Mariia S. Evdokimova — Lecturer and PhD student, School of Finance, National Research University Higher School of Economics, Moscow, Russia
http://orcid.org/0000-0002-7915-4736
evdokimovamary@mail.ru

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 30.07.2021; revised on 15.08.2021 and accepted for publication on 17.12.2021.
The authors read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-85-109
UDC 336.76:330.4:330.332(045)
JEL E17, G11, G17, G24, G35

Construction of a Dynamic Normative Model of the Stock Valuation Rating

A.S. Tonkikh[a], S.A. Tonkikh[b], E.Yu. Maslova[c]
[a, b] Admiral Ushakov State Maritime University, Novorossiysk, Russia;
[c] Kuban State University Novorossiysk Branch, Novorossiysk, Russia

ABSTRACT

Ratings are widely used in stock analysis, as they increase information transparency, simplify the assessment and investment decision-making in the stock markets, and increase their effectiveness. However, the use of ratings is constrained by subjective factors, which necessitates replacing expert assessments with objective characteristics. The aim of the study is to develop a stock valuation rating model, which allows obtaining additional information about the qualitative characteristics of shares in relation to existing rating methods, ensuring the transparency of the rating methodology and limiting the influence of subjectivity, affiliation and bias of an expert analyst. On the basis of the method of dynamic standard, a rating score is constructed, which expresses the measure of the proximity of the normative and actual order of the growth rates of indicators characterizing the investment attractiveness of stocks. When choosing current indicators for inclusion in the rating, the principle of their dynamic co-subordination (normative ordering by growth rates) proposed by I.M. Syroezhin and successfully used in the analysis of economic situations has been applied. The authors have built a share valuation rating model, which is a system of inequalities for the growth rates of six indicators (reference aggregated indicator — benchmark): the market price of a share, the size of the dividend per share, the company's net profit, revenue, equity and the number of shares in circulation. One of the advantages of the developed model is its universality: it is applicable to any industry, any market, and any accounting system. The model will make it possible to address the problems of expert subjectivism, the use of projected values and the availability of initial information (indicators of official reporting). The authors note the limitation of the model — it is not applicable to the valuation of early-stage companies. The model has been tested by compiling ratings of common stocks from the MICEX index10 for one-year and five-year periods. A comparative analysis of the ratings showed that a higher rating of stocks is observed in the long term. This also determines the applicability of the developed model as a stock analysis tool for long-term fundamental investments.

***Keywords*:** stock ratings; criteria for investment attractiveness of stocks; economic and mathematical modeling; dynamic co-subordination of indicators; market price of stocks

For citation: **Tonkikh A.S., Tonkikh S.A., Maslova E.Yu. Construction of a dynamic normative model of the stock valuation rating. *Finance: Theory and Practice*. 2022;26(3):85-109. DOI: 10.26794/2587-5671-2022-26-3-85-109**

INTRODUCTION

Ratings and rating assessments are widely used by people both in everyday life and in professional activities. Ratings of films, quality of life in the regions, universities, smartphones emitting the most radiation, etc. — this is a small list of those ratings that people turn to in everyday life. Among the ratings used by specialists in their professional activities, there are credit ratings, ratings of companies according to certain indicators, ratings of most reliable banks, etc. It is no coincidence that ratings in the field of economics and finance are presented as examples: the main objective in these areas is to choose the best investment project, increase the efficiency of resources, and ratings solve the object ranking problem, for example, potential investment objects, according to some criterion.

Let us define the initial concepts.

The meaning of the concept of rating, with various approaches to its understanding, comes down to two interpretations:

1. Rating as a rating score is a general indicator (number, rank, interval level) that reflects the quality of an object or the quality of an object's functioning, its results based on certain initial information.

2. Rating in the form of a ranked or classification list — placement of rating objects according to a general indicator that reflects the quality of the object or the quality of the object's functioning, and its results. The ranking is usually carried out from a more significant object according to a certain criterion to a less significant one.

These two approaches should not be opposed, they should be considered as two complementary or, moreover, as two consecutive stages of the rating process: first, a rating score is calculated, and then a ranked list of objects is formed based on the rating score.

The ranked list may not be compiled according to the calculated rating score, but according to some basic indicators that do not require additional processing, such as, sales revenue and profit. In this case, such ranked lists are often referred to as ratings.

Having determined the essence of the rating process, let us consider its application in relation to the stock ratings.

Stock ratings are one of the most used analysis tools in stock market decision-making. Hence, almost every investment company (bank), an investment consultant has its own stock ratings. Moreover, the stock ratings are drivers of the market value of stocks. In particular, this phenomenon is confirmed in [1]. Using the example of Polish stock indices, the author substantiates that rating agencies are one of the most important sources of information, and their rating reports are considered an indicator of confidence in the organization's securities. The rating itself, coupled with information about the projected rating, can have a significant impact on investment decisions and, as a result, on stock indices. This is also true for the markets of other countries. Thus, a comprehensive analysis of the impact of stock ratings on Japanese markets, which confirms the main conclusions of the above study, is considered in [2].

The growing demand for ratings in the United States led to the development of the financial industry regulator FINRA (Financial Industry Regulatory Authority) rule 2212 on the use of ratings by investment companies.[1] A fragment of a similar rating published by the Goldman Sachs Group analysts is shown in *Fig 1*.

As a result of the increasing spread of ratings as a tool for analyzing stocks, special web services have appeared — stock screeners (for example, the finviz[2]), stock screener), which allow investors to create their own ratings (rankings) of stocks according to the parameters they are interested in: dividend yield, P/E and P/S ratios, earnings per share, etc.

The popularity of stock ratings can be explained by the following reasons:

[1] FINRA Rule 2212. URL: https://www.finra.org/rules-guidance/rulebooks/finra-rules/2212 (accessed on 21.04.2021).
[2] Finviz. URL: https://finviz.com/screener.ashx (accessed on 21.04.2021).

Company	Date	Action	Brokerage	Rating	Price Target
RHM Rheinmetall	4/21/2021	Target Set by	The Goldman Sachs Group	Buy	€133.00
HSY The Hershey	4/21/2021	Upgraded by	The Goldman Sachs Group	Neutral -> Buy	$171.00 -> $181.00
PENN Penn National Gaming	4/21/2021	Target Raised by	The Goldman Sachs Group	Buy	$153.00 -> $158.00
RRR Red Rock Resorts	4/21/2021	Downgraded by	The Goldman Sachs Group	Buy -> Neutral	$33.00 -> $36.00
RCL Royal Caribbean Group	4/21/2021	Target Raised by	The Goldman Sachs Group	Neutral	$76.00 -> $95.00
CCL Carnival Co. &	4/21/2021	Target Raised by	The Goldman Sachs Group	Neutral	$21.00 -> $26.00
KER Kering	4/21/2021	Target Set by	The Goldman Sachs Group	Buy	€750.00
AAPL Apple	4/21/2021	Target Set by	The Goldman Sachs Group	Sell	$83.00

Fig. 1. **Fragment of the Goldman Sachs Group analyst stock ratings**

Source: MarketBeat. URL: https://www.marketbeat.com/ratings/by-issuer/goldman-sachs-group-stock-recommendations/ (accessed on 21.04.2021).

1. Reducing labor intensity and saving time for collecting and processing a large amount of information about the issuer of shares.

2. Obtaining information about the quality and characteristics of shares in a concise form.

3. Simplicity of comparison of stocks according to one final, instead of several separate indicators.

4. Ranking of stocks according to the degree of urgency of their acquisition.

5. Additional justification, regardless of the subject of investment, of the expediency of investing in specific shares.

6. Reduction of information uncertainty in relation to investment objects.

7. Search for new investment ideas in the stock market.

8. Ability to choose the object of investment in shares without expertise.

It is for these reasons that the stock ratings, compiled in the form of recommendations — buy, actively buy, hold, sell, actively sell, have become most widespread. Non-fundamental variations of this scale are possible: strong selling, strong buying, numbers from 1 to 5, above the market, below the market, etc. In most cases, this type of rating is compiled by investment companies that are interested in the growth of purchases of shares by their clients, so the rating scale contains an incentive for users of ratings to take specific actions.

Such recommendations may be based on the idea of the target (predicted) stock price, which in this case will be an integral element of such ratings. *Fig. 1* shows an example of such recommendations. Predicted stock prices are expert, often contradictory assessments of various stock analysts. The inconsistency of ratings forces users of ratings to compare, generalize and derive a final rating based on them, which deprives ratings obtained in this way of their main advantages: reducing the complexity and reliability of information when analyzing and investing in company stocks.

These contradictions are caused by a number of factors:

1. *Subjectivity of estimates of target prices.*

First, experts may use different methods of evaluation, the results of which do not match. Second, even if experts use the same methods, they may not have the same understanding of the variables from

which the resulting value is calculated. For example, when applying the discounted cash flow method, the target stock price in the most general form is found by formula (1):

$$PP = \sum_{t=1}^{\infty} \frac{Div_t}{(1+i)^t}, \quad (1)$$

where PP — price prediction, monetary unit;

Div_t — dividend in period t (predicted), monetary unit;

i — acceptable (required) rate of return for an investor, % per annum.

Experts' opinions can differ both in terms of expected dividends and in terms of the required rate of return, and their perceptions of the target price can vary greatly.

2. *Information asymmetry*.

Information about the issuer is unevenly distributed among experts, or some of the necessary information is missing. Consequently, experts, using unequal information, receive different estimates of the same object.

3. *Expert affiliation or bias*.

This issue is well covered in various sources. For example, [3] shows that optimistic analysts have more favorable career prospects. In [4], analysts' bias is considered as a rational property of company financial performance forecasts. Sometimes analyst bias is a reaction to the expectations of the market, which tends to listen to the decisions of analysts with a higher percentage of "buy" ratings even in the face of a downgrade [5]. Equity analysts are also prone to bias: they tie their earnings-per-share forecasts to the industry average without making sufficient adjustments, which encourages hiding negative information about stocks, especially for companies with worse ratings [6]. Affiliated analysts whose securities companies are mortgagees to the firms that have pledged shares issue optimistic rating reports for them due to conflicts of interest [7].

The study [8] is devoted to confirming the hypothesis that the detail and tone of financial analyst reports are much more indicative of an analyst's sentiments than their analysis of quantitative indicators, and can be used to assess the extent to which conflicts of interest of analysts interfere with the display of real estimates of the value of the firm in stock recommendations. In [9], the problem of the herding behavior of analysts is studied. The authors argue that analysts in some cases usually tend to give the same recommendations as their colleagues, thus forming a "prevailing consensus forecast", which is less accurate than "bold" forecasts.

An independent analyst sometimes works with an affiliate analyst to improve the accuracy of their recommendations because, compared to an independent analyst, an affiliate analyst has more information but faces a conflict of interest. Due to the conflict of interest of the affiliated analyst, the value that the independent analyst expects to receive in order to avoid the herding decision may motivate him to get more information [10].

Other aspects of expert bias can also be identified.

4. *The proximity of methods for assigning ratings*.

Evaluators are not inclined to disclose in detail the content of the rating methodology. This is their know-how — a competitive advantage over other raters, which has commercial value due to its obscurity to third parties. The current situation makes it impossible to verify the ratings, which hypothetically creates the preconditions for the dishonesty of the rating subjects in a broad sense: from arbitrary, poorly substantiated assessments to frankly erroneous, unconditional judgments or even fraud.

The foregoing necessitates the weakening of the influence of the subjective opinion of an expert on the rating score by expanding the use of objective characteristics.

For example, many experts see the solution to the problem of the objectivity of a rating compiled on the basis of stock price predictions in improving forecasting

methods using regression methods or machine learning [11–16]. But this does not completely remove the problem of subjectivity: starting from the choice of the type of specific dependence, model variables, their significance, and ending with fitting the model to empirical material. In addition, the non-obvious interpretability of such models reduces confidence in the results obtained by rated consumers: how can an investor be sure that an analyst is right about the target price of a share, if it is not always clear how this price was obtained? It is argued in [17] that the justifications that support the analyst's opinion reduce, and in some models exclude, the importance of profit forecasts and revision of recommendations. The rationale for the target price is the only essential element for an investor to accept an analyst's opinion. There is no correlation between forecasting methodology and analyst accuracy or market reaction to a report. Moreover, although the use of machine learning methods improves the efficiency of forecasts, it significantly complicates the forecasting process. However, the complexity of such models is often a serious obstacle to their practical application. In [18], the myth that complex predictive models outperform simpler ones is questioned.

Developing the idea that the use of predictive models in the formation of stock ratings should be treated with caution, we also note the work [19] with reference to [20, 21]. In particular, research [20] presents the results of a study of 296 published significant factors that are used by financial economists when predicting stock returns. The authors conclude that between 80 and 158 of these are likely to be false discoveries, and this casts doubt on the reliability of such predictive models.

Criticism of approaches to stock price forecasting based on regressions began in 2011 in a message from the president of the AFA (American Finance Association) [21]. J. H. Cochrane suggested that the presence of a large number of noisy and highly correlated predictors does not allow reliable forecasting of returns using cross-data regressions and portfolio sorting, which entails the need to use alternative methods. As a similar approach, [19] proposes machine learning using regularization methods. However, machine learning methods face the usual problem of cross-predictability and opacity of the methods themselves. In any case, predicting stock prices using machine learning methods, the effectiveness and stability of predictive signals, and the lack of transparency in complex machine learning algorithms require additional research.

The problem of applying ratings in the Russian stock market is widely covered in domestic studies. First, we note the work [22], which is a systematic generalization of the rating methodology, the rationale for using ratings to improve the efficiency of management decisions, including investment, practical solutions in various fields of activity: economics, government regulation, sports, etc. As for the rating of valuable securities, the authors focus on the regulation and control of the debt market, while the ratings of shares remain without their attention. In [23], an option of the formation of an investment rating of shares based on the construction of a logistic regression is presented. However, the solution of the problem of independence of variables, which, as was shown earlier, is inherent in some regression functions, was left outside the scope of the presented work by the authors. At the same time, as the authors note, the results do not always correlate with empirical patterns, which in some cases forced them to resort to "manual" tuning of the model.

The study [24] developed a matrix rating model for evaluating the most liquid Russian stocks, taking into account the industry specifics of issuers. This model is based on the dependence of the market stock prices on the dividend policy of companies, primarily on the size of dividends paid. It should be noted that the amount of dividends paid is one of the most important, but not the

only factor in the stock attractiveness, it is advisable to include other indicators in the calculation of the stock rating.

A significantly expanded list of indicators for the formation of stock ratings is presented in [25–27]. They note that one of the most convenient tools for analyzing securities is the rating method, which allows presenting an impressive amount of incoming information in a concise, systematized form [25]. At the same time, the rating of shares should take into account both the indicators of profitability and risk of securities, and the quality of corporate governance.

In [26], the author proposed a rating algorithm based on the adaptation of the Graham-Rea model to the Russian stock market: selection of indicators, normalization and standardization of their values relative to generally accepted standards, calculation of the rating score as the sum of weighted normalized values of indicators. Meanwhile, the conditionality and unconvincingness of generally accepted values of indicators — for some companies, the share of equity less than 50% of total assets is really critical, for others — not (the same applies to other similar standards), as well as the need to determine the significance of indicators based on expert (according to essentially subjective) assessments create prerequisites for the use of alternative rating algorithms. Similar reasoning is appropriate for the score-rating assessment, as well as various combinations of the mentioned rating methods [27], in which the final assessment is impossible without the participation of experts.

The work [28] presents, in our opinion, a rather promising approach to the stock rating process, based on a hierarchical analysis of ranked indicators, taking into account their priority. The priority of indicators is a natural principle of building a stock rating and means that different indicators have different meaning for an investor when assessing the stock attractiveness. In this context, the market stock price is a more significant indicator than the inventory turnover ratio. It is clear that the first indicator makes a more significant contribution to the attractiveness of the stock than the second. Moreover, many analysts generally do not consider the turnover ratio to be a significant factor in the investment quality of a share, then its defining characteristic is the market value of the share or indicators derived from it in the overwhelming majority of cases. Less obvious examples of the hierarchical ordering of indicators can be given. However, when justifying the priority of indicators, in our opinion, one should rely more not on the rules of fuzzy logic (accumulated experience, intuition, etc.), as, for example, in [28], but on a more formal argumentation.

Thus, this study is aimed at creating a model for the rating assessment of stocks, which allows obtaining additional information about the qualitative characteristics of stocks in relation to existing rating methods, ensuring the transparency of the rating assignment methodology and limiting the influence of subjectivity, affiliation and bias of an expert analyst.

METHODOLOGY

Let us formulate the basic principles for the formation of the desired model for the rating evaluation of stocks.

1. *Objectivity.* The assessment should be completely independent of the opinion of the subject of assessment, their experience, intuition, preferences, interests, etc. In other words, when forming a model, it is desirable to exclude expert judgments.

2. *Retrospective*. The rating assessment must be built on the basis of achieved, not predicted results. This principle is a logical continuation of the previous one. As mentioned above, most forecasts are to some extent based on the subjective judgments of the analyst, and this contradicts the principle of objectivity. There will certainly be a lot of objections that a rating without indicating the prospects of the rating

Table 1

Example of assessing the dynamics of indicators excluding the principle of dynamic co-subordination

Indicator, monetary unit	Option 1			
	Previous period	Current period	Growth rate, %	Recommended dynamics
1	2	3	4	5
Dividend per share (*DPS*)	4	5	125	Growth
Earnings per share (*EPS*)	8	12	150	Growth
Conclusion	The dynamics correspond to the recommended, the assessment is positive			
Indicator, monetary unit	**Option 2**			
	Previous period	**Current period**	**Growth rate, %**	**Recommended dynamics**
1	6	7	8	9
Dividend Per Share (*DPS*)	4	3	75	Рост
Earnings per share (*EPS*)	8	4	50	Рост
Conclusion	The dynamics does not correspond to the recommended, the assessment is negative			

Source: compiled by the authors.

subject, forecasting its values in the future will inevitably lose practical significance and applied value, since when making investment decisions, not only the results achieved at the moment are important, but also their subsequent changes. Therefore, we will give a few more arguments in favor of the principle of retrospective rating.

It is noted in [29] that there is no single forecasting method for the profitability of all stock markets; different types of markets require their own methods. Based on this, we cannot reject the assumption that when predicting the quantitative characteristics of some stocks, certain forecasting methods work, others do not. This leads to doubts about the validity of comparisons of such stocks and their ranking relative to each other. And this, as we have determined, is one of the main advantages and purposes of the rating.

Further, statistical forecasting methods, one of the most common in the stock market, are based on extrapolation of historical data. It is logical to assume that if the rating score is based on previously achieved results, then the rating score itself should be formed on their basis, and not on hypothetical future values of indicators. The possibility of extrapolation and the trend towards stability of the current rating values in the future will be determined by the phenomenon of economic inertia. Economic inertia is the desire of an economic object to maintain its current state, the inability to quickly change it. The attractiveness of stocks cannot decrease instantly, this process will stretch over time, in the same way, the attractiveness of stocks cannot increase in a short time. Various aspects of the phenomenon of inertia, confirming its existence in the economy, are considered in [30–32].

In accordance with the above arguments, we consider it expedient to base the developed method of stock ratings on the retrospective principle. We are not against forecasting in the stock market, moreover, we consider it an integral part of the market analysis, but in order to build the required rating, we will adhere to an approach that is not based on forecast estimates.

3. *Methodological simplicity*. The principle of simplicity, sometimes generalized by

the principle of economy, has various interpretations [33–38]. In its most general form, the principle of simplicity is a heuristic principle that generalizes the experience of cognition, according to which, other things being equal, the simplest cognitive construction is preferable.[3] This principle has both supporters and opponents, who argue that in the process of development, science becomes more complicated, not simplified, moreover, the concept of simplicity is relative. We will discuss the validity of certain judgments. At the same time, in striving for a more reliable and convincing interpretation of the ratings obtained, in order to choose a simpler theoretical scheme for explaining the phenomenon of the attractiveness of shares, to meet the criterion of scientific rationality, we will use a modification of this principle — the principle of methodological simplicity [39]. According to this principle, the more complex the calculations and statistical techniques, the more difficult and arbitrary the interpretation of the data obtained. That is why we based the developed rating model on the principle of methodological simplicity.

4. *Generality of the model*. The desired rating model should be valid for the widest possible range of stocks, regardless of country, market, scale of operations, capitalization, industry affiliation, accounting system used, and other similar factors.

5. *Dynamic co-subordination and comparability of indicators*. There are two points to note here. Firstly, the attractiveness of stocks is determined by the growth potential of key market indicators, such as stock price, dividends, price/earnings ratio, book value of shares, etc. In other words, when choosing stocks for an investor, it is important what dynamics this or that indicator has, so the rating of stocks should be based on a dynamic criterion of attractiveness. Secondly, it is not just the positive dynamics of individual indicators that is important, but their orderly growth relative to each other. Let us explain the latter on the example of *Table 1*.

Table 1 presents two hypothetical options for dividends per share *(DPS)* and earnings per share *(EPS)*. At the top of the table, an example of an increase in indicators is considered: *DPS* by 25%, *EPS* by 50% (column 4). If we consider the indicators separately, i.e. without taking into account the principle of dynamic co-subordination, then their dynamics corresponds to the recommended one, and the final assessment will be positive. And vice versa, in the lower part of the table, the decrease in *DPS* and *EPS* was 25 and 50%, respectively (column 8), their dynamics does not correspond to the recommended one, and without taking into account the principle of dynamic co-subordination, the final assessment is negative.

Let us supplement our reasoning with the principle of dynamic co-subordination of indicators. It is obvious that the stocks of a company that has kept the share of profits allocated for payment of dividends (payout ratio) at a level not lower than in the previous period are more attractive to the investor than the stocks of a company that has reduced it. As a confirmation of this thesis, we can cite the situation with the payment of dividends by TATNEFT at the end of 2020.[4] The formulated criterion for the attractiveness of stocks can be written as an inequality ordering the growth rates of *DPS* and *EPS* indicators (2):

$$\text{Growth rate } (DPS) \geqslant \text{Growth rate } (EPS). \quad (2)$$

The growth rate (GR) of indicator A, is calculated by formula (3).

$$GR(A) = \frac{A_2}{A_1}, \quad (3)$$

where $GR(A)$ — growth rate of indicator A;

[3] Website: New Philosophical Encyclopedia. Electronic Library of the Institute of Philosophy RAS. The principle of simplicity. URL: https://iphlib.ru/library/collection/newphilenc/document/HASH2120954aaecac8b1b76221 (accessed on 02.08.2021).

[4] Tatneft will share half of the profits. Kommersant (Kazan) No. 76 of 04/29/2021. URL: https://www.kommersant.ru/doc/4793856 (accessed on 02.08.2021).

A_2 — the value of indicator A in the current period;

A_1 — the value of indicator A in the previous period.

Compliance with inequality (2) indicates at least a non-decrease in the share of the company's profits directed to the payment of dividends. The validity of (2) can be questioned by supporters of low, even zero dividends, which, in their opinion, lead to a much greater increase in the value of stocks by increasing the amount of profits reinvested in the company's assets. We will return to this thesis later, in the course of substantiating the desired rating model. Now we use inequality (2) as an illustration of the principle of dynamic co-subordination of indicators.

The co-subordination (mutual ordering) of indicators in dynamics, similar to inequality (1), allows us to make significant adjustments to our reasoning about the attractiveness of the stocks presented in *Table 1*. Thus, the authors, without hesitation, gave a positive assessment of the attractiveness of the stock for the first option: the actual dynamics of market indicators corresponds to the recommended one. However, taking into account inequality (2) does not allow such an unambiguous assessment, since for a given stock *GR* (*EPS*) > *GR (DPS)*, and according to (2) it should be vice versa.

The same picture with the second option. We negatively assessed the attractiveness of stocks, whose market characteristics have worsened in dynamics. At the same time, the normative relation (2) is observed for this action. In practice, the following interpretation of this situation is possible: an unfavorable market situation has developed, which has led to an objective decrease in the company's net profit and, as a result, to the inability, primarily at the legislative level, not only to increase dividend payments, but even to keep them in the amount of previous year. However, the organization has found a way to pay its shareholders a larger share of net income compared to the previous period, as evidenced by the smaller decline in *DPS* compared to *EPS*. In this sense, the stocks of the second option look quite attractive: despite the objective decrease in the amount of dividends paid, the company seeks to minimize the loss of current income of shareholders.

As a result, we have: the implementation of the principle of dynamic co-subordination of indicators significantly affects the result of assessing the attractiveness of stocks, which must be taken into account when constructing a rating model.

The principle of dynamic comparability of indicators means that indicators that are not comparable in statics become comparable in dynamics. Thus, for example, the indicator of the company's equity and the indicator of the number of outstanding stocks cannot be compared with each other if they are considered as static, including because of their different dimensions: the first indicator is measured in monetary units, the second is a dimensionless indicator. However, the growth rates of these indicators, used as a quantitative characteristic of their dynamics, are quite comparable values. We consider the book value per share. One of the ways to calculate it is presented in the form of formula (4).

$$\text{Book value per share (BV)} = \frac{\text{Equity (Eq)}}{\text{Number of shares outstanding (NSO)}}. \quad (4)$$

The book value per share characterizes, among other things, the amount that the shareholders of the enterprise will receive in the event of its liquidation. The higher this amount, the better for shareholders. Therefore, this indicator in dynamics is attributed to growth. For the growth of *BV*, a faster growth of the numerator relative to the denominator is required, which can be written as a dynamic relation (5).

$$\text{Growth rate } (Eq) > \text{Growth rate } (NSO). \quad (5)$$

Here, the principle of dynamic comparability of indicators is manifested:

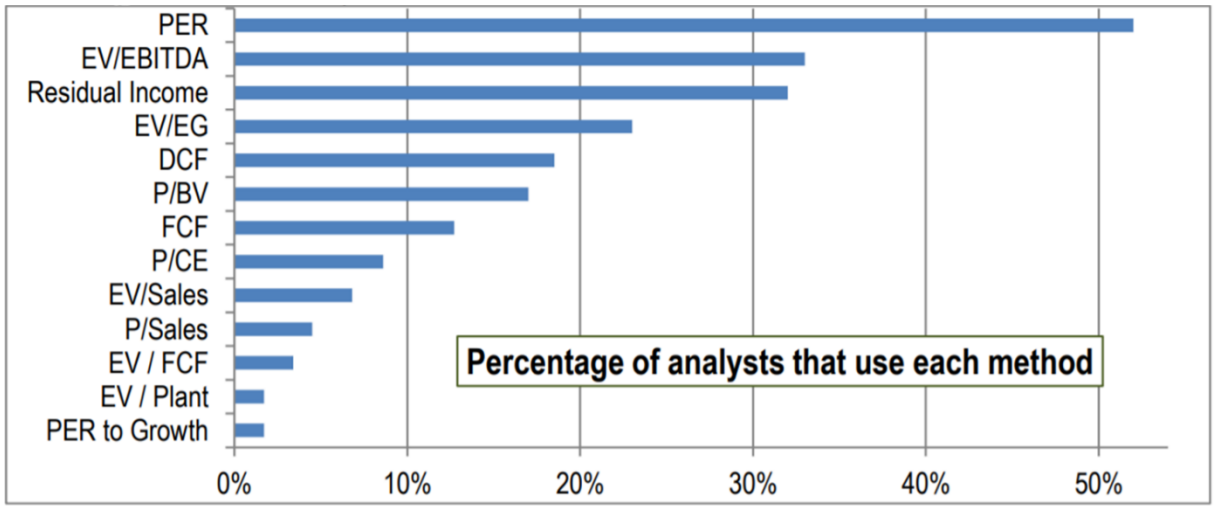


Fig. 2. **Most widely used stock valuation methods**

Source: P. Fernandez [46].

two indicators that are not comparable in statics become comparable in dynamics, which is an important methodological technique when building composite ratings when it is necessary to reduce several heterogeneous characteristics into one.

Dynamic co-subordination and comparability of indicators are the fundamental concepts of the theory of a dynamic standard, and expressions (2) and (5) can be considered as an example of the simplest dynamic standards or, in other words, the reference dynamics of indicators. Dynamic standards are called a system of indicators, normatively ordered by growth rates relative to each other. The idea of ordering indicators of the dynamics of economic systems belongs to I.M. Syroezhin [40]. In the future, this approach was repeatedly tested in the analysis of various economic situations, for example, in assessing sustainable development and the balance of economic systems [41, 42], modeling the development of production and economic structures [43], measuring and monitoring the parameters of various economic processes and phenomena [44, 45] and others. The situations considered in one form or another are based on the calculation of integral (rating) estimates, so the choice of a dynamic standard as a methodological method for building an exchange rating seems natural, logical and justified.

Let us move on to the content of the stock rating model.

RESULTS AND DISCUSSION

To select the indicators of the constructed rating model, we will use the results of the study [46]. In this article, with reference to Morgan Stanley Dean Witter Research, we present the most popular methods for assessing the attractiveness of stocks among stock analysts (*Fig. 2*).

The authors note that the preference for one method or another, i.e. the percentage of analysts using one method or another, depends on the industry and the country of the companies. At the same time, the composition of the calculated indicators basically remains the same — similar to the composition in *Fig. 2*. Description and formulas for calculating these indicators are presented in *Table 2*.

We do not consider in this paper which of these methods better or worse. We are primarily interested in the list of indicators that determine the attractiveness of stocks, according to stock analysts. Moreover, in the context of the principle of dynamic co-subordination of indicators, we are interested not so much in the indicators,

Table 2

The most commonly used indicators in stock valuation

Designation	Indicator	Calculation formula
PER, P/E	Earnings Multiplier	Market capitalization / Net profit = Market price per share / earnings per share
EV/EBITDA	Enterprise Value-to-EBITDA ratio	Enterprise Value / EBITDA (earnings before interest, taxes and depreciation)
Residual Income, RI	Residual Income	Operating Profit – Minimum Required Rate of Return × Operating Assets
EV/EG	Enterprise Value-to-Projected EBITDA ratio	Enterprise value / EBITDA (actual) / projected EBITDA growth
DCF	Discounted cash flow	$\sum_{i=1}^{n} \frac{\text{projected annual cash flow}_i}{(1+\text{discount rate})^i}$
P/BV	Price-to-Book Value ratio	Market capitalization / Book value of equity
FCF	Free Cash Flow	Earnings before interest and taxes + depreciation – change in working capital requirement – capital expenditures
P/CE	Price-to-Cash Earnings ratio	Market capitalization / Net profit before depreciation
EV/Sales, EV/S	Enterprise Value-to-Sales ratio	Enterprise Value / Sales
P/Sales, P/S	Sales Multiplier	Market capitalization / Sales = Market price per share / Sales per share
EV/FCF	Enterprise Value-to-Free Cash Flow ratio	Enterprise Value / Free Cash Flow
PER to Growth, P/EG, PEG	Earnings Multiplier-to-Earnings Growth	PER / Projected EPS Growth

Source: compiled by the authors based on [46].

but in what indicators are compared with each other in their calculation process. This will allow the formation of a set of dynamic standards that will be taken as the basis for the desired model for stock ratings.

To begin with, in accordance with the principle of objectivity, we will exclude from the list of the most used indicators those of them that are calculated on the basis of predictive and subjective assessments. These include EV/EG, DCF, PEG, RI.

The following important point is which of the indicators characterizes the investment attractiveness of shares to a greater extent — *P* (market capitalization or market value per share) or *EV* (fair enterprise value / share)? The fair enterprise value is calculated using the formula (6).

Table 3

PER and P for Apple Inc stocks in 2019–2021

Indicator	2019	2020	September 2021
P/E Ratio (PER)	18,41	34,28	29,74
Market value (P), USD.	55	113	154

Source: compiled by the authors based on finanz.ru and finviz.com. URL: https://www.finanz.ru/balans/apple; URL: https://finviz.com/quote.ashx?t=AAPL&ty=c&ta=0&p=m (accessed on 31.08.2021).

$$EV = \text{Market Capitalization} + \\ + \text{Net Debt} + \text{Minority Interest}. \quad (6)$$

Net debt is the difference between the sum of a company's long-term and short-term liabilities and cash.

Some stock analysts believe that to assess the attractiveness of shares, it is necessary to use the *EV* indicator, and not *P*. This is motivated by the fact that the investor when buying stocks, i.e. stocks in the company's business, acquires from this the obligations and free cash of the company, which increase or decrease the value of the company. An example is often given with the purchase of an apartment with a mortgage on it, the value of which increases due to the fact that, along with the apartment, the obligation to pay the balance of the debt to the bank passes to the new owner. At the same time, if a cache of money is found, the new owner can appropriate it, thereby compensating for part of the cost of buying an apartment. However, there is a significant difference between buying an apartment and buying company stocks: in the case of buying stocks, their new owner is not liable for the company's obligations and cannot directly claim the funds at its disposal. The investor, regardless of the method of the stock valuation, is interested in the amount of actual income per share, which is the sum of the difference between the sale and purchase price of a share, as well as dividends received during the period of ownership. At the same time, the owner of the share does not pay anything additional for the debts of the company and also does not receive anything, except dividends, from the funds at his disposal. The purchase and sale of shares take place at a real market price, and not at a hypothetical fair value, i.e. the investor's profit is formed as a result of transactions at market value.

We also add that the increase in *EV* due to debt may not affect the market value of the shares owned by the investor. In this case, the increase in fair value will be of little interest to the investor, since it is not accompanied by an increase in his income. This type of stock will not be attractive to an investor, despite the *EV* growth. Therefore, the investment attractiveness of stocks should be assessed not by the *EV* indicator, but by the market capitalization indicator *P*. That is why, when forming the stock rating model, we will not consider the indicators obtained on the basis of a comparison with the fair value of the company presented in *Table 2* — EV/EBITDA, EV/S, EV/FCF. We note that when comparing *P* and *EV*, we are not talking in the context of identifying undervalued stocks, but in the context of generating income per share.

From the indicators remaining in the list (Table 2), we will form dynamic standards that will characterize the investment attractiveness of stocks.

The most popular indicator among analysts and investors is the PER or P/E ratio. Its goal is to identify undervalued stocks with upside potential. The formula for calculating the P/E ratio is P/E = Market share price / Earnings per share.

To answer the question of which price/earnings ratio is good or bad, standard P/E values are introduced, which can be set:

1. Absolute values, for example, if P/E > 20, then the company's shares are overvalued.

2. P/E values of a peer company, for example, if the P/E of a company (1) is less than the P/E of a similar company (2), then the company (1) is undervalued compared to the company (2).

3. Average P/E values for the industry, for example, if the P/E of the company (1) is less than the average P/E for the industry, then the company (1) is undervalued compared to the industry as a whole.

The validity of the presented standards is not obvious. For example, if we focus on the first version of the standard, then Apple stocks have been overvalued for several years since P/E starting from 2018 has been either at the border of the standard or above it (*Table 3*).

Given the above standards, Apple stocks were not attractive enough for investors in 2018 — there were many other interesting stocks in the market in terms of P/E. However, Table 3 shows that the market value of stocks has risen by almost 100% in 2020 and by almost 40% in 2021. The situation is exactly the same with the other two benchmarks: there are examples of "undervalued" stocks with low and even declining P/E during a long period of time, as well as "overvalued" stocks with high and rising P/E for a long time. Such "overvalued" stocks have greater investment attractiveness than similar "undervalued" ones. Therefore, a clarification or a different presentation of this criterion is required.

Regardless of which of the three criteria the investors are guided by, they buy stocks that they consider undervalued, hoping for future growth in their value, believing that their P/E will "catch up" with the industry average or similar companies. In other words, after buying stocks, investors expect P/E growth, which, given the popularity of this indicator, should be taken as one of the main criteria for their attractiveness. The growth of P/E implies a faster growth of *P* (market value, capitalization) over *E* (net profit), which can be written as a dynamic standard (7):

$$GR\,(P) > GR\,(E), \qquad (7)$$

where $GR\,(A)$ — the growth rate of indicator *A*;

P — market price per share;

E — company's net profit.

Presentation of the P/E indicator in the form (7) makes it possible not to use its insufficiently substantiated normative values, expressed either in absolute or industry average or in indicators of peer-competitors.

A distinctive feature of the dynamic standard (7) is its focus not on comparison with the stocks of peers, but on assessing the dynamics of the company's own indicators. This means that it is more important to improve one's own results compared to the previous period than to try to meet or beat the industry average or a competitor's performance. There are no absolute analogs, each company is unique in terms of a combination of strengths and weaknesses, the ability to seize opportunities and withstand threats from the external environment, risk appetite, shareholder structure, goals, interests, etc. What is acceptable for one company may not fit for another. For example, one competitor achieved a significant increase in market value through leveraged buybacks. These actions are accompanied by a significant increase in risk, which can lead to a significant deterioration in financial and market results in the future, and even bankruptcy. However, by doing so, the competitor establishes a high level of PER, which is accepted by the market as a benchmark. As a result, market information is distorted, and on the basis of distorted information, conclusions are drawn about the investment attractiveness of certain shares. Also, a regular, albeit relatively small, increase in P/E can result in a much larger net increase in the long term than a significant increase in the short term followed by a decline. In this sense, the dynamic standard (7) makes it possible to assess the attractiveness of stocks with

Fig. 3. **Stock price dynamics of Gazprom, Sberbank, Apple Inc and Johnson & Johnson**
Source: Investing.com. URL: https://ru.investing.com/equities/ (accessed on 31.08.2021).

greater certainty than the traditional interpretation of the P/E ratio.

Similar reasoning can be given for P/S and P/BV. P/S = Market price per share / earnings per share. Corresponding dynamic standard:

$$GR\,(P) > GR\,(S), \qquad (8)$$

where P — market price per share;

S — sales revenue.

P/BV = Market capitalization / book value of equity. The requirement for the growth of this coefficient leads to a dynamic standard (9).

$$GR\,(P) > GR\,(Eq), \qquad (9)$$

where Eq — is equity.

Earlier we substantiated the dynamic standard (10).

$$GR\,(Eq) > GR\,(NSO), \qquad (10)$$

where NSO — the number of shares outstanding.

According to the theory of business valuation, one of the methods for assessing the market value of a company can be a cost method. In accordance with it, the market value of the company is equated to the book value of its equity. Therefore, an increase in equity means an increase in market value and, consequently, an increase in the investment attractiveness of stocks. The requirement for equity growth can be written as a normative ratio (11)

$$GR\,(Eq) > 1. \qquad (11)$$

The growth of equity can occur not only due to its quantitative increase but also due to the acceleration of its turnover. In this case, there is a faster release and, accordingly, the involvement of funds in circulation, which leads to an increase in the company's profit, and therefore, as we noted earlier, to an increase in the market value and investment attractiveness of stocks. The equity turnover rate (ET) is expressed as the equity turnover ratio (12).

$$ET = \frac{S}{Eq}. \qquad (12)$$

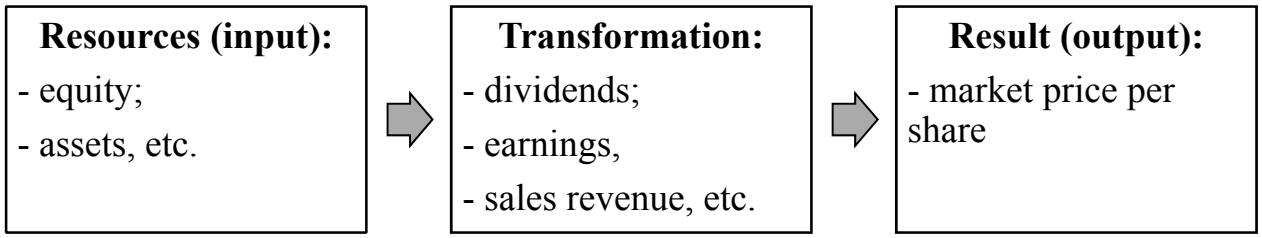


Fig. 4. **Formation of the stock market price**
Source: compiled by the authors.

The acceleration of the turnover of equity means faster growth in sales (*S*) in relation to the growth of equity (*Eq*), which can be written as a dynamic standard (13).

$$GR\,(S) > GR\,(Eq). \qquad (13)$$

Next, we will consider the normative ratio of earnings (*E*) and sales revenue (*S*), which is sometimes called the "golden" rule of the enterprise economy. But we will consider it not in the context of the enterprise economy, but through its impact on the market value and investment attractiveness of stocks. Under the income approach, the market value of a business is defined as the present value of future cash flows, which in turn are calculated based on the company's earnings. The greater the profit, the greater, ceteris paribus, the cash flows, and the higher the valuation of the company, which entails an increase in the investment attractiveness of stocks. Earnings growth means that sales revenue (income) is growing faster than expenses, and this is tantamount to faster growth of earnings in relation to sales revenue (14).

$$GR\,(E) > GR\,(S). \qquad (14)$$

It remains to determine the place of the indicator of paid dividends in the stock rating model in accordance with the principles of dynamic co-subordination and dynamic comparability of indicators.

The question of the impact of dividends on market value is one of the most contentious issues in corporate finance. Thus, in [47], an extensive review of the literature on this issue was carried out, which allowed the authors to identify three different approaches. The first approach states that an increase in dividend payments increases the company's value (price per share). Proponents of the second approach believe that increasing the payment of dividends reduces the value of the firm. The third approach supports the Modigliani-Miller theory that dividend policy does not affect the market value of a company or price per share. To date, no consensus has been reached on this issue and the results are inconclusive.

However, our task in constructing a rating model is not to assess the impact of dividends on the market value of a share, but to assess its attractiveness for an investor, i.e. those properties of the stock that arouse interest, the desire to buy it and allow getting additional benefits.

In this sense, the payment of dividends has the properties necessary in the context of attractiveness, namely, it determines the desire, depending on the amount of declared dividends, to buy or sell shares and provides an opportunity to receive additional income. There is a lot of empirical evidence for this. As an example, let's take the stocks of Gazprom, Sberbank, Apple Inc, and Johnson & Johnson, the price dynamics of which are shown in *Fig. 3*.

The letter "D" in the charts indicates the date of closing the register for receiving dividends (the date of compiling the list of persons entitled to receive dividends).

The stocks of the selected companies belong to different industries, circulate on the markets of different countries, and are

listed on different exchanges (MOEX, Nasdaq, and NYSE). Nevertheless, in each case, the influence of dividends on the interest of investors in the stock can be traced. In the case of Gazprom, Sberbank, and Johnson & Johnson, the amount of declared dividends attracted investors, which was reflected in the increase in the volume of purchases and the corresponding increase in share prices before the register was closed. In the case of Apple, investors' expectations for dividends were initially optimistic — the market price of the stocks was growing, but then the announced dividends disappointed investors to some extent and there is a slight decrease in the stock price before the registration closing date. Earlier we gave a similar example of the negative impact of dividends on the demand for stocks in relation to Tatneft. We have other examples. Consequently, there are numerous confirmations that the dividends paid on a share are a factor in its attractiveness for an investor.

Another argument in favor of including dividends in the calculation of stock ratings. The model we are developing is based on a normative (reference) idea of a stock. In practice, it is customary to distinguish between growth stocks, which provide their owners with a high income from the growth of their market value (capital income), and dividend stocks, which provide their owners with a regularly increasing income in the form of dividends from the company's profit (current income). But a reference, a perfect share should bring to its owner all possible types of income — both capital and current income at the same time i.e. should combine the benefits of both growth stocks and dividend stocks. In the example of the stocks shown in *Fig. 3*, we see that this is achievable and is not an isolated case. Therefore, the indicator of dividend payments will be taken into account by us when forming the model for the rating assessment of stocks.

Illustrating the principle of dynamic co-subordination, we substantiated the normative ratio of the growth rates of dividend and profit indicators (2). Let us write this relation in the accepted notation (15).

$$GR\,(D) \geqslant GR\,(E), \qquad (15)$$

where D — dividends per share;

E — the company's net profit.

We present arguments in favor of the validity of inequality (15).

1. On the one hand, it is necessary to leave a part of the company's profit for development ("get into the pocket of shareholders").

2. On the other hand, the owners of stocks would not like to see more money in their pocket in the current period than in the previous one, i.e. to reduce their current income.

3. However, during a period of declining profits, shareholders must accept that the dividends paid will also decrease.

We write the ratio of the growth rates of the market price per share and dividends in the form of inequality (16).

$$GR\,(P) > GR\,(D), \qquad (16)$$

where P — market price per share.

To justify the validity of inequality (16), let us imagine the formation of the market price per share in the form of a processor (*Fig. 4*).

According to the concept of a dynamic standard, the growth rates of indicators characterizing the result should outpace the growth rates of indicators characterizing the transformation, and those, in turn, the growth rates of indicators characterizing resources. This is the principle of increasing processor performance. This principle formally substantiates inequality (16).

This concludes the list of indicators characterizing the investment attractiveness of stocks, which should form the basis of the rating model.

It is necessary to explain why we did not include the FCF indicator in the stock rating indicators, although more than 10%

of analysts use it to assess the attractiveness of stocks (*Fig. 2*). They consider it a more informative indicator than net income, since the company's shareholders have access (including for the payment of dividends) only to cash, and not to profit reflected in the financial statements. FCF characterizes the amount of cash for the reporting period remaining at the disposal of the enterprise after the implementation of current cash payments and capital expenditures (CAPEX). The formula for calculating the indicator is presented in *Table 2*.

We agree that this is an important indicator for assessing the financial condition of a company, but, in our opinion, it is not suitable for assessing the investment attractiveness of stocks. Dividends are based on net profit, not FCF, although sometimes the shareholders may limit their amount to FCF. This means that a company can have positive free cash flow but no net profit, in which case it will not be able to pay dividends. In addition, hypothetically, and in some cases practically, a company can borrow to pay dividends and thereby pay dividends even in the absence of free cash flow. It should also be taken into account that the payment of dividends for the reporting period occurs much later than its end, and by this time the situation with free cash in the company may change significantly — the company will be able to pay dividends, although at the end of the reporting period it did not have FCF. In view of the foregoing, when forming the rating model for shares, we will take into account the net profit indicator, and not free cash flow.

As a result, generalizing inequalities (7)–(11), (13)–(15) and (16), and also taking into account the transitivity of «>» and «⩾», operators, we will form the desired model for the stock rating (17).

$$\begin{cases} GR(P) > GR(D) \geq GR(E) > GR(S) > GR(Eq) > GR(NSO) \\ GR(Eq) > 1 \end{cases}, \quad (17)$$

where $GR\ (A)$ — (A) is the growth rate of indicator A;

P — market price per share;

D — dividends per share;

E — the company's net profit;

S — sales revenue;

Eq — equity;

NSO — the number of shares outstanding.

The second inequality in the system (17) for all indicators, with the exception of *NSO*, is imputed to growth in dynamics.

The constructed stock rating model (17) has the following features:

1. Forecast and expert estimates are not used, a high level of formalization of the model, due to which the required level of objectivity is achieved.

2. Only achieved and confirmed indicators of official reporting are used. This ensures, firstly, the availability of initial information, and secondly, an increase in the reliability of the rating.

3. The attractiveness of stocks is assessed for compliance with the normative (reference) dynamics of indicators. In essence, model (17) is a benchmark, a reference aggregated indicator, which, when compared, allows assessing the real attractiveness of a particular stock for an investor.

4. The presented stock valuation model makes it possible to compare companies in different industries, different markets, different accounting systems, etc. For example, despite the serious differences between Russian (RAS) and international (IFRS) accounting standards both in accounting rules and in the formation of financial results, the main inequalities (17) are valid for both one and the other standard. In particular, the requirement for faster earnings growth compared to sales revenue growth will be in place for both RAS and IFRS. The same can be said about the other relations used in (17). These ratios are also valid for companies in any industry and any market.

5. It is not required to determine the significance of indicators. The formation of a hierarchy of indicators is a mandatory component of most rating methods. The

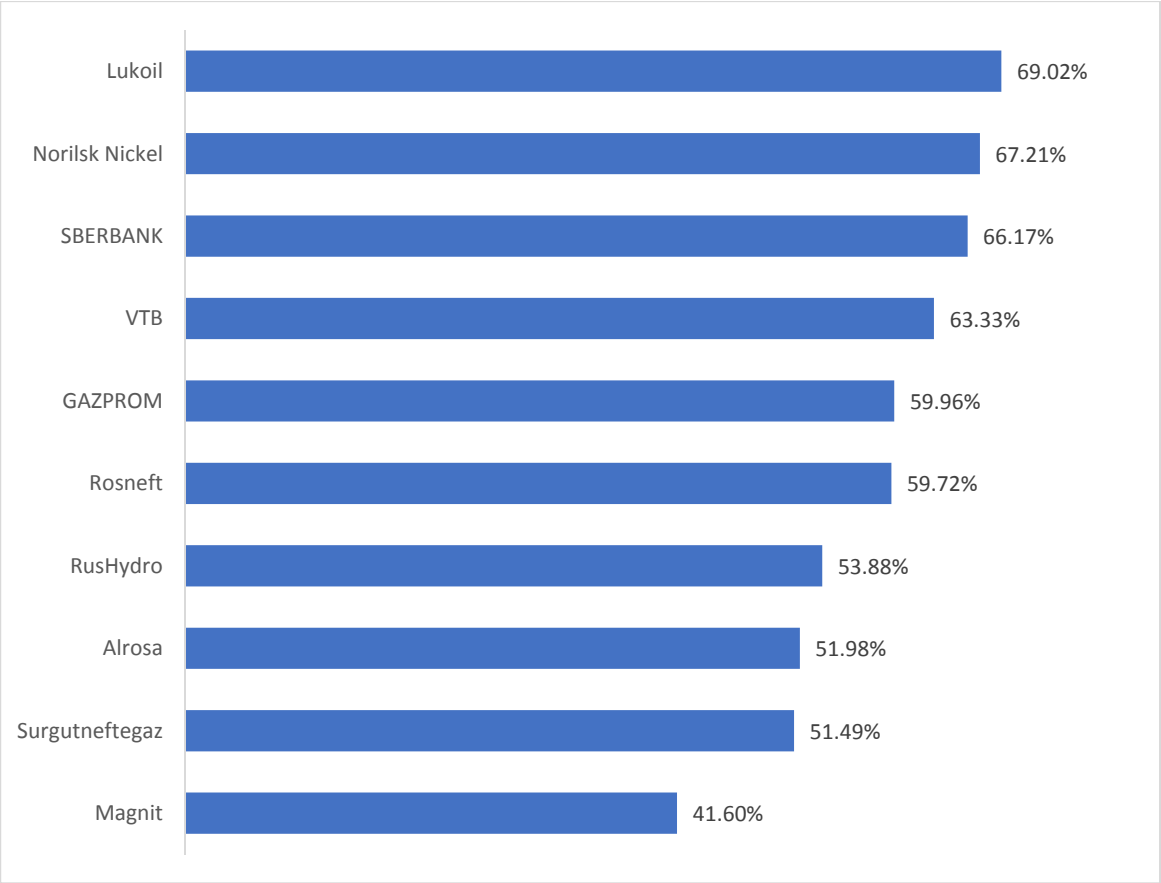


Fig. 5. **Rating of common stocks from the MICEX10 index for a one-year period**

Source: compiled by the authors.

significance of indicators is quantitatively determined on the basis of expert methods. This introduces subjectivity into the assessment process, reduces the transparency of the methodology and the interpretability of the results, and leads to the emergence of various options for constructing a rating score, the advantages of which are not obvious. In addition, the significance of indicators may change over time along with market conditions, business practices, and the external environment of the company, which will lead to a revision of the model and hence to the problem of comparability of ratings over time. Therefore, the absence of the need to determine the significance of the indicators included in the model for evaluating the attractiveness of stocks should be considered a significant advantage of the proposed approach.

6. The criterion for a stock's attractiveness is positive performance relative to its own performance in the previous period, not relative to competitors or industry averages, and, in addition, insufficiently justified absolute standards (for example, that the P/E ratio should be less than 10). The justification for the importance of this thesis is given above. Comparison with stocks of other companies is carried out only by how much the performance of each stock has improved in relation to their own values in the past.

7. A more rigid criterion for the attractiveness of stocks is compliance with the benchmark dynamics and not simple growth or achievement of generally accepted normative values of individual indicators. Presenting the criterion in the form (17) makes it much more difficult to manipulate the market and "embellish" reporting, which makes possible to exclude stocks that grow

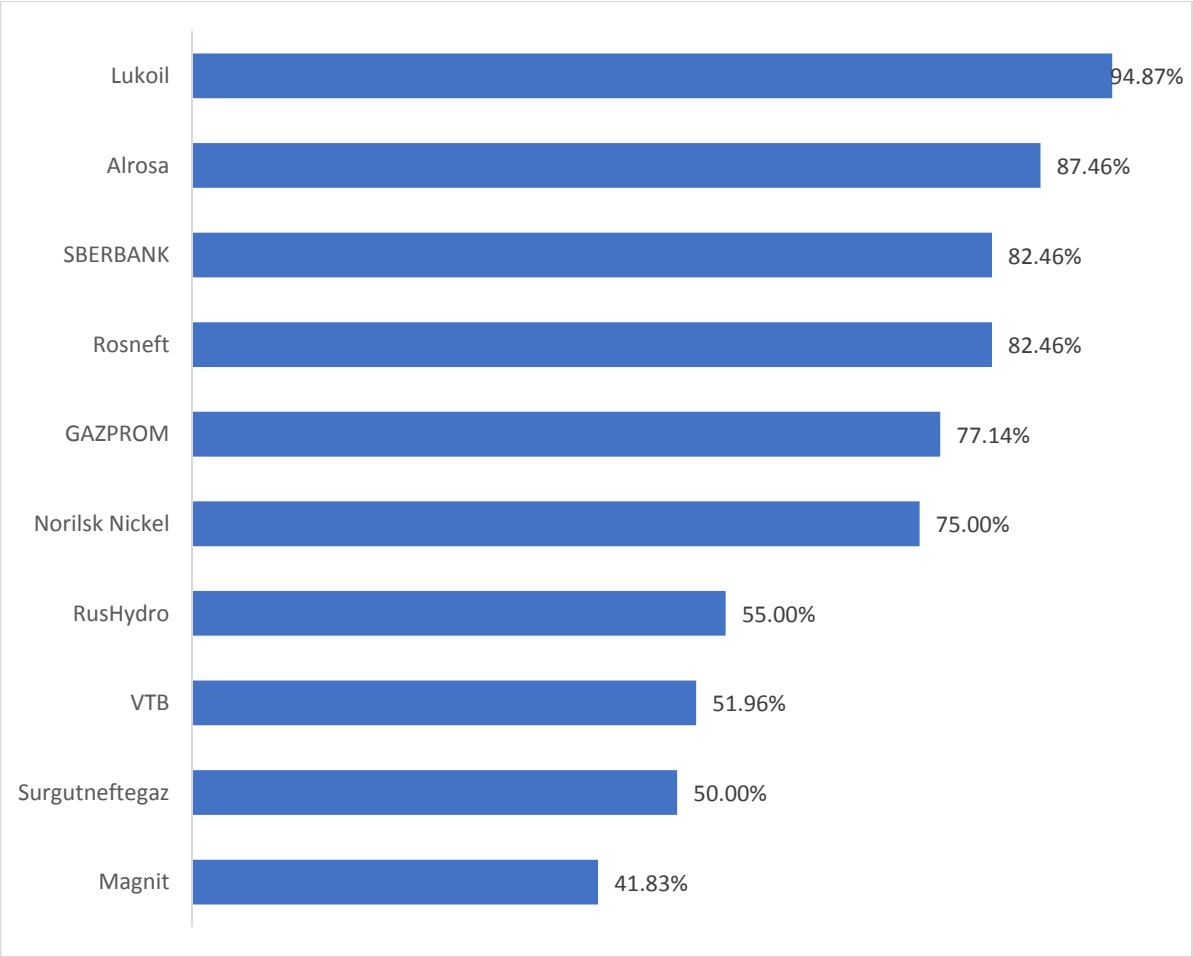


Fig. 6. **Rating of common stocks from the MICEX10 index for a five-year period**

Source: compiled by the authors.

in the short term, the price of which is artificially "accelerated" in order to sell at an overpriced price or increase the size of the bonus.

8. The developed model makes it possible to operate with negative indicators (losses, equity, etc.), although traditionally the growth rate, as well as some indicators of *Table 2* for negative values are not calculated. To do this, it suffices to change the order of indicators in the system of inequalities (17). For example, if in the base period the net profit of a company is negative (in this case, the traditional growth rates cannot be calculated), then either the profit will be positive, or a decrease in the amount of loss in the current period, which is equivalent to the inequality (18):

$$1 > GR'(E), \qquad (18)$$

where $GR'(E)$ — an indicator that replaces the traditional growth rate, but which can be calculated for negative values of the indicators. For convenience, we call it the quasi-tempo. $GR'(E) = E_2/E_1$, where E_2 is net profit in the current period; E_1 is the net profit in the previous period.

Compliance with inequality (18) in the case of $E_1 < 0$ can mean:

1. $E_2 > 0$ and the whole fraction $GR'(E) < 0$, i.e. during the period under review, losses were replaced by profit, which is considered a significant factor in the growth of the attractiveness of stocks of a company.

2. $E_1 < E_2 < 0$ and the whole fraction $0 < GR'(E) < 1$, i.e. in the period under review, there was a decrease in losses, which is also considered a positive factor in the attractiveness of stocks.

Taking into account (18), the benchmark (17) will take the form (19).

$$\begin{cases} GR(Eq) > GR(NSO) \\ GR(P) > GR(D) > GR(S) > GR(Eq) > 1 > GR'(E). \end{cases} (19)$$

By changing the order of the exponents in (17), it is possible to operate with other negative exponents.

Similarly, division by zero can be bypassed if dividends in the base period are zero. The growth of dividends by any amount in the next period is considered a very positive factor in the growth of the attractiveness of stocks, and it can be assigned an arbitrary quasi-rate, the numerical value of which will be the maximum in relation to the growth rates of the indicators that make up model (17). As a result, benchmark (17) will take the form (20).

$$\begin{cases} GR'(D) > GR(P) > GR(E) > GR(S) > GR(Eq) > GR(NSO). \\ GR(Eq) > 1 \end{cases} (20)$$

9. Model (17) is not suitable for evaluating new companies (established less than 2 years ago), since the assessment of the attractiveness of shares is based on the analysis of their dynamics, i.e. by comparing the indicators of the reporting and previous period. In the case of valuation of the shares of a new enterprise of the previous period, there may not be a history of a relatively newly created company. This feature of the model characterizes the limitation of its use.

Thus, we have developed and substantiated a model for rating the company stocks, and also highlighted its features. The next step is the calculation of the rating score.

PRACTICAL APPLICATION

The quantitative calculation of the rating score is carried out on the basis of a measure of the coincidence of the actual order of the growth rates of the calculated indicators with the normative order (standard) (17), (19), or (20), depending on the values of the indicators in the base period. This measure can be calculated in different ways. Since the orders (17), (19), and (20) are non-linear, it is advisable to use the normalized Hamming distance between the matrices corresponding to the normative and actual order of the values of the indicators to calculate the rating score. The procedure for calculating such a measure is given by us in [48]. The calculation of the similarity measure of the normative and actual order based on the Hamming distance makes it possible to set the dimension of the final indicator that is convenient for interpretation. The indicator varies from 0 to 100% and characterizes the percentage of the actual dynamics that coincide with the benchmark. This is the desired rating of a stock that meets our requirements, including the ability to compare stocks of companies in different industries, markets, or with a certain specificity. At the same time, instead of many different indicators of the stock attractiveness— P/E, P/S, P/BV, return on equity, dividend yield ratio, dividend yield — we got one, expressed as a percentage, which will be a generalized assessment of the attractiveness of stocks.

It should also be noted that at time intervals of different lengths, the dynamics of the indicators of the same stock can differ significantly. Since investors purchase stocks for different periods, we consider it expedient to form separate rating estimates for time intervals of various lengths. In addition, to obtain a generalized assessment for a number of consecutive periods, the geometric average of private rating indicators is calculated.

Let us test the developed model on the shares of Russian companies from the MICEX 10 Index based on the results of 2019. The stock rating for a one-year period is shown in *Fig. 5*.

A one-year period means that the previous year is taken as the reference period when determining the actual order of growth rates of indicators; biennium — the year before last; three-year — the previous year before last, etc.

The rating of the Lukoil stocks of 69.02% implies that, on average, the actual dynamics of the stocks' attractiveness indicators annually coincide with the normative (benchmark) by 69%. This is the largest value for stocks shown in *Fig. 5*, so they are the most attractive in terms of compliance with the regulatory dynamics. This statement is true for an investment period of 1 year. If the investment period is more than one year, for example, 5 years, then the rating list will be different from the previous one (*Fig. 6*).

We see that the ratings of stocks at different time intervals can differ significantly, both in terms of rating scores and in terms of the distribution of places in the ranked list. Thus, the Lukoil stocks annually have an average rating of 69.02%, while over a five-year interval the rating reaches a very high level of 94.87%. This means that the attractiveness of the Lukoil stocks with investments for a period of 5 years is much higher than with investments for 1 year. We also notice that some stocks that are the least attractive for investment in the short term become attractive in the long term. For example, the Alrosa stocks are in 7th place out of 10 in the rating for a one-year period with a result of 51.98%, but the situation changes when considering them for a 5-year period — they are already in 2nd place with a fairly high rating of 87.46%.

We also note that 8 stocks out of 10 in the list improve their ratings in the long term relative to the short term. At the same time, their relative attractiveness may decrease. For example, the rating of the Norilsk Nickel stocks in the transition from a one-year interval to a five-year interval increases from 67.21 to 75.00%, while their relative attractiveness decreases — the stocks take 2nd place in the ranked list for a one-year period and only 6th for a five-year period.

The higher rating of stocks in the long term is explained by the smoothing of the volatility of the Russian stock market, which in some short-term periods can lead to a significant deterioration in the rating, as in 2017 against the backdrop of an annual market decline of more than 5%. This is another confirmation that company stocks are a long-term investment tool.

It is possible to significantly expand the list of rated stocks both at the expense of stocks remaining on the Russian market and foreign stocks. The presented approach makes it possible to do this.

CONCLUSIONS

Ratings occupy an important informational niche in the analysis of the investment attractiveness of stocks. The ratings, by increasing information transparency, simplify the assessment and adoption of investment decisions in the stock markets and increase their efficiency. However, the use of ratings is constrained by the subjective component in determining the significance of the indicators formed by the rating, as well as in predicting their target values, which necessitates the replacement of expert assessments with objective characteristics. A large number of studies have been carried out in this area, which in their own way solve this problem.

In this study, the increase in the objectivity and informativeness of the stock ratings is decided on the basis of the principles of dynamic co-subordination and dynamic comparability of estimated indicators. The application of these principles made it possible to present traditional indicators of attractiveness of stocks in a new way, to propose specific criteria for their evaluation and an original method of processing.

The developed model was tested on the example of the Russian stocks from the MICEX 10 index. As a result, a list of 10 stocks was compiled, ranked according to the integral characteristic of the investment attractiveness of the selected securities, depending on the expected investment period. The estimates obtained have a high level of objectivity, solidity, dynamism, interpretability, and comparability. Comparability means that the presented algorithm makes it possible to compare

the investment attractiveness of stocks of companies in different industries, scales of activity, markets, and countries without modification.

We believe that our model will be useful both as a primary and as an additional stock analysis tool for long-term fundamental investments. Its application is not focused on obtaining immediate benefits, but on a long-term and sustainable improvement in investment results. The present study was carried out in this context.

An increase in the level of algorithmization of evaluation process in the stock market determines the constant development of this area, so the search for the most relevant ratings will continue. In the future, the results of this study will be tested in other markets. In particular, ratings are currently being formed for a wider range of Russian stocks, as well as US stocks from the S&P500 index. The results will be posted on the Alternative Stock Rating website (URL: http://alteratingstock.biz/currently under construction).

Thus, we have presented a working methodology for the process of investment decision-making in the stock market, which is quite easy to interpret and put into practice, consolidating various investor ideas about the investment attractiveness of stocks. Approbation of the methodology showed that this approach has the potential for widespread use both together with existing assessment systems and independent use.

REFERENCES

1. Adamczyk M. The impact of ratings and other information on the fluctuation of Polish stock indexes. *Bank i Kredyt*. 2020;51(3):239–262. URL: https://bankikredyt.nbp.pl/content/2020/03/BIK_03_2020_02.pdf
2. Ishigami S., Takeda F. Market reactions to stock rating and target price changes in analyst reports: Evidence from Japan. *Journal of International Financial Markets, Institutions and Money*. 2018;52:134–151. DOI: 10.1016/j.intfin.2017.09.014
3. Hong H., Kubik J.D. Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance*. 2003;58(1):313–351. DOI: 10.1111/1540–6261.00526
4. Lim T. Rationality and analysts' forecast bias. *The Journal of Finance*. 2001;56(1):369–385. DOI: 10.1111/0022–1082.00329
5. Chan C.-Y., Lo H.-C., Su Y.-R. Distribution of stock ratings and analyst recommendation revision. *The North American Journal of Economics and Finance*. 2014;28:273–286. DOI: 10.1016/j.najef.2014.03.004
6. Ashour S., Hao G.Q. Do analysts really anchor? Evidence from credit risk and suppressed negative information. *Journal of Banking & Finance*. 2019;98:183–197. DOI: 10.1016/j.jbankfin.2018.11.006
7. Zhang C., Qian A., Shi X. The effect of affiliated analysts on stock recommendations: Evidence from share pledges in China. *China Journal of Accounting Research*. 2020;13(1):79–107. DOI: 10.1016/j.cjar.2020.02.001
8. Twedt B., Rees L. Reading between the lines: An empirical examination of qualitative attributes of financial analysts' reports. *Journal of Accounting and Public Policy*. 2012;31(1):1–21. DOI: 10.1016/j.jaccpubpol.2011.10.010
9. Clement M.B., Tse S.Y. Financial analyst characteristics and herding behavior in forecasting. *The Journal of Finance*. 2005;60(1):307–341. DOI: 10.1111/j.1540–6261.2005.00731.x
10. Xue H. Independent and affiliated analysts: Disciplining and herding. *The Accounting Review*. 2017;92(4):243–267. DOI: 10.2308/accr-51631
11. Asghar M.Z., Rahman F., Kundi F.M., Ahmad S. Development of stock market trend prediction system using multiple regression. *Computational and Mathematical Organization Theory*. 2019;25(5):271–301. DOI: 10.1007/s10588–019–09292–7
12. Bilev N.A. Modelling of stock market security price dynamics using market microstructure data. *Finansy: teoriya i praktika = Finance: Theory and Practice*. 2018;22(5):141–153. (In Russ.). DOI: 10.26794/2587–5671–2018–22–5–141–153

13. Meher B.K., Hawaldar I.T., Spulbar C., Birau R. Forecasting stock market prices using mixed ARIMA model: A case study of Indian pharmaceutical companies. *Investment Management and Financial Innovations*. 2021;18(1):42–54. DOI: 10.21511/imfi.18(1).2021.04
14. Sharma S., Kaushik B. Quantitative analysis of stock market prediction for accurate investment decisions in future. *Journal of Artificial Intelligence*. 2018;11(1):48–54. DOI: 10.3923/jai.2018.48.54
15. Alzheev A.V., Kochkarov R.A. Comparative analysis of ARIMA and lSTM predictive models: Evidence from Russian stocks. *Finance: Theory and Practice*. 2020;24(1):14–23. DOI: 10.26794/2587–5671–2020–24–1–14–23
16. Singh J., Khushi M. Feature learning for stock price prediction shows a significant role of analyst rating. *Applied System Innovation*. 2021;4(1):17. DOI: 10.3390/asi4010017
17. Asquith P., Mikhail M.B., Au A.S. Information content of equity analyst reports. *Journal of Financial Economics*. 2005;75(2):245–282. DOI: 10.1016/j.jfineco.2004.01.002
18. Gosiewska A., Kozak A., Biecek P. Simpler is better: Lifting interpretability-performance trade-off via automated feature engineering. *Decision Support Systems*. 2021;150:113556. DOI: 10.1016/j.dss.2021.113556
19. Avramov D., Cheng S., Lior M. Machine learning versus economic restrictions: Evidence from stock return predictability. *SSRN Electronic Journal*. 2019. DOI: 10.2139/ssrn.3450322
20. Harvey C.R., Liu Y., Zhu H. ...and the cross-section of expected returns. *The Review of Financial Studies*. 2016;29(1):5–68. DOI: 10.1093/rfs/hhv059
21. Cochrane J.H. Discount rates. *The Journal of Finance*. 2011;66(4):1047–1108. DOI: 10.1111/j.1540–6261.2011.01671.x
22. Karminsky A., Polozov A. Handbook of ratings: Approaches to ratings in the economy, sports, and society. Cham: Springer-Verlag; 2016. 356 p. DOI: 10.1007/978–3–319–39261–5
23. Digo S.N., Sokolova A.M. Shaping the stock portfolio by the investment rating method. *Vestnik Rossiiskogo ekonomicheskogo universiteta imeni G.V. Plekhanova = Vestnik of the Plekhanov Russian University of Economics*. 2018;(1):75–89. (In Russ.). DOI: 10.21686/2413–2829–2018–1–75–89
24. Badokina E.A., Maksimenko O.I. The rating approach to assessing the dividend policy effectiveness of Russian public companies. *Korporativnoe upravlenie i innovatsionnoe razvitie ekonomiki Severa: Vestnik Nauchno-issledovatel'skogo tsentra korporativnogo prava, upravleniya i venchurnogo investirovaniya Syktyvkarskogo gosudarstvennogo universiteta = Corporate Governance and Innovative Economic Development of the North: Bulletin of Research Center of Corporate Law, Management and Venture Investment of Syktyvkar State University*. 2018;(4):36–46. (In Russ.).
25. Gorbunova N.A. Using the methodology of rating evaluation of risk and return of securities public company. *Kazanskii sotsial'no-gumanitarnyi vestnik = The Kazan Socially-Humanitarian Bulletin*. 2016;(5):13–16. (In Russ.).
26. Pyatnitsky D.V. Rating and valuation of investment attractiveness of companies. *Izvestiya vysshikh uchebnykh zavedenii. Seriya: Ekonomika, finansy i upravlenie proizvodstvom = News of Higher Educational Institutions. Series: Economy, Finance and Production Management*. 2018;(1):52–60. (In Russ.).
27. Samsonov A.V., Vasiliev I.V. Comparative analysis of investment attractiveness of oil and gas industry enterprises in Russia. *Moskovskii ekonomicheskii zhurnal = Moscow Economic Journal*. 2020;(2):50. (In Russ.). DOI: 10.24411/2413–046X-2020–10088
28. Vygodchikova I. Toolkit of decisions making about investment of large Russian companies using hierarchical procedure of ranking and minimax approach. *Prikladnaya informatika = Journal of Applied Informatics*. 2019;14(6):123–137. (In Russ.). DOI: 10.24411/1993–8314–2019–10054
29. Mallikarjuna M., Rao R.P. Evaluation of forecasting methods from selected stock market returns. *Financial Innovation*. 2019;5:40. DOI: 10.1186/s40854–019–0157-x
30. Różycki M. Inertia in assessing the possibilities of economic development: Limits in modelling economies. In: Sarfraz M., Adbullah M.I., Rauf A., Shah S.G.M., eds. Sustainable management practices. London: IntechOpen; 2019:65–76. DOI: 10.5772/intechopen.86351

31. Seth H., Talwar S., Bhatia A., Saxena A., Dhir A. Consumer resistance and inertia of retail investors: Development of the resistance adoption inertia continuance (RAIC) framework. *Journal of Retailing and Consumer Services*. 2020;55:102071. DOI: 10.1016/j.jretconser.2020.102071
32. Zhuravsky Yu.A. Economic inertia and acceleration in the methodology of economic dynamics. *Ekonomika i upravlenie innovatsiyami = Economics and Innovation Management*. 2019;(4):20–34. (In Russ.). DOI: 10.26730/2587–5574–2019–4–20–34
33. Gunitsky S. Rival visions of parsimony. *International Studies Quarterly*. 2019;63(3):707–716. DOI: 10.1093/isq/sqz009
34. Becchio G. The two blades of Occam's razor in economics: Logical and heuristic. *Economic Thought*. 2020;9(1):1–17. URL: https://core.ac.uk/download/pdf/326909303.pdf
35. Cohen M.R., Nagel E. An introduction to logic and scientific method. Germantown, MD: Hughes Press; 2013. 480 p.
36. Mamchur E.A., Ovchinnikov N.F., Uemov A.I. The principle of simplicity and measures of complexity. Moscow: Nauka; 1989. 304 p. (In Russ.).
37. Opolev P.V. The logical principle of simplicity in sciences about complexity. *Vestnik Omskogo universiteta = Herald of Omsk University*. 2014;(4):87–90. (In Russ.).
38. Anikin A.V., Yashina N.I., Kashina O.I., Pronchatova-Rubtsova N.N. The methodical aspects of analysis of the dependence of the gross regional product of high-tech branches from financial factors in the conditions of knowledge economy. *Fundamental'nye issledovaniya = Fundamental Research*. 2019;(7):7–12. (In Russ.).
39. Allakhverdov V.M., Karmin A.S., Shilkov Yu.M. Introduction to the heading "History and philosophy of science". *Metodologiya i istoriya psikhologii = Methodology and History of Psychology*. 2007;2(1):227–230. (In Russ.).
40. Syroezhin I.M. Improving the system of indicators of efficiency and quality. Moscow: Ekonomika; 1980. 192 p. (In Russ.).
41. Tret'yakova E.A., Osipova M. Yu. Evaluation of sustainable development indicators for regions of Russia. *Studies on Russian Economic Development*. 2018;29(2):124–134. DOI: 10.1134/S 1075700718020144 (In Russ.: *Problemy prognozirovaniya.* 2018;(2):24–35.).
42. Polyanskaya I.G., Yurak V.V. Balanced natural resource management of a region: Estimation by dynamic normal technique. *Ekonomika regiona = Economy of Regions*. 2018;14(3):851–869. (In Russ.). DOI: 10.17059/2018–3–12
43. Biryukova V.V. Production system management based on a balanced development model. *IOP Conference Series: Materials Science and Engineering*. 2020;753:062014. DOI: 10.1088/1757–899X/753/6/062014
44. Pogostinskaya N.N., Pogostinskiy Yu.A., Vlasova M.S. Measuring the strategy for socio-economic development of the Russian Arctic zone. *Arktika: ekologiya i ekonomika = Arctic: Ecology and Economy*. 2019;(1):21–33. (In Russ.). DOI: 10.25283/2223–4594–2019–1–21–33
45. Kazakova N.A., Kogdenko V.G. Monitoring of the main parameters of environmental safety industrial production. *Ekologiya i promyshlennost' Rossii = Ecology and Industry of Russia*. 2021;25(3):60–65. (In Russ.). DOI: 10.18412/1816–0395–2021–3–60–65
46. Fernandez P. Valuation using multiples: Dispersion. Useful to compare and to negotiate. *SSRN Electronic Journal*. 2019. DOI: 10.2139/ssrn.274972
47. Raza H., Ramakrishnan S., Gillani S., Ahmad H. The effect of dividend policy on share price: A conceptual review. *International Journal of Engineering and Technology*. 2018;7(4):34–39. DOI: 10.14419/ijet.v7i4.28.22386
48. Tonkikh A.S., Ionov A.V., Ionov S.A. Methodological tools of increase of firm market value and the compliance of the balance of interests. *Audit i finansovyi analiz = Audit and Financial Analysis*. 2013;(2):280–319. (In Russ.).

ABOUT THE AUTHORS

Andrei S. Tonkikh — Dr. Sci. (Econ.), Assoc. Prof.; Prof., Admiral Ushakov State Maritime University, Novorossiysk, Russia
https://orcid.org/0000-0001-9085-5691
Corresponding author
andrew.tonkih@gmail.com

Svetlana A. Tonkikh — Cand. Sci. (Econ.), Lecturer, Admiral Ushakov State Maritime University, Novorossiysk, Russia svetlana.tonkih@gmail.com
https://orcid.org/0000-0001-6358-8364
svetlana.tonkih@gmail.com

Elena Yu. Maslova — Lecturer, Kuban State University Novorossiysk Branch, Novorossiysk, Russia maslovaey@mail.ru
https://orcid.org/0000-0003-1559-3251
maslovaey@mail.ru

Authors' declared contribution:
A. S. Tonkikh — statement of the problem, development of the conceptual framework of the article, critical analysis of the literature, substantiation of the research methodology.
S. A. Tonkikh — model development and validation.
E. Yu. Maslova — collection and analysis of data, description of the results, formation of research conclusions.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 17.09.2021; revised on 07.10.2021 and accepted for publication on 25.02.2022.
The authors read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-110-128
UDC 336.27(045)
JEL C20, E62, H63, H74, H81

Debt Sustainability Assessment of Regional Budgets

E.G. Kiseleva
Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia

ABSTRACT

The **aim** of the paper is to assess the debt sustainability of the budget system of the regions of the Northwestern Federal District (NWFD) of Russia and establish differentiated values of the budget constraints. The author uses **methods** of comparison and grouping, correlation and cluster analysis. The author's methodology based on the methods of multivariate statistical analysis made it possible to include indicators that have the nature of leading indicators in the assessment, evaluate the integral indicator of debt sustainability, and determine the limit values of these indicators. Approbation of the author's methodology for assessing the debt sustainability of the budget system on the statistics of the NWFD regions allowed dividing the regions into three clusters according to the types of debt sustainability and determining the threshold values of indicators for each cluster. The research results substantiate the need to reduce the high debt burden for 70% of the NWFD regions. The results indicate that the regions with a high level of debt sustainability include St. Petersburg, Leningrad and Kaliningrad regions. The integral indicator of debt sustainability of these regions is greater than one and the regions are well grouped according to six indicators of debt sustainability into a group that is homogeneous in terms of characteristics. The Republic of Komi, the Republic of Karelia, and the Arkhangelsk and Pskov regions have a low level of debt sustainability: the standard limiting values for the "debt-to-GRP ratio" indicator of the regions is less than 5%, and the "debt-to-revenue ratio" indicator — less than 42%, the value of all indicators of debt sustainability exceeds the values of the cluster centroids. For these regions, the author recommends pursuing a targeted budget policy with a mandatory debt reduction to the level of threshold values for the indicators of the regions of this cluster. The paper **concludes** that it is necessary to reduce the debt burden of most regions of the NWFD, as well as to establish differentiated values of budget constraint on public debt, taking into account the indicators of socio-economic development of these regions.

Keywords: debt sustainability; budget system; public debt; indicators; debt burden limit; econometric methods; GDP; economic growth; the Northwestern Federal District

For citation: **Kiseleva E.G. Debt sustainability assessment of regional budgets.** ***Finance: Theory and Practice.*** **2022;26(3):110-128. DOI: 10.26794/2587-5671-2022-26-3-110-128**

INTRODUCTION

The issues of managing the debt burden by imposing budgetary restrictions are relevant for different countries. The use of debt financing, on the one hand, increases investment activity and social and economic objectives, on the other hand, inevitably increases credit risk and the likelihood of default. Therefore, issues of assessment of debt sustainability of the budget system of the Russian Federation are a subject of discussion and are actively discussed by the world scientific community. In Russia, the volume of public debt has been steadily increasing over the past ten years, while the statistical validity of the limits of the debt burden indicators of the budgets of the constituent entities of the Russian Federation remains open. The limits of the state internal debt of the budgets of the constituent entities of the Russian Federation are regulated by the Budget Code of the Russian Federation, but this practice does not take into account the differentiation of indicators of socio-economic development. In this regard, the justification of limits of indicators of debt burden of the budget system of regions on the basis of systematization and processing of statistical data on subjects of the Russian Federation using classification methods seems to be an interesting, promising and relevant task.

ANALYSIS OF THE STATUS AND TRENDS OF THE DEVELOPMENT OF THE STATE DEBT OF THE RUSSIAN FEDERATION AND REGIONS OF THE NORTHWESTERN FEDERAL DISTRICT

According to the data of the Ministry of Finance of the Russian Federation and the Federal Service of State Statistics, the volume of external and internal public debt tends to increase. This dynamic is well evident in the use of debt financing in the form of securities *(fig. 1)*.

The *figure 1* show that the Ministry of Finance of the Russian Federation borrowed significantly over the last 20 years, with the second decade growing faster than the first. The strongest growth occurred between 2012 and 2020. The increase in the issue of government securities is a consequence of the liberalization of the debt market, modernization of market infrastructure, increased market transparency, simplification of the mechanism for the acquisition of securities, etc. The increased informativeness of securities transactions has made the Russian debt market more attractive and accessible to both domestic and foreign investors. A gradual decline in the yield of the government bond portfolio, which, on the one hand, means an increase in the issuer's debt rating and, on the other hand, a reduction in the cost of debt servicing.

The size of the Russian government's domestic debt also tends to increase [1]. The composition of the state internal debt of the constituent entities of the Russian Federation is dominated by credits of special organizations and international financial organizations, as well as other budgets of the budget system of the Russian Federation, the volume of government securities in the debt structure is insignificant.

At the same time, the main direction of the use of domestic public debt is to cover the deficit of regional budgets. The regional budget deficit is a consequence of the decline in budget revenues as a result of the 2014 crisis, after the introduction of the policy of economic sanctions. In addition, experts cite changes in the tax legislation of the Russian Federation with regard to the payment of taxes by taxpayers in the consolidated group as the reason for the decline in revenues, which reduced the income of many regions [1, p. 97]. The regional budget deficit is also the result of the increase in social expenditures related to the implementation of the May decrees of the President of the Russian Federation. All this has encouraged regional authorities to resort to a policy of borrowing money, which contributes to the growth of the public debt as well as the cost of its servicing. At the

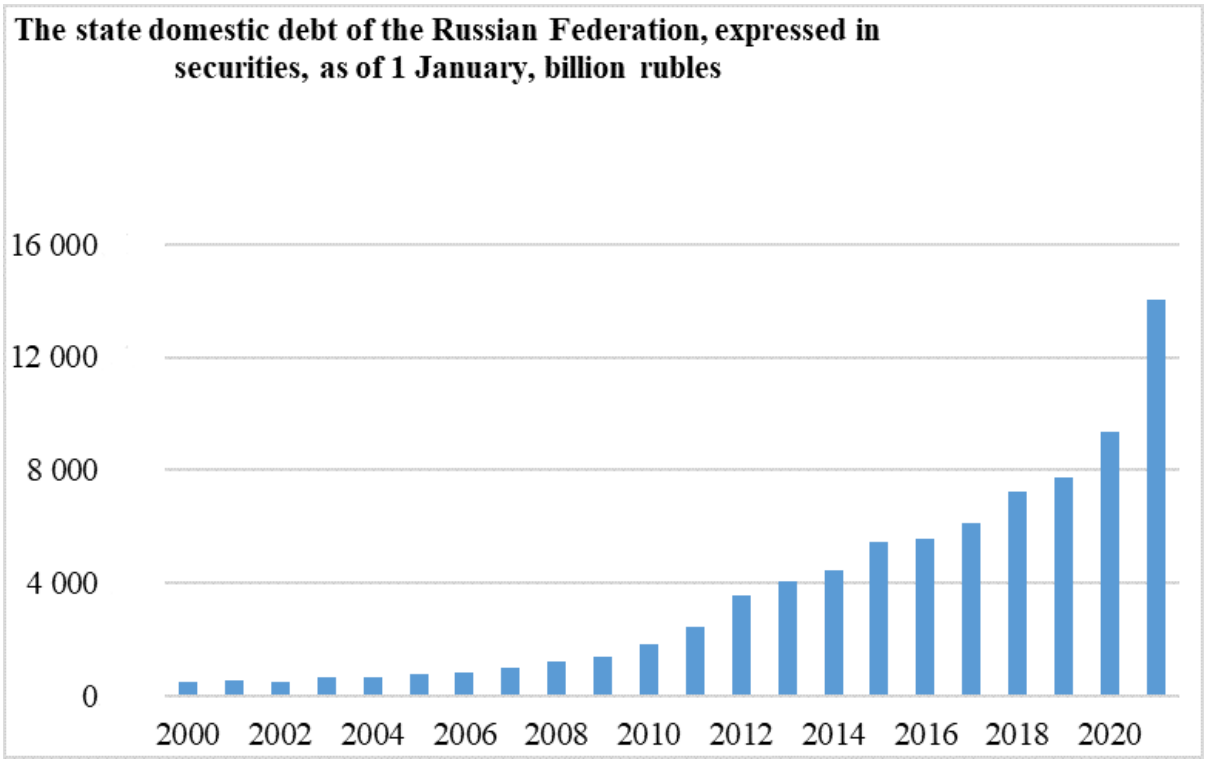


Fig. 1. **Dynamics of the volume of the government internal debt of Russia expressed in government securities**

Source: compiled by the author based on the Ministry of Finance of the Russian Federation. URL: https://minfin.gov.ru/ru/perfomance/public_debt/internal/structure/duty/ (accessed on 20.04.2021).

same time, it is important to note that the regional policy provides for the refinancing of loans of the constituent entities of the Russian Federation through the use of the federal budget (budget credits) at the rate of 0.1% (according to the Ministry of Finance of the Russian Federation in 2020, 31% of the budget deficit was financed).[1]

The use of indicative approach in assessment of the state and dynamics of the development of the public debt of the Russian Federation showed the growth of values of key indicators (*table 1*). The selection of key indicators is based on the results of analysis of previous research of Russian and foreign scientists [2–5]. *Table 1*.

One of the world's main indicators 'debt-to-GDP ratio', calculated on the internal debt of the constituent entities of the Russian Federation, increased to 3.92% in 2020, and on the gross debt of the Russian Federation increased to 25.18%. Note that the calculations made by the author on the indicator of GDP at comparable prices show a high debt burden in comparison with official data of the Ministry of Finance of the Russian Federation. At the same time, analysis of the level of the debt burden for the countries of the world allows us to assert that, despite the growth of debt dependence, Russia has a sustainable level of public debt in GDP.

Japan, the US, the Eurozone stand out for traditionally high public debt. Since the country's GDP is one of the basic macroeconomic indicators and is used for international comparisons of the level of welfare of different countries, the indicator 'debt-to-GDP ratio' is the main global indicator of the state's debt burden. Looking at the evolution of country borrowing over the past twenty years, it is safe to say that

[1] The rules for granting budget loans to the regions have been clarified: press centre of the Ministry of Finance of the Russian Federation. 2021. URL: https://minfin.gov.ru/ru/press-center/?id_4=37435-utochneny_pravila_predostavleniya_regionam_byudzhetnykh_kreditov (accessed on 05.06.2021).

Table 1

Indicators of the financial system's debt sustainability of the Russian Federation (as of the end of the year)

Indicator	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Share of public debt of constituent entities of the Russian Federation in GDP, %	1.78	1.86	2.35	2.84	3.51	3.54	3.94	3.61	3.52	3.92
Share of the state domestic debt of the Russian Federation. expressed in government securities in GDP*, %	5.88	6.50	6.97	8.55	8.92	9.79	13.34	13.13	15.55	25.18
The volume of public debt of the Russian Federation in GDP **, %	–	–	10.6	13.2	13.5	13.2	14.6	14.9	15.3	19.1
The volume of public debt of the Russian Federation in GDP ***, %	5.86	7.54	11.87	11.78	16.49	17.57	20.44	21.49	22.61	36.55
Share of public and municipal debt servicing costs in GDP,%	0.44	0.51	0.57	0.65	0.83	1.00	1.30	1.37	1.22	1.40
Share of debt servicing costs in the total amount of domestic debt of the Russian Federation, %	5.69	6.12	6.08	5.70	6.68	7.48	7.55	8.16	6.39	4.83
Ratio of the state debt of the subject of the Russian Federation to the annual volume of budget revenues, %	9.43	9.06	11.46	12.57	16.04	16.41	14.19	10.95	10.47	11.68
The ratio of public domestic debt expressed in government securities to the export of goods and services, % ****	21.80	24.82	26.25	28.01	26.13	31.36	34.63	27.27	34.81	59.37
Ratio of public domestic debt expressed in government securities to international reserves, %	16.31	22.78	23.57	22.13	22.87	22.21	24.10	24.56	22.02	22.38

Source: compiled by the author based on the Federal State Statistics Service and Ministry of Finance of the Russian Federation.

Notes: * government securities;

** based on data from the Ministry of Finance (The main directions of the state debt policy of the Russian Federation for 2017–2019, Ministry of Finance of the Russian Federation. 2017. 63 p.);

*** based on the author's calculations of GDP in 2011 prices;

**** the export volume is calculated in rubles at the exchange rate at the end of the corresponding year.

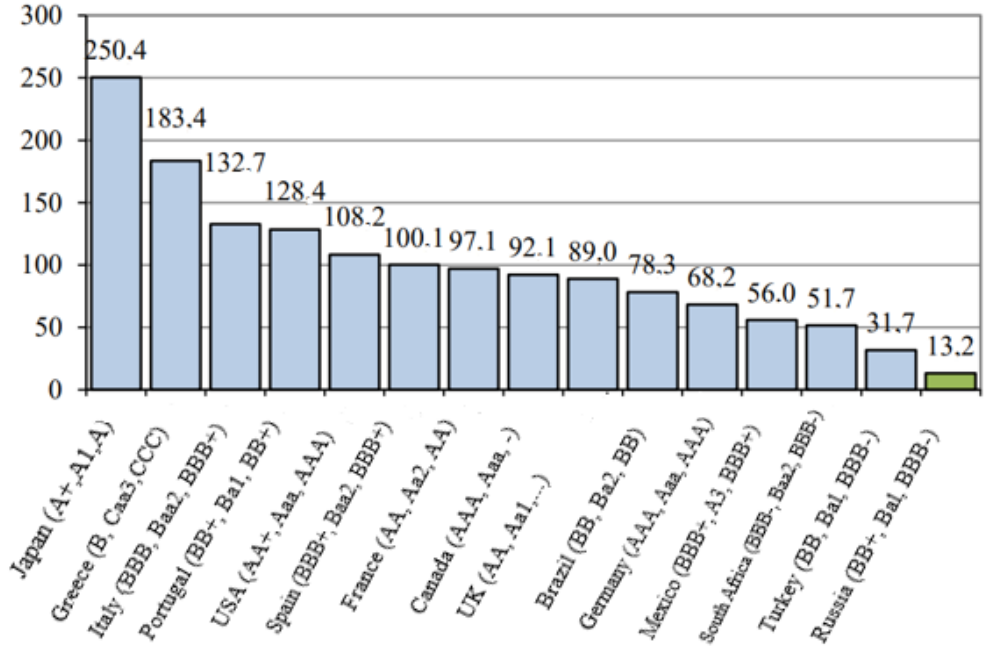


Fig. 2. **Debt burden and credit ratings of countries in 2016**

Source: The main directions of the state debt policy of the Russian Federation for 2017–2019, Ministry of Finance of the Russian Federation. 2017. 63 p.

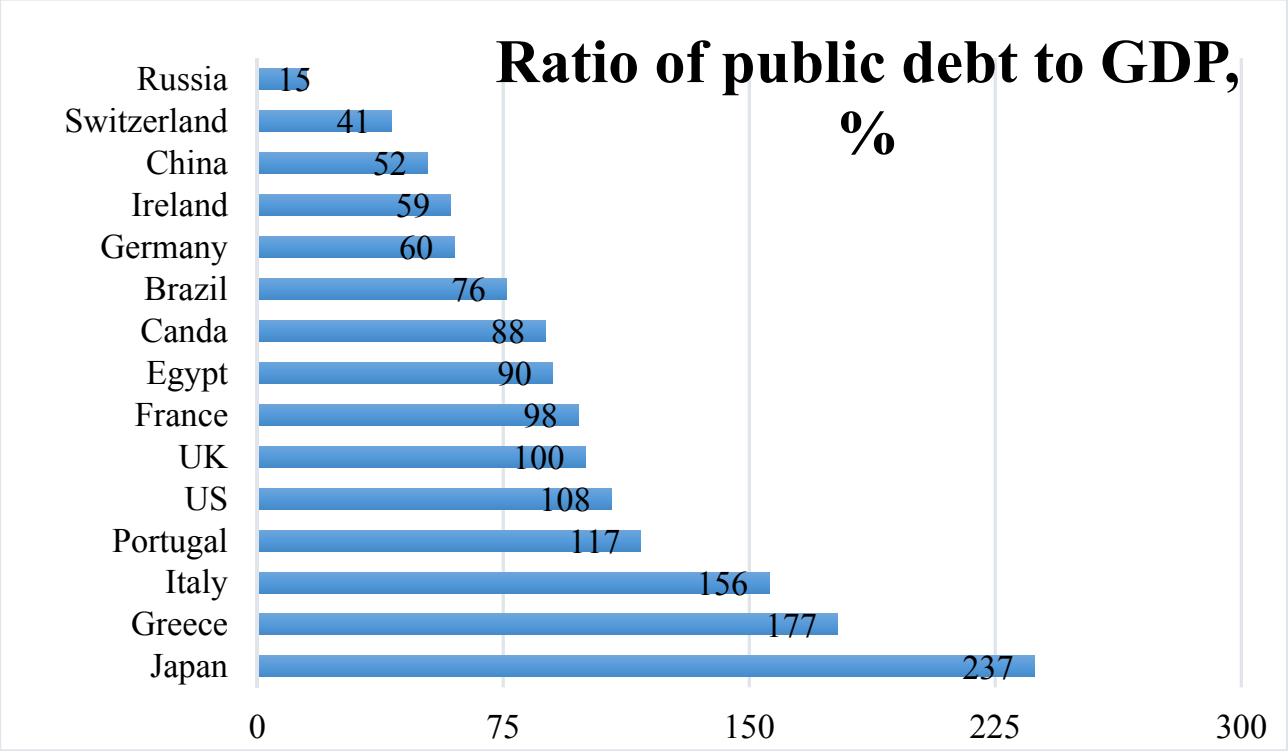


Fig. 3. **Debt-to-GDP ratio by countries in December 2020**

Source: compiled by the author based on Trading Economics. URL: https://ru.tradingeconomics.com/country-list/government-debt-to-gdp (accessed on 01.05.2021).

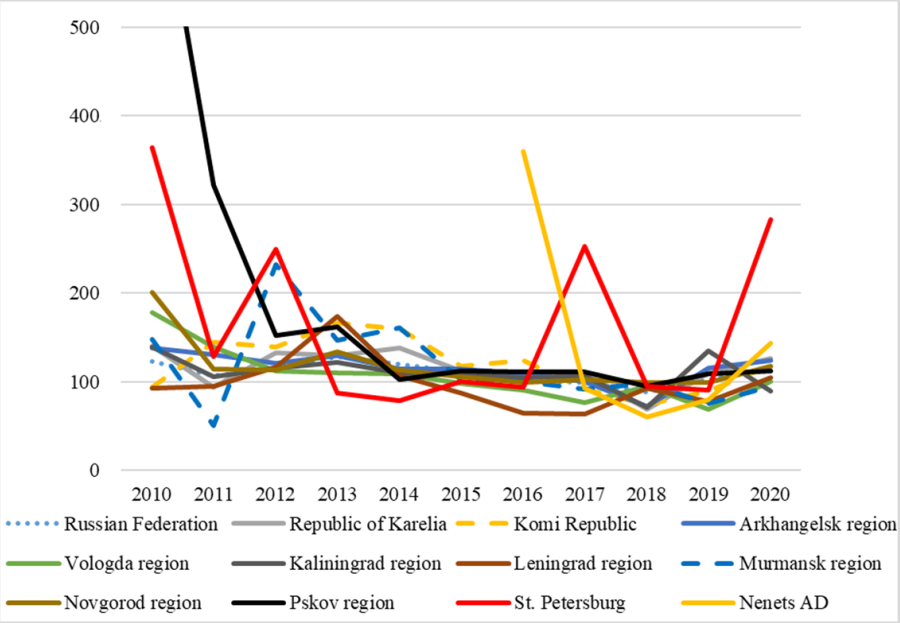


Fig. 4. **Growth rates of the government debt of the subjects of the Northwestern Federal District and Russia**

Source: compiled by the author based on the Ministry of Finance of the Russian Federation. URL: https://www.minfin.ru/ru/perfomance/public_debt/subdbt (accessed on 25.04.2021).

global debt has increased. Thus, in 2012, the Japanese government's gross debt to GDP was 361%, Italy — 258%, U.S. — 252%, Brazil — 129%, Greece — 150%, India — 73%, France — 284%, Russia — 65%. In 2015, debt accounted for 365% of Japan's GDP, and Greece — 299%, Italy — 270%, U.S. — 249%, France — 292%, Mexico — 76%, Russia — 89%.[2] Some countries' borrowings in 2016 can be found at *fig. 2*.

In 2021, as a result of the economic crisis caused by the COVID-19 epidemic, global debt rose to 89.6 trillion USD, from 83.5% to 97.6% of global GDP. Japan, Greece, Italy, Portugal, and U.S. became the leading countries in terms of public debt to GDP in 2020 (*fig. 3*).[3]

More serious concerns about the growth of debt dependence should be related to the increase in the indicator' debt in exports. This indicator is often used in studies of domestic and foreign scientists to characterize the processes of debt financing of the economy, because it demonstrates a country's relative long-term ability to accumulate foreign exchange earnings without pressure on its balance of payments balance [2, p. 127]. After a slight decline to 27.3% in 2018, the indicator increased to 38% in 2019, this may be due to a decline in export volumes due to the prolongation of the agreement between the OPEC+ countries to limit oil production on the background of an increase in borrowing by the Ministry of Finance of the Russian Federation. The growth of the indicator in 2020 occurred due to the global world lockdown in the spring of 2020, this significantly reduced the volume of not only Russian exports, but also global trade turnover in general. World trade, according to one UN organisation (ESCAP), decreased by 14.5%, exports to Russia fell by 27%.[4]

The regions of the North-Western Federal District (NWFD) also show a rising trend in public domestic debt. At the same time,

[2] According to the research centre McKinsey Global Institute. URL: www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/visualizing-global-debt (accessed on 23.05.2021).

[3] According to the information and analytical portal 'World Finance'. URL: http://global-finances.ru/gosdolg-mira-2021 (accessed on 20.03.2021).

[4] According to financial and analytical news portal Investing.com. URL: https://ru.investing.com/news/economy/article-2024987 (accessed on 20.03.2021).

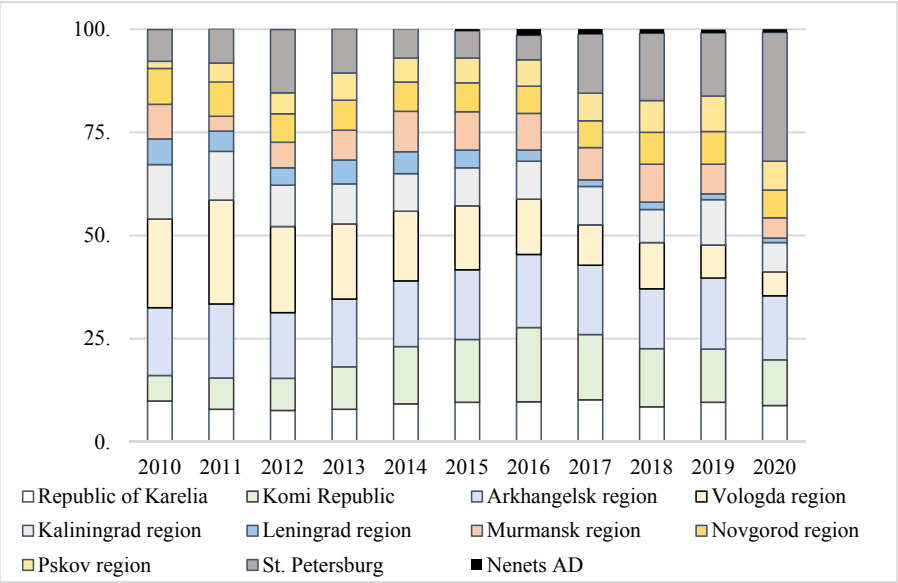


Fig. 5. **Structure of the government debt by the subjects of the Northwestern Federal District**

Source: compiled by the author based on the Ministry of Finance of the Russian Federation. URL: https://www.minfin.ru/ru/perfomance/public_debt/subdbt (accessed on 25.04.2021).

the level of the debt burden of the North-Western regions is growing more rapidly: the average growth rate for the last ten years, including 2020, is 112.2%, compared to the value of the indicator in the Russian Federation — 108.4% (*fig. 4*).

The Pskov oblast and the city of St. Petersburg (average growth rate for the last 10 years is 29%), as well as the Republic of Karelia (18.9%) and the Arkhangelsk oblast (10.8%) use debt financing more actively than other regions. In two North-Western regions there is a small trend of debt reduction: Vologda and Leningrad oblasts (growth rates of 98.4 and 94.3%, respectively).

Analysis of the dynamics and structure of debt by Northwestern regions allows us to assert that Pskov, Arkhangelsk, Vologda regions, Komi Republic, as well as the city of St. Petersburg are more active in raising public debt. In 2020, St. Petersburg became the leader in the volume of public debt (31.2%) among the regions, the second and third place are the Arkhangelsk region and the Komi Republic. At the same time, extrapolation of the results leads to the conclusion that regions with a high growth rate of public debt do not always have a greater share in the structure of district debt. This confirms once again the need to study in detail the debt sustainability of the budgetary system of regions, as well as the relevance of defining safe lines of the debt burden according to statistics.

ANALYSIS OF APPROACHES TO DEBT SUSTAINABILITY ASSESSMENT

A number of works of domestic and foreign authors are devoted to debt sustainability of systems [6–9]. Among foreign researches, the relationship between economic growth and public debt is of great importance. Such works are based on extensive empirical data from different countries (China, Eurozone countries, BRICS countries, Great Britain, Malaysia, etc.) for a period of 20 to 50 years and have high scientific validity of the results obtained [10–12]. Indeed, there are a number of scientific works that prove the significant impact of public debt on sustainable economic growth [10, 11, 13]. At the same time, there are works that claim that there is no evidence of such regularity [14], as well as works where the level of debt burden is statistically substantiated, in which there is a negative correlation between public debt and the country's GDP [15, 16].

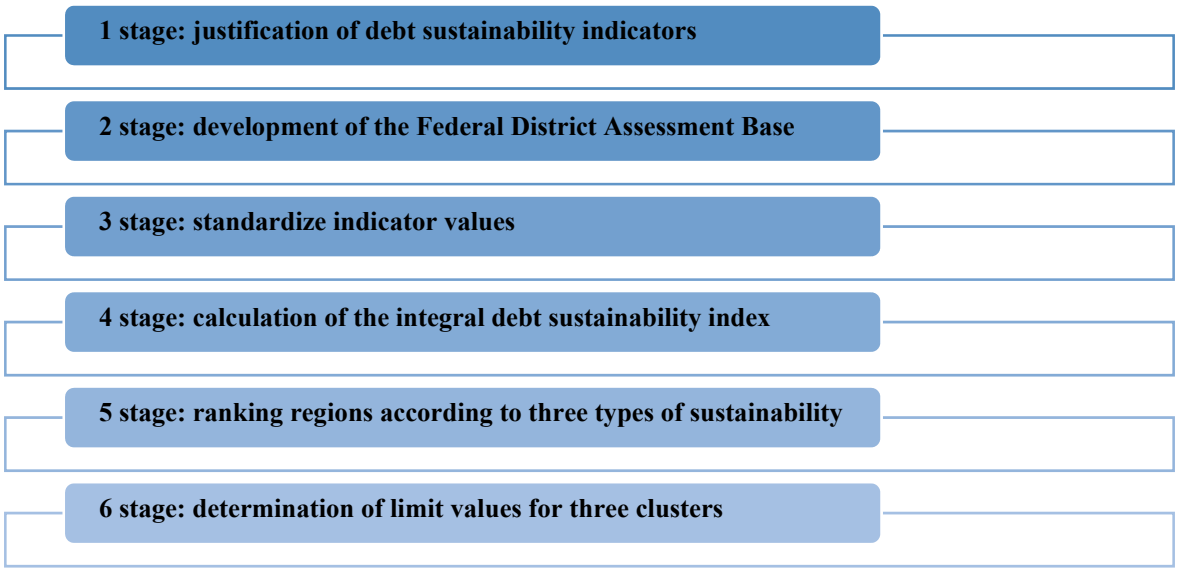


Fig. 6. **Methodology for the debt sustainability assessment of Russian regions**
Source: compiled by the author.

In this context, the results of the study on the sustainability of the budget policy of Austrian municipalities, which proved the effectiveness of using municipal debt limits to reduce the budget deficit on the basis of an adaptive version of the Bon stability testing method [9]. In study by A. Chudik, K. Mohaddes et al. statistically significant thresholds of the debt burden of budgets for countries with growing debt were obtained and the need to reduce the country's debt burden to a sustainable level was confirmed [17].

A comparative analysis of scientific articles by Russian scientists shows that a set of indicators (relative indicators) on the basis of which such an assessment is carried out plays a paramount role in issues of debt sustainability. At the same time, researchers systematized indicators and their criteria values. Thus, in the work of E.V. Rogatenyuk is given description of indicators of debt security of the Russian Federation based on the analysis of different sources of scientific literature [2, c. 125]. In total, about 20 indicators have been allocated, the concepts of the indicator of debt sustainability and debt security of the country have been defined, the importance of indicative analysis of debt security of Russia and the need for statistical evaluation of thresholds has been emphasized. In the work of A.V. Kalina, I.P. Savelieva noted that the "optimal approach to diagnostics of economic security is the use of the method of indicative analysis" [18, c. 15]. The indicators are grouped by production, financial and socio-demographic factors of regional development. In total, 47 indicators with thresholds are given in the work. At the same time, A.A. Kurilova, A.N. Kiryushkina notes that «parameters for optimising the structure of public debt are not widely represented in the literature» [19, p. 157], and S.N. Soldatkin emphasises the need to expand the debt sustainability indicators of Russian regions [20, 21]. Other authors emphasise the need to assess debt sustainability through the indicators established in the Budget Code of the Russian Federation, which significantly narrows the set of indicators [2, 22]. The methodology of the Ministry of Finance of the Russian Federation is based on these indicators.[5] According to the developed classification, the regions of the Russian

[5] Rules for the assessment of debt sustainability of the constituent entities of the Russian Federation: approved by the Decree of the Government of the Russian Federation from 04 March 2020 No. 227.

Table 2

Correlation matrix of indicator values

	Public debt	Budget revenues	Export	Budget expenditures	Debt repayment	Population size	GRP
Public debt (Y)	1.0000						
Income (X1)	0.9991	1.0000					
Export (X2)	0.9989	0.9999	1.0000				
Costs (X3)	0.9990	1.0000	0.9999	1.0000			
Repayment (X4)	0.9971	0.9976	0.9970	0.9976	1.0000		
Population(X5)	0.9995	0.9998	0.9998	0.9998	0.9969	1.0000	
GRP(X6)	0.9989	0.9999	1.0000	0.9999	0.9970	0.9998	1.0000

Source: compiled by the author.

Note: the null hypothesis is rejected for the significance level $\alpha = 0.05$; $t_r > t_{tabl}$.

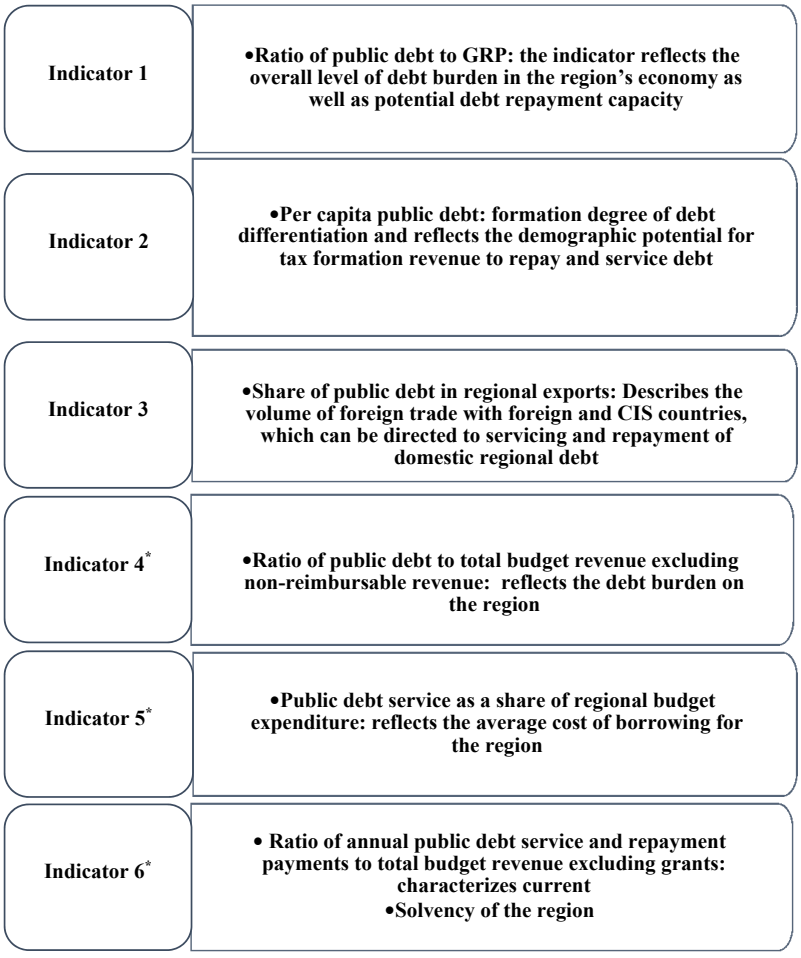


Fig. 7. **Composition of the region's debt sustainability indicators**

Source: compiled by the author.

Note: * referred to indicators of debt sustainability of the Budget Code of the Russian Federation.

Table 3

Partial research results: Indicative analysis of debt sustainability of the Northwestern Federal District

Period	Регионы СЗФО / NWFD Regions											
	NWFD	RKA	RKO	AO	VO	KO	LO	MO	NO	PO	SPB	NAO
Indicator 1: Ratio of public debt to GRP, %												
2016	3.08	10.04	6.39	9.36	7.17	5.82	1.12	5.15	6.6	9.84	0.44	0.44
2017	3.01	9.92	6.7	8.76	4.68	5.41	0.41	4.29	6.24	10.81	0.91	1.28
2018	2.25	6.15	4.3	5.71	3.89	3.5	0.33	3.86	5.98	9.48	0.79	0.66
2019	2.1	5.82	3.53	6.08	2.49	4.19	0.23	2.28	5.68	8.64	0.59	0.48
Indicator 2: The amount of public debt per capita. rub./people												
2016	16.7	36.1	49.2	36.6	26.2	21.8	34.6	27.1	24.9	22.9	2.63	81.8
2017	17.50	39.99	45.58	36.65	20.19	22.80	2.17	25.12	25.85	25.61	6.55	75.63
2018	14.5	27.8	34.2	26.5	19.3	16.2	1.9	24.8	26	24.6	6.2	45.54
2019	14.1	30.7	30. 9	31	13.5	21.6	1.5	18. 9	25.9	27.1	5.6	36.4
Indicator 3: Share of public debt in regional exports, %												
2016	12.03	51.9	87.97	30.44	18.58	27	2.02	12.68	27.43	409.12	1.48	0
2017	11.35	41.74	79	30.59	14.74	32.26	1.28	9.61	28.01	491.86	3.13	0
2018	9.01	160.13	241.63	218.23	64.85	3.21	1.46	90.59	61.74	94.46	2.43	0
2019	6.81	31.08	46.3	22.15	6.76	26.08	0.71	5.41	16.7	277.42	1.93	4.01
Indicator 4: Ratio of public debt to total budget revenue excluding budget expenditures, %												
2016	23.92	79.35	63.14	67.4	54.44	50.91	4.83	31.2	55.34	74.91	2.98	27.53
2017	23.55	90.41	50.44	61.88	37.94	50.98	3.11	28.89	59.12	78.3	6.95	17.98
2018	14.72	37.32	37.35	35.99	171.68	13.98	2.27	25.58	66.04	131.87	5.63	8.6
2019	15.21	53.91	28.17	41.65	18.67	39.02	1.71	15.88	51.01	71.58	4.85	7.63
Indicator 5: Share of state and municipal debt servicing expenses in regional budget expenditures, %												
2016	1	2.9	2.6	1.37	1.63	0.38	1.42	1.34	3.3	3.02	0.1	1.8
2017	1.06	2.52	6.1	1.89	1.06	0.29	1.52	0.8	1.57	2.8	0.04	1.55
2018	1.22	1.77	3.75	1.41	0.61	0.23	3.51	0.77	1.08	2.23	0.4	2.83
2019	1.37	1	4.61	0.67	0.16	0.21	5.5	0.9	0.97	2.22	0.35	2.57
Indicator 6: Ratio of annual payments for servicing and repayment of public debt to total budget revenues excluding budget expenditure, %												
2016	14.6	25.6	55.47	51.8	21.03	28.57	1.53	26.09	22.3	59.25	0.32	28.16
2017	17.66	50.54	55.87	69.01	17.36	23.49	2.72	34.91	20.57	69.88	1.85	20.16
2018	20.35	47.46	44.87	102.51	134.46	22.98	3.52	65.38	42.88	162.67	1.2	26.84
2019	19.36	64.77	14.71	108.78	15.79	24.19	6.02	68.93	16.69	80.67	0.37	7.12

Source: author's calculations.

Note: RKA — Republic of Karelia, RKO — Komi Republic, AO — Arkhangelsk Region, VO — Vologda Region, KO — Kaliningrad Region, LO — Leningrad Region, MO — Murmansk Region, NO — Novgorod Region, PO — Pskov Region, SPB — Saint Petersburg, NAO — Nenets Autonomous Okrug, BP — gratuitous receipts.

Table 4

Results of the debt sustainability assessment for the Northwestern Federal District

NWFD	Standardized values of indicators												Debt Sustainability Index, I_{DS}			
	2019						2018									
	I2	I3	I3	I4	I5	I6	I2	I2	I3	I4	I5	I6	2019	2018	2017	2016
Republic of Karelia	0.04	0.05	0.02	0.06	0.16	0.01	0.05	0.07	0.01	0.06	0.13	0.03	0.24	0.27	0.22	0.19
Republic of Komi	0.07	0.24	0.02	0.04	0.03	0.03	0.08	0.06	0.01	0.06	0.06	0.03	0.32	0.23	0.19	0.18
Arkhangelsk oblast	0.04	0.12	0.03	0.09	0.24	0.01	0.06	0.07	0.01	0.06	0.16	0.01	0.36	0.28	0.26	0.25
Vologda oblast	0.09	0.09	0.11	0.04	1	0.02	0.08	0.10	0.02	0.01	0.38	0.01	0.88	0.42	0.45	0.32
Kaliningradskaya oblast	0.05	0.12	0.03	1	0.76	0.02	0.09	0.12	0.45	0.16	1.00	0.05	1.55	1.36	1.32	0.53
Leningradskaya oblast	1	0.44	1	0.11	0.03	0.06	1.00	1.00	1.00	1.00	0.07	0.34	2.20	3.78	3.45	2.53
Murmansk oblast	0.1	0.16	0.13	0.03	0.18	0.01	0.09	0.08	0.02	0.09	0.30	0.02	0.44	0.42	0.45	0.41
Novgorod oblast	0.04	0.13	0.04	0.02	0.16	0.02	0.06	0.08	0.02	0.03	0.21	0.03	0.29	0.32	0.32	0.27
Pskov oblast	0.03	0.69	0.01	0.35	0.07	0.005	0.03	0.08	0.02	0.02	0.10	0.01	0.82	0.20	0.55	0.19
St. Petersburg	0.39	1	0.37	0.22	0.46	1	0.42	0.32	0.60	0.40	0.58	1.00	3.22	2.97	3.41	5.70
Nenets AO	0.48	0.04	0.18	0.01	0.06	0.05	0.50	0.04	0.00	0.26	0.08	0.04	0.68	0.74	0.37	0.41
Average value	0.21	0.28	0.18	0.18	0.29	0.11	0.22	0.19	0.2	0.2	0.28	0.14	1.00	1.00	1.00	1.00

– Three maximum values of indicators;

– Three minimum values of indicators

Source: author's calculations.

Table 5

Scale of values of the regional debt sustainability index

Value range by year				Type of sustainability
2016	2017	2018	2019	
⩾ 1	⩾ 1	⩾ 1	⩾ 1	Higher
< 1; ⩾ 0.32	< 1; ⩾ 0.32	< 1; ⩾ 0.42	< 1; ⩾ 0.44	Medium
< 0.32	< 0.32	< 0.42	< 0.44	Low

Source: author's calculations.

Federation, based on the values of indicators, are in one of three groups by the type of debt sustainability: with a high, medium and low level. Since the procedure for assessing the debt sustainability of budgets of regions, established in the framework of the legislation, presupposes the same thresholds for all subjects of the Russian Federation, that definition of allowable limits on the debt burden of regional budgets by systematising and processing statistical data using classification methods is an interesting, promising and urgent task.

ASSESSMENT OF DEBT SUSTAINABILITY OF THE BUDGET SYSTEM OF THE NORTHWEST REGIONS

Assessment of debt sustainability of the budget system of the regions of the Russian Federation is carried out using the author's methodology, which includes several stages (*fig. 6*). The developed methodology is a set of stages and methods of estimation, as well as a certain algorithm of their application: involves the selection and systematisation of debt sustainability indicators, standardisation of their values based on the Euclidean distance method, calculation of the integral indicator of debt sustainability and its ranking using the formula of multidimensional average, and the division of regions into three debt sustainability groups with the definition of limit values (threshold values, limits) for each group.

At the first stage, indicators of debt sustainability were systematized, as well as their selection using the method of correlation analysis were substantiated. Initially, about 15 indicators were selected, which are the most common in the scientific literature, and are also among the indicators of the Budget Code of the Russian Federation. Further, on the basis of data from statistical indicators for North-West regions for the period from 2010 to 2019, a correlation analysis was carried out in order to establish a close relationship between the amount of public debt and macroeconomic indicators characterizing the state of the budgetary system of the region (*table 2*).

According to the results of the correlation analysis, a strong link between the amount of public debt of the Russian Federation entities and absolute indicators, which in most cases are included in the calculation of indicators of debt sustainability, which allows to substantiate their composition (*fig. 7*).

In the second phase of the study, based on open data from the Federal Service of State Statistics and the Ministry of Finance of the Russian Federation, six debt sustainability indicators were calculated for 11 North-Western regions from 2010 to 2019. A fragment of the results obtained is given in *table 3*.

Analysis of indicator values suggests that the two regions stand out from the other low level of indicators, which allows them to be previously classified as regions with high debt sustainability: St. Petersburg and Leningrad Oblast. High values of indicators

for the analyzed period are most often observed in the Republic of Karelia, the Republic of Komi, Vologda, Arkhangelsk and Pskov regions. Therefore, in the future it can be expected that these entities of the Russian Federation will be classified among the regions with a low level of debt sustainability.

In the third stage of the study, the values of the indicators were standardized. The Euclidean distance method was used to account for the degree of difference of each indicator for federal districts. The procedure of the chosen method is widely described in the scientific literature [23]. Since all six indicators are inverse, standardization has been applied to the minimum value of the corresponding year indicator. This resulted in a set of standardized values for six indicators for the period from 2010 to 2019, which in the fourth phase of the research was averaged over a dynamic series using the multivariate mean formula. This approach made it possible to identify regions with low/high debt sustainability for each indicator and to reflect in the integral assessment the level of each region in comparison with the average value of NWFD (*table 4*).

Standardised values of indicators reflect the high level of stability of St. Petersburg, Leningrad and Kaliningrad regions, which naturally affects the value of the integral stability rating of these regions. Regions with a low level of debt sustainability include the Republic of Karelia, the Republic of Komi and the Pskov region. The remaining regions of the NWFD at this stage of the research can be classified as regions with medium debt sustainability.

At the fourth stage of the study, the integrated debt sustainability index for the regions of the NWFD was calculated. The use of the multidimensional average formula, which is often used to assess the level of investment potential and takes into account not only the ranking of feature values, but also correlates it with the average by region in the estimated year, allows you to determine regions whose level of stability is above average. St. Petersburg, Leningrad and Kaliningrad regions have been steadily in such regions for four years. In order to gradate the regions under study by three types of debt sustainability, it is necessary to set intervals for such an assessment: regions whose debt sustainability index is greater than or equal to one are classified in a group with high debt sustainability; the subsequent gradation was carried out taking into account the differentiation of annual values within the two groups (*table 5*).

As can be seen from the results obtained for the period from 2016 to 2018, the group of regions with low debt sustainability consistently includes the Republic of Karelia and the Republic of Komi, as well as the Pskov region, which allows us to conclude about the high risk of insolvency and the need to establish targeted budget restrictions on debt obligations. Novgorod and Arkhangelsk regions are also in a credit risk zone, as they fall into a group of regions with low debt sustainability in 2019 and 2017.

In order to determine borrowing standards (upper limits of restrictions) for each region of the Northwestern Federal District, a hierarchical cluster analysis was carried out at the fifth stage of the study, as it is one of the methods of multidimensional classification, which allows you to distinguish areas of accumulation of objects from this population and combine them into homogeneous groups (segment) [24, p. 9]. Using the distance matrix of indicators by Northwestern Federal District for each analysed period, it is necessary to divide the data set into three clusters, and then define cluster centroids for each indicator. This will confirm the results obtained at the previous stage of the study, as well as solve the problem of targeted statistical justification of the norms of budgetary restrictions on the debt burden. The method of intergroup communication was used as a clustering method, as a measure of similarity between objects — Euclidean distance. The analysis

Table 6

Agglomeration schedule and cluster membership by Debt-to-GRP ratio to regions

Order of agglomeration (clusters)

Stage	Unified cluster		Rate	Cluster first appearance stage		Next stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	7	11	2,915	0	0	2
2	7	12	5,615	1	0	11
3	2	4	6,446	0	0	6
4	6	9	12,821	0	0	8
5	3	8	18,167	0	0	7
6	2	10	28,969	3	0	10
7	1	3	39,715	0	5	9
8	5	6	40,297	0	4	9
9	1	5	81,158	7	8	10
10	1	2	140,984	9	6	11
11	1	7	311,931	10	2	0

was carried out using the SPSS software product. The results of clustering by one of the indicators are presented in *table 6, 7.*

When performing cluster analysis, three clusters are forcibly specified, the value of the agglomeration order coefficient shows that this number coincides with the difference in the number of observations and the number of steps, after which the coefficient increases abruptly. Thus, it can be concluded that the observation array is well clustered into three groups. At the same time, the cluster belongs to a particular type of stability can be judged on the basis of the values of cluster centrifuges.

Table 7 shows that the value of cluster centroids is changing, but the amplitude of oscillations is low. This indicates a change in the cluster centroids of the indicator over the years under the influence of the macroeconomic situation in the country, which determines the level of income and expenditure parts of regional budgets. The importance of the indicator is also affected by fiscal policy in terms of debt management.

Clustering

Monitoring	Cluster 3
1: North-West DF	1
2: Republic of Karelia	2
3: Komi Republic	1
4: Arkhangelsk region	2
5: Vologda region	1
6: Kaliningrad region	1
7: Leningrad region	3
8: Murmansk region	1
9: Novgorod region	1
10: Pskov region	2
11: St. Petersburg	3
12: Nenets AD	3

Source: author's calculations.

Table 7

Cluster centroids by Debt-to-GRP ratio for the period from 2011 to 2019

Average Linkage (Between Groups)	2011	2012	2013	2014	2015	2016	2017	2018	2019
1	3.9750	4.5000	5.5850	6.2550	5.7017	5.6300	5.0550	3.9633	3.3783
2	5.5133	6.7500	8.8533	9.5933	9.7467	9.8567	9.8300	7.1133	6.8467
3	.4267	.6000	.7500	.7000	.6667	.8200	.8667	.5933	.4333
Total	3.4725	4.0875	5.1933	5.7008	5.4542	5.4842	5.2017	3.9083	3.5092

Source: author's calculations.

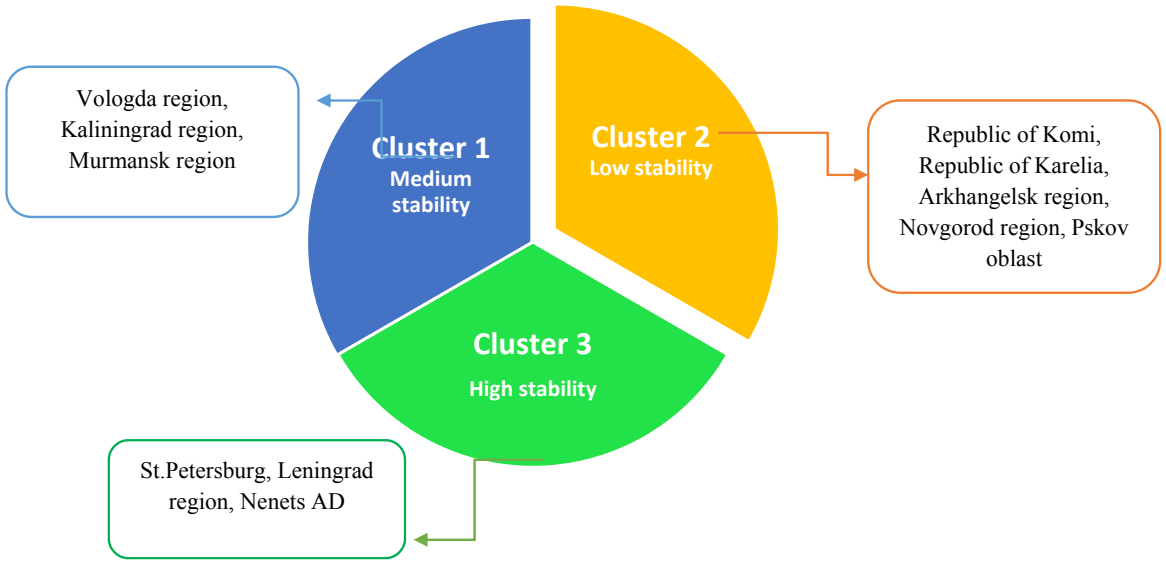


Fig. 8. **Clusters by type of debt sustainability**

Source: compiled by the author.

Using the data of dispersion analysis and R-square, we can say that all indicators of the table are significant, R-square in 2014, 2015, 2019 is more important, so when determining the boundaries of debt sustainability groups it is necessary to focus more on the values of these years (tables of descriptive statistics are not given due to the cumbersomeness of The values of the centroids of the third cluster determine the lower permissible boundary for the regions of the group, which according to the indicator I1 for the Komi Republic, Vologda, Leningrad Murmansk and Novgorod regions is no more than 10%.

In the same way, the results of clustering according to other indicators were analysed. Analysis of the agglomeration order, cluster membership, cluster centroid report, dispersion table allows us to conclude that the results of clustering regions for each indicator are heterogeneous, and regions change their belonging to clusters depending on the year being analysed.

Table 8

Results of the study of clusters by regions of the Northwestern Federal District and volume of budget constraints

Low debt sustainability cluster	Budget constraints for regions with low debt sustainability
Indicator 1: ratio of public debt to GRP, %	
2019	
Komi Republic, Republic of Karelia, Arkhangelsk region, Novgorod region, Pskov region, Nenets AO	> 5.04
2018	
Komi Republic, Karelia Republic, Arkhangelsk Region, Murmansk Region, Pskov Region	> 5.01
Indicator 2: the amount of public debt per capita, rub./people	
2019	
Komi Republic, Republic of Karelia, Arkhangelsk Region, Novgorod Region, Pskov Region, Nenets AO	30 337
2018	
Nenets AO	45 544
Indicator 3: share of public debt in regional exports, %	
2019	
Komi Republic, Republic of Karelia, Arkhangelsk Region, Novgorod Region, Pskov Region, Nenets AO	66.28
2018	
Komi Republic, Karelia Republic, Arkhangelsk Region, Murmansk Region, Pskov Region	104.87
Indicator 4: ratio of public debt to total budget revenue excluding BE, %	
2019	
Komi Republic, Republic of Karelia, Arkhangelsk Oblast, Novgorod Oblast, Pskov Oblast, Nenets AO	42.33
2018	
Komi Republic, Karelia Republic, Arkhangelsk Region, Murmansk Region, Pskov Region	59.39
Indicator 5: share of state and municipal debt servicing expenditures in regional budget expenditures, %	
2019	
St. Petersburg, Leningrad region	2.92
2018	
Komi Republic, Karelia Republic, Arkhangelsk Region, Murmansk Region, Pskov Region	2.83
Indicator 6: ratio of annual payments for servicing and repayment of public debt to total budget revenues excluding BE, %	
2019	
Komi Republic, Republic of Karelia, Arkhangelsk Region, Novgorod Region, Pskov Region, Nenets AO	48.79
2018	
Nenets AO	51.5

Source: compiled by the author.

Therefore, the identification of types of debt sustainability and determination of threshold values of debt obligations of Northwestern Federal District regions was based on the results of clustering indicators for 2019 and 2018. Generalisation of the results of cluster analysis made it possible to determine the current affiliation of the regions to the three clusters (*fig. 8*) and assess the type of debt sustainability, as well as to determine the marginal debt burden standards (*table 8*).

Analysis of the data of *table 8* allows us to draw two important conclusions: firstly, the results of clustering are consistent with the results of the assessment of the integral debt sustainability index; secondly, belonging to clusters of regions changes over the years. *Table 8* also shows that regions such as the Republic of Karelia, the Komi Republic, Arkhangelsk and Pskov regions are combined into a cluster with high indicators, which allows them to be classified as regions with low debt sustainability. Of particular interest are the values of indicators I4, I5, I6, as the upper limits of these indicators are regulated by the Budget Code of the Russian Federation.

CONCLUSION

The analysis of the state and development of debt financing of the budget system of the Russian Federation showed that over the past ten years the volumes of internal public debt of the constituent entities of the Russian Federation tend to moderately increase (7.4%). At the same time, the main direction of use of domestic public debt is the coverage of the deficit of regional budgets, which is due to changes in tax legislation, the policy of economic sanctions against Russia, the growth of social expenditures for the implementation of May' Presidential Decrees. Analysis of one of the main indicators of the country's debt security 'debt-to-GDP ratio' showed that Russia's debt policy is restrained compared to world borrowing practises. Traditionally, Japan, the United States and Eurozone countries are characterised by high public debt. Nevertheless, the relevance of the assessment of the debt sustainability of the budgets of the Russian regions is confirmed by the results of the indicative analysis, which showed a high heterogeneity of the values of six indicators for the regions of the Northwestern Federal District.

The application of the author's methodology using the standardisation of indicator values and the multidimensional average formula for the calculation of the integral index made it possible to rank regions by the level of debt sustainability, and the use of hierarchical cluster analysis allowed to group regions by three types of debt sustainability with the definition of threshold values (permissible boundaries) of debt sustainability indicators of each.

As a result of approbation of the methodology on the basis of open data of the Federal State Statistics Service and the Ministry of Finance of the Russian Federation for the regions of the Northwestern Federal District, the group of regions with low debt sustainability included the Republic of Karelia and the Republic of Komi, as well as the Pskov region, which allows us to conclude about the high risk of insolvency and the need to control the level of debt burden using the thresholds obtained in the study by region clustera. Novgorod and Arkhangelsk regions are also in a credit risk zone, as they fall into the group of regions with low debt sustainability in 2019 and 2017.

The proposed methodology for assessing the debt sustainability of the budget system of the region is universal and can be applied to data on other constituent entities of the Russian Federation for an objective assessment of the permissible level of debt burden and levelling the amount of public debt in case of debt stability reduction. Directions of further research within the framework of a given topic are determined by the need for annual statistical justification of differentiated values of the upper limits of public domestic debt, taking

into account the macroeconomic situation in the country and the priorities of socio-economic development of each subject of the Russian Federation.

ACKNOWLEDGMENTS

The article was written based on the results of budgetary-supported research according to the state assignment carried out by the Financial University as part of the research project "Ensuring the stability of the financial system of the region: Mechanism, tools, evaluation indicators", Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia.

REFERENCES

1. Sanginova L. D. Efficient debt policy of the Russian Federation subjects: Theory and practice. *Ekonomika. Nalogi. Pravo = Economics, Taxes & Law*. 2018;11(1):96–105. (In Russ.). DOI: 10.26794/1999–849X-2018–11–1–96–105
2. Rogatenyuk E. V. Indicative analysis of a debt security of the Russian Federation. *Uchenye zapiski Krymskogo federal'nogo universiteta imebi V. I. Vernadskogo. Ekonomika i upravlenie = Scientific Notes of V. I. Vernadsky Crimean Federal University. Economics and Management.* 2018;4(2):123–134. (In Russ.).
3. Galukhin A. V. Assessment of debt sustainability of budgets of constituent entities of the Russian Federation. *Voprosy territorial'nogo razvitiya = Territorial Development Issues*. 2016;(5):6. (In Russ.).
4. Zotova A. I., Kirichenko M. V. The stability of the financial system of the region: Essence, factors, indicators. *Teoriya i praktika obshchestvennogo razvitiya = Theory and Practice of Social Development*. 2017;(5):76–82. (In Russ.). DOI: 10.24158/tipor.2017.5.19
5. Checherita-Westphal C., Hughes Hallett A., Rother P. Fiscal sustainability using growth-maximizing debt targets. *Applied Economics*. 2014;46(6):638–647. DOI: 10.1080/00036846.2013.861590
6. Vavilov A., Kovalishin E. The principles of state debt policy. *Voprosy ekonomiki*. 2001;(8):46–64.
7. Kleiner G. B. State — region — field — enterprise: Framework of economics system stability of Russia. Part 2. *Ekonomika regiona = Economy of Region*. 2015;(3):9–17. (In Russ.). DOI: 10.17059/2015–3–1
8. Kluza K. Sustainability of local government sector debt. Evidence from Monte-Carlo simulations. *Lex Localis*. 2016;1(1):115–132. DOI: 10.4335/14.1.115–132(2016)
9. Bröthaler J., Getzner M., Haber G. Sustainability of local government debt: A case study of Austrian municipalities. *Empirica.* 2015;42(3):521–546. DOI: 10.1007/s10663–014–9261–3
10. Zhao R., Tian Y., Lei A., Boadu F., Ren Z. The effect of local government debt on regional economic growth in China: A nonlinear relationship approach. *Sustainability.* 2019;11(11):3065. DOI: 10.3390/su11113065
11. Checherita-Westphal C., Rother P. The impact of high government debt on economic growth and its channels: An empirical investigation for the euro Area. *European Economic Review*. 2012;56(7):1392–1405. DOI: 10.1016/j.euroecorev.2012.06.007
12. Joy J., Panda P. K. An empirical analysis of sustainability of public debt among BRICS nations. Journal of Public Affairs. 2021;21(2):112–126. DOI: 10.1002/pa.2170
13. Baharumshah A. Z., Soon S.-V., Lau E. Fiscal sustainability in an emerging market economy: When does public debt turn bad? Journal of Policy Modeling. 2017;39(1):99–113. DOI: 10.1016/j.jpolmod.2016.11.002
14. Égert B. Public debt, economic growth and nonlinear effects: Myth or reality? Journal of Macroeconomics. 2015;43:226–238. DOI: 10.1016/j.jmacro.2014.11.006
15. Eberhardt M., Presbitero A. F. Public debt and growth: Heterogeneity and non-linearity. Journal of International Economics. 2015;97(1):45–58. DOI: 10.1016/j.jinteco.2015.04.005
16. Reinhart C. M., Rogoff K. S. Growth in a time of debt. *American Economic Review*. 2010;100(2):573–578. DOI: 10.1257/aer.100.2.573
17. Chudik A., Mohaddes K., Pesaran M. H., Raissi M. Is there a debt-threshold effect on output growth? *The Review of Economics and Statistics.* 2017;99(1):135–150. DOI: 10.1162/REST_a_00593

18. Kalina A.V., Savelieva P. Formation of threshold values of the economic security of Russia and its regions. *Vestnik Yuzhno-Ural'skogo gosudarstvennogo universiteta. Seriya: Ekonomika i menedzhment = Bulletin of South Ural State University. Series "Economics and Management"*. 2014;8(4):15–24. (In Russ.).
19. Kurilova A.A., Kiryushkina A.N. Assessment of debt sustainability of subjects of the Russian Federation as managing sovereign debt at the regional level. *Karel'skii nauchnyi zhurnal = Karelian Scientific Journal*. 2016;5(4):156–160. (In Russ.).
20. Soldatkin S.N. Debt sustainability of Russian regions: The overall assessment and the adequacy of regulation. *Politematicheskii setevoi elektronnyi nauchnyi zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta = Scientific Journal of KubSAU*. 2017;(128):1325–1342. (In Russ.). DOI: 10.21515/1990–4665–128–094
21. Soldatkin S.N. Evaluation of the effectiveness of regional debt policy (on the example of the subjects of the Russian Federation included in the Far Eastern Federal District). *Vestnik Khabarovskogo gosudarstvennogo universiteta ekonomiki i prava = Bulletin of Khabarovsk State University of Economics and Law*. 2019;(1):59–66. (In Russ.).
22. Buklemishev O.V., Mikhaylov D.L, Solovyev V.V., Fetisov I.G. Effectiveness of legislative restrictions on debt burden of the Russian regions. *Finansovyi zhurnal* = Financial Journal. 2018;(3):34–48. (In Russ.). DOI: 10.31107/2075–1990–2018–3–34–48
23. Uskova T.V. Managing the sustainable development of the region. Vologda: Institute for Socio-Economic Development of Territories RAS; 2009. 355 p. (In Russ.).
24. Kurushin D.S., Vasilyeva E.E. The construction of a typology of branch banks on the basis of cluster analysis and neural networks. *Finansy: teoriya i praktika = Finance: Theory and Practice*. 2017;21(6):6–19. (In Russ.). DOI: 10.26794/2587–5671–2017–21–6–6–19

ABOUT THE AUTHOR

Elena G. Kiseleva — Cand. Sci. (Econ.), Assoc. Prof., Graduate School of Business and Management of the Institute of Industrial Management, Economics and Trade, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia
https://orcid.org/0000-0002-6334-5562
chachina_eg@spbstu.ru

Conflicts of Interest Statement: The author has no conflicts of interest to declare.

The article was submitted on 12.08.2021; revised on 30.08.2021 and accepted for publication on 01.12.2021.
The author read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-129-145
UDC 332.14,330.356(045)
JEL C01, C23

Econometric Analysis of the Effectiveness of Government Incentive Measures for the Development of the Region

I.V. Tregub[a], M. De Iaco[b]
[a] Financial University, Moscow, Russia; [b] University of Studies Turin, Turin, Italy

ABSTRACT

This study aims to identify the main factors that can influence the growth of the regional economy in order to assess the effectiveness of the government support measures for regions, designed to stimulate the regional domestic product growth. The subject of the study is the relationship of socio-economic indicators in the Northwestern Federal District of Russia. The authors apply the method of correlation-regression analysis. The scientific novelty lies in the development of an econometric model based on statistical data of the constituent entities of the Northwestern Federal District to forecast the levels of regional domestic product and consumer spending per capita. The main socio-economic indicators of the development of the Northwestern Federal District have been identified, which make it possible to quantitatively assess the measures taken by the government to financially support the population and the economy of the region. The Menges model was chosen as the basis for the study, which allows analyzing the relationship between such significant financial and economic indicators of an economic entity as gross regional product, investments, profit of organizations and consumption. The article developed an econometric model in the form of a system of interconnected econometric equations, the coefficients of which were estimated in the Gretl package of applied econometric programs. Based on the analysis and modification of the classical Menges model, the authors were able to determine the key indicators that have a significant impact on the dynamics of the economy of the Northwestern Federal District. These include the following factors: investments; consumer spending per capita; taxes; social payments; household deposits; personal loans; industrial production; income from business activities; profit of organizations. The authors conclude that financial support for entrepreneurs and businesses does not have a significant impact on the economic growth of the region, and measures aimed at increasing the wages of employees of organizations for the economy of the Northwestern Federal District are more effective than measures associated with an increase in social benefits. The research results will be useful to those who make regional management decisions to stabilize the post-pandemic economic situation in the Northwestern Federal District.

Keywords: financial and economic indicators; Northwestern Federal District; Menges model; econometric analysis; forecasting; the effectiveness of government support measures; investments; profit of organizations; consumption; entrepreneurial activity

For citation: **Tregub I.V., Iaco M. De. Econometric analysis of the effectiveness of government incentive measures for the development of the region.** ***Finance: Theory and Practice.*** **2022;26(3):129-145. DOI: 10.26794/2587-5671-2022-26-3-129-145**

INTRODUCTION

The 2020 pandemic had a negative impact on the global economy. Suspension of production, reduction of services in the consumer's sector, falling incomes — these are the economic consequences of the quarantine, which led to a slowdown, up to a complete stop, of economic development of both the country as a whole and its regions.

In the current conditions of the global crisis and the constantly changing external environment, the Government of the Russian Federation pays special attention to supporting the population and business. The socio-economic measures being taken are aimed at stabilizing the economic situation in the country, supporting the regional economy and the social well-being of citizens. This determines the importance of studying the drivers of economic growth in the regions, the impact on which can increase the effectiveness of government support measures.

The Northwestern Federal District, formed by the Decree of the President of the Russian Federation dated May 13, 2000, was chosen as the object of study. It is located in the north and northwest of the European part of the Russian Federation. The district includes 11 subjects with a population of 13,941,959 people (9.54% of the population of Russia as of January 1, 2021) and an area of 1,686,972 km^2 (9.85% of the territory of the Russian Federation).[1] The center of the region is the federal city of St. Petersburg. The Northwestern Federal District includes: the Republic of Karelia, the Komi Republic, the Arkhangelsk, Vologda, Kaliningrad, Leningrad, Murmansk, Novgorod, Pskov regions, the city of St. Petersburg, the Nenets Autonomous Okrug.

The industrial production index in the Northwestern Federal District in January-November 2020, according to Rosstat, amounted to 97.2% compared to the corresponding period last year. This coincides with the all-Russian data (97%). According to this indicator, the Northwestern Federal District ranks fifth among eight federal districts.

From an economic perspective, 2020 has become the most challenging year for the Northwestern Federal District in the last decade. Although there were no obvious economic collapses, the recession in a number of industries was significant.

The oil-producing Nenets Autonomous Okrug (NAO) and the Komi Republic (–7%), as well as the Kaliningrad Region (–8%), suffered the largest losses (–11% in the first eleven months of 2020). The structure of industrial production in the NAO and Komi is such that the extractive industry makes a significant contribution to their regional domestic product. Russia's accession to the OPEC+ deal has led to a decline in oil production in almost the entire country. In the NAO, the decline in production amounted to 11.2% compared to January-November 2019, and in the Komi Republic — 8.7%. As for the economy of the Kaliningrad region, its decline was largely due to the automotive cluster, which reduced production by 27.2%.

The industry showed growth only in two regions of the North-West — the Vologdskaya Oblast (+1.8%) and Karelia (+22.8%). In the first case, this is primarily the result of the stable operation of the Cherepovets Metallurgical Plant, the main asset of Severstal.

According to the results of January-November 2020, the Republic of Karelia ranked second in terms of the dynamics of the industrial production index after North Ossetia. The extraction of crushed rock increased 5.3 times, the output of the food industry increased by 16.6%, and the volume of beverage production increased by 72.1%. But the industries that seem traditional for Karelia — woodworking and paper production — did not give a significant increase. This determines the relevance of studying the reasons for this situation, both as a result of the efforts of regional authorities to diversify the economy, and the effect of a

[1] Northwestern Federal District. Official site of the Plenipotentiary Representative of the President of the Russian Federation in the Northwestern Federal District. URL: http://szfo.gov.ru/ (accessed on 23.03.2021).

low base since in previous years the industry of the republic was in a rather depressed state. The rational use of economic resources can lead to significant economic development in the region and an increase in the standard of living of the population. Therefore, it is very important to understand the relationship between the main financial and economic indicators of the region and to identify the key factors influencing the economic development of the region.

PREVIOUS RESEARCH

The study of the development of the economy, both individual regions and national economies, is today a topical problem, the solution of which is devoted to a lot of scientific articles by modern domestic and foreign authors.

In 2001, Academician A. G. Granberg outlined the strategy of the territorial socio-economic development of Russia [1], the main provisions of which were developed in the scientific works of many domestic researchers and were included in the main documents that form the state policy.

The works of P. A. Minakir and A. N. Demyanenko discussed the evolution of approaches and methodology of spatial economics [2].

G. B. Kleiner and M. A. Rybachuk [3] presented the results of the analysis of the systemic balance of the constituent entities of Russia, based on the calculation of the systemic balance index. The authors showed a significant scatter of regions in terms of the degree of balance and a high proportion of structurally unbalanced regions. In the region under study, certain regions of the Northwestern Federal District, for example, the Nenets Autonomous Okrug, fell into the category of outsiders. A.N. Bereznyatskiy and B.E. Brodskiy [4] described St. Petersburg as a leading region.

S.A. Aivazian, M. Yu. Afanasiev, A.V. Kudrov [5] developed indicators that characterize the socio-economic state of the regions of the Russian economy. The differentiation is based on eight indicators that characterize the five main areas of socio-economic development of the constituent entities of Russia, including production of goods and services, material well-being, quality of the population, quality of social services, and internal security. St. Petersburg is ranked among the leaders in terms of the main parameters for ensuring stable development.

We will use the understanding of the systemic balance and stability of various areas of the Northwestern Federal District, obtained by previous researchers, to analyze the validity of government support measures for individual regions of the district.

The use of econometric methods for studying various economic problems makes it possible to obtain a quantitative assessment of the influence of various endogenous and exogenous factors on the socio-economic characteristics of regions. Problems of this type for the regions of Russia were solved using econometric modeling in the publications of many authors.

In the article by A.N. Bereznyatskiy, and B.E. Brodskiy [4], the authors were interested in the problem of modeling the macro dynamics of the Russian economic region. At the same time, the authors took into account only intra-regional factors of economic development. They built models where the gross regional product was used as a simulated indicator; world oil prices, world prices for products of the domestically oriented sector, tariffs for products of the natural monopoly sector, real investment in fixed assets, the real exchange rate of the ruble, subsidies were exogenous variables. However, for the purposes of our study, we need a set of factors that reflect the internal parameters of regional development.

The article by N.E. Egorov [6] analyzed and assessed the creative potential of the Far Eastern Federal District in the field of scientific, educational, and innovative activities for the period 2010–2014. The results of the econometric analysis are presented in the form of tables, allowing the formulation of realistic management recommendations.

In the works [7, 8], econometric methods are used to develop managerial decisions based on the analysis of the determinants of the socio-economic development of various economic entities. Panel data models for studying the economic development of regions are also found in modern literature.

The study of the influence of initial conditions on the properties of the parameters of dynamic panel data models was considered in the work of R. J. Barro et al. (2017), where, along with studies of the determinants of economic growth, the authors examined this growth using econometric models of the economies of individual countries [9].

S.A. Aivazian et al. [10] developed dynamic multiple linear regression models in which the variables were taken in logarithmic form. The endogenous variable was the natural logarithm of the regional domestic product, and the natural logarithms of the value of fixed assets, the economically active population, the number of organizations performing scientific research, the number of the time period, the product of the logarithms of the above indicators by the number of the period time. Panel data for 80 Russian regions for the period 2009–2013 became the empirical basis for modeling. As significant conclusions of their study, the authors announced a significant increase in the elasticity of the regional domestic product by the number of organizations performing scientific research. The use of the logarithm of empirical data makes it possible to deal with the effects of the so-called spurious regressions. However, this smooths out the effect of the impact of crises, which can be included in the models in the form of structural shift variables. In this paper, we will use "pure" data and substantiate the significance of the obtained simulation results using special tests. At the same time, the results of S.A. Aivazian et al. [10], with the appropriate transformation of the initial data, can be used to verify our own research results.

Among all the works studied, we are close to the approach presented by G.P. Litvintseva, I.N. Karelin [11], H.T. Hoang, L.T.D. Huynh, and G.S. Chen [12].

G. P. Litvintseva and I. N. Karelin [11] attempted to determine the impact of digital factors on the main indicator of the social well-being of the population (gross regional product per capita) for the period from 2015 to 2018 in the constituent entities of Russia. To assess the effects, panel data models were used, showing the relationship of parameters simultaneously in space and time for the period 2015–2018.

H.T. Hoang, L.T.D. Huynh, and G.S. Chen [12] estimated the structural model using data on the provinces of Vietnam for the period 2000–2012.

P.S. Bravok and L.E. Pynko [13] estimated the expected values of the region's development in terms of the regional domestic product based on an econometric analysis of the report's panel data, which proposes an analysis of the main economic indicators of the Far Eastern Federal District. The purpose of this analysis was to forecast the gross regional product of the Far Eastern Federal District for 2019–2021, using a time series model and panel data, and comparing the accuracy of time series forecasts and panel data models. The authors stated that a fixed effects model can be used for forecasting. Due to the peculiarities of data accounting, it is possible to make forecasts simultaneously for all regions of the Far Eastern Federal District. According to their hypothesis, for a more accurate and efficient forecast of the regional domestic product, the simultaneous use of many types of models is necessary. It is possible to determine the model that offers the most reliable prediction by comparing their characteristics.

In this paper, we will apply a similar approach to the Northwestern Federal District.

As for research on the development of the Northwest region itself, there are not so many studies. S.V. Kuznetsov proposed a strategy for the scientific and technological development of the Northwest as an instrument of regional policy [14]. B.M. Grinchel' and E.A. Nazarova, assessing the sustainability of the innovative development of the regions of the Northwestern Federal District according to the

data of 2013–2016, concluded that a decrease in the costs of scientific research and the introduction of technological innovations into production indicates a decrease in the pace of innovation in the region [15]. I.V. Tregub [16] studied the development of regions from the standpoint of demography and identified factors that affect the effectiveness of government support measures.

Among the works of interest, in our opinion, are the works of P.V. Stroev et al., V.K. Kashin and S.V. Makar [17, 18], who studied the socio-economic development of the regions of the Russian Federation under the conditions of increased volatility of the ongoing processes in the Russian and global economies. At the same time, V.K. Kashin and S.V. Makar substantiated [18] that the Northwestern Federal District can be considered as a model of Russia, where various types of regional dynamics are presented. This conclusion is also very important for our study since it allows us to extrapolate the obtained data to the Russian economy as a whole.

Thus, summing up the review of scientific articles devoted to the analysis of regional development, it can be noted that with all the variety of ongoing research on econometric modeling of the dynamics of economic indicators of regional development and the identification of growth factors for regional economies, the topic of this article continues to be relevant due to the constantly changing economic situation and the presence of crisis phenomena in the global economy, which were not taken into account by previous researchers.

MATERIALS AND METHODS

To study the effectiveness of government support measures for the regions, econometric models were developed to identify the relationship between the characteristics of the economic development of the district and the factors that have a significant impact on these indicators.

Initially, it was assumed that the analysis of the economic development of the region can be carried out using the Menges macro econometric model [19, 20], which can be successfully used to model both national economies and individual regions.

Hypothesis 1. Measures of government support measures for the economic development of the Northwestern Federal District can be assessed by the results of identifying the relationship between the socio-economic indicators of the development of the district using regression analysis based on the Menges model, the essence of which is that:

- the current level of regional domestic product can be explained by the value of GRP at the previous point in time and investments in fixed assets;
- investments in fixed capital are explained by the value of the regional domestic product and the gross profit in economics;
- current consumption depends on the previous level of consumption, the cost of the standard of living, and the value of the regional domestic product;
- the gross profit in economics depends on its previous value and the volume of industrial production.

It should be noted that in this model, not all exogenous factors are associated with the government and their influence cannot be interpreted from the standpoint of evaluating the effectiveness of government support measures to stimulate regional development. However, the presence of these variables in the equations is necessary to obtain unbiased effective, and consistent estimates of the regression coefficients.

Sample Description

To conduct the study, data were collected on the main indicators of development and standard of living of the population of ten subjects of the Northwestern Federal District for 2015–2020. Thus, the total sample size is 60 people. Panel data for estimating the parameters of econometric models are taken from the Russian Statistical Yearbooks of Rosstat.[2]

[2] Regions of Russia. Socio-economic indicators. URL: http://www/https://rosstat.gov.ru/folder/210/document/13204 (accessed on 23.08.2021).

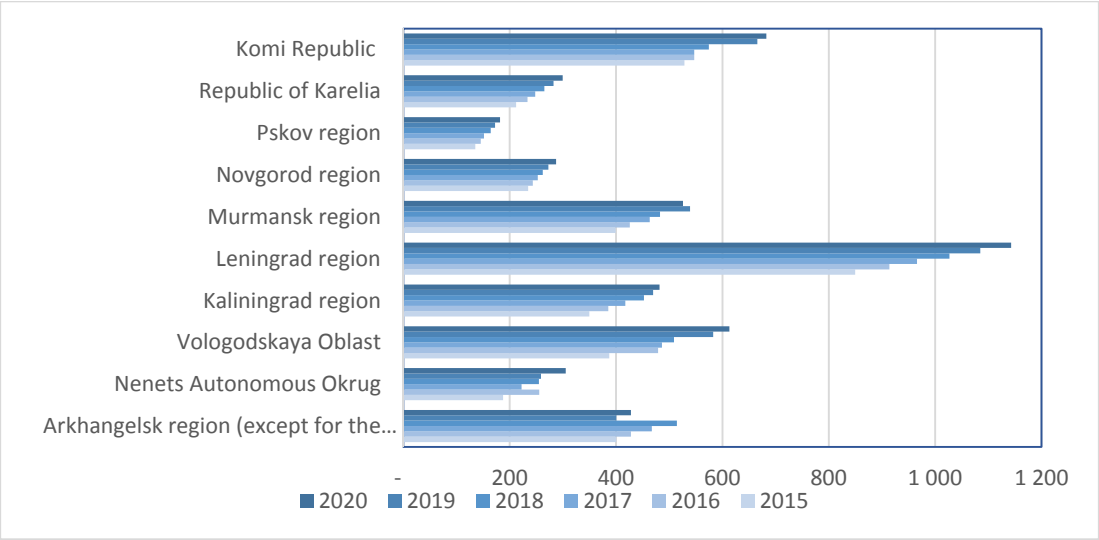


Fig. 1. **Regional domestic product, 2015–2020, billion rubles**

Source: compiled by the authors. Regions of Russia. Socio-economic indicators. URL: http://www/https://rosstat.gov.ru/folder/210/document/13204 (accessed on 23.08.2021).

The dynamics of the regional domestic product for various subjects of the district are shown in *Fig. 1*.

The mathematical notation of the classical Menges model is a system of four simultaneous equations:

$$Y_t = a_0 + a_1 \cdot Y_{t-1} + a_2 \cdot I_t + \mu_t, \quad (1)$$

$$I_t = b_0 + b_1 \cdot Y_t + b_2 \cdot Q_t + \vartheta_t, \quad (2)$$

$$C_t = c_0 + c_1 \cdot Y_t + c_2 \cdot C_{t-1} + c_3 \cdot P_t + \varphi_t, \quad (3)$$

$$Q_t = d_0 + d_1 \cdot Q_{t-1} + d_2 \cdot R_t + \varepsilon_t. \quad (4)$$

The following designations are used in the classical model: Y — the national income; I — investments; C — personal consumption expenditures; Q — the gross profit in economics; P — the cost of living index; R — the volume of industrial production.

The first equation (1) of the Menges model shows the impact of investment and the lagged value of national income on the current value of national income. The second equation (2) of the model reflects the dependence of investments on national income and the gross profit in economics. The third equation (3) reflects the impact of national income, cost of living index, and previous consumption on the current value of the consumption function. The last equation (4) allows modeling the current value of the gross profit of the economy depending on its lagged value and the volume of industrial production. It should be noted that the system of equations (1)–(3) are simultaneous equations, the coefficients of which can be estimated, for example, by the two-stage least squares method (2SLS). Equation (4) contains variables independent of equations (1)–(3). Estimates of the coefficients of equation (4) can be found using the method of ordinary least squares (OLS). An alternative to the use of 2SLS can be a technique for transforming a system of simultaneous equations to a reduced form when each equation contains only one endogenous variable.

For greater convenience in interpreting the simulation results, we used our own notation for variables. Some variables have been replaced by their counterparts. With regard to the panel data of the Northwestern Federal District, the list of variables used to model the economy of the region, their description, and their dimension are given in *Table 1*.

The index t denotes the current moment of time, and the index i — is the number of the subject of the district.

Table 1

Indicators used in the econometric equations of the Menges model

Variable abbreviation in the classical model	Variable abbreviation in the developed model	Description of the variable in the developed model	Dimension of a variable
Endogenous variables			
Y_t	GRP_{it}	Regional domestic product	Billion rubles
I_t	I_{it}	Investments in fixed assets	Billion rubles
C_t	C_{it}	Consumption expenditure per capita	Thousand rubles
Q_t	$Profit_{it}$	Profit of organizations	Billion rubles
Predefined variables			
Y_{t-1}	GRP_{it-1}	Regional domestic product, lagged value	Billion rubles
C_{t-1}	C_{it-1}	Consumption expenditure per capita, lagged value	Thousand rubles
Q_{t-1}	$Profit_{it-1}$	Profit of organizations, lagged value	Billion rubles
R_t	Vol_{it}	Volume of industrial production	Billion rubles
P_t	CPI_{it}	Consumer price index	Percentage points
Random perturbation			
$\mu_t, \vartheta_t, \varphi_t, \varepsilon_t$		Reflects the influence of all factors not taken into account in the model	Has the dimension of the dependent variable

Source: compiled by the authors.

It should be noted that the "Cost of Living Index" indicator in the classical model was replaced by the "Consumer Price Index". This is possible since the cost of living index in the methodology of Rosstat is understood as an indicator that measures the relative cost of a set of goods and services in individual cities compared to its average Russian value. The result of comparing these indicators allows us to calculate how much more expensive (or cheaper) the same specific set of consumer goods and services will be with the same volumes of their consumption in different cities of Russia, which, in fact, coincides with the consumer price index indicator. Statistics on the cost of the living index are conducted only by cities and are not provided for individual subjects of federal districts. That's why we used CPI instead of the Cost of Living Index.

Model parameters (1) were estimated taking into account their panel structure. General methods for estimating the parameters of a linear regression equation, as well as the features of estimating regression parameters, taking into account their spatial structure, are given in the work of I.V. Tregub [21]. In this paper, we used the ordinary least squares algorithm modified to take into account the structure of panel data and the presence of an autoregressive component in the models [22].

RESULTS OF ECONOMETRIC ANALYSIS AND DISCUSSION

The analysis of state support measures is based on the results of modeling the development of the economy of the Northwestern Federal District using the Menges model.

Since the Menges model is a system of interrelated equations, at the first stage of modeling, the system of equations (1–3) was estimated by a two-stage least squares method with the choice of instrumental variables from a list of predefined variables (*Table 1*). This approach is justified since when building

Table 2

Results of testing the classical Menges model on data of the Northwestern Federal District

	GRP_{it}	I_{it}	C_{it}	$Profit_{it}$
Const	–4.087	–103.025***	–555.756	41.5
GRP_{it-1}	1.080***	–	–	–
I_{it}	–0.034	–	–	–
GRP_{it}	–	0.485***	–0.014	–
C_{it-1}	–	–	1.862	–
Vol_{it}	–	–	–	–1.306*
CPI_{it}	–	–	529.175	–
$Profit_{it}$	–	0.25*	0.28	–
$Profit_{it-1}$	–	–	–	0.968***
R^2-*adjusted*	0.98	0.82	**0.13**	0.85

Source: compiled by the authors.

Note: * – the coefficient is significant at α = 0.1; ** – the coefficient is significant at α = 0.05; *** – the coefficient is significant at α = 0.01.

panel data models, they usually start with the pooled regression model, which does not take into account the structure of panel data. After that, the fixed and random effects models are evaluated and the presence of individual effects is checked.

It should be noted that tests for the presence of intragroup and intergroup individual effects in the Menges model showed their absence in the sample of panel data for the Northwestern Federal District.

According to *Table 2*, the results of the equation for modeling the regional domestic product, investments in fixed assets, and profits of organizations have high explanatory power. Thus, the dynamics of the regional domestic product are 98% explained by the change in investments in fixed assets and the previous value of GRP. However, the only significant factor in this equation is the lagged value of the regional domestic product.

The amount of investment in fixed capital, in turn, depends on the values of the regional domestic product and the profit of organizations by 82%. The profit of organizations in the Northwestern Federal District is explained by the past value of profit and the volume of industrial output by 85%.

At the same time, the consumption level equation in the Menges model is absolutely inapplicable for analysis, since it has low explanatory power (only 13%). Also, the variables in the equation are not significant.

The absence of asterisks next to the coefficient in *Table 2* means that this coefficient is not significant for an acceptable level of significance. Sign "–" in *Table 2* means that this coefficient is absent in equations (1–4) of the Menges model.

Due to the low explanatory power of individual equations of the Menges model and a large number of insignificant variables, it was decided to modify the model by adding additional variables to the equations for the consumption function and the regional domestic product function. Equation (2) for the investment function of the classical Menges model and equation (4) for the gross profit of the county economy are left unchanged. The impact of the pandemic was also further studied to justify the need to include structural shift dummy variables in the model.

Table 1 shows that, despite the pandemic and the economic downturn of the Russian economy in 2020, there are no obvious

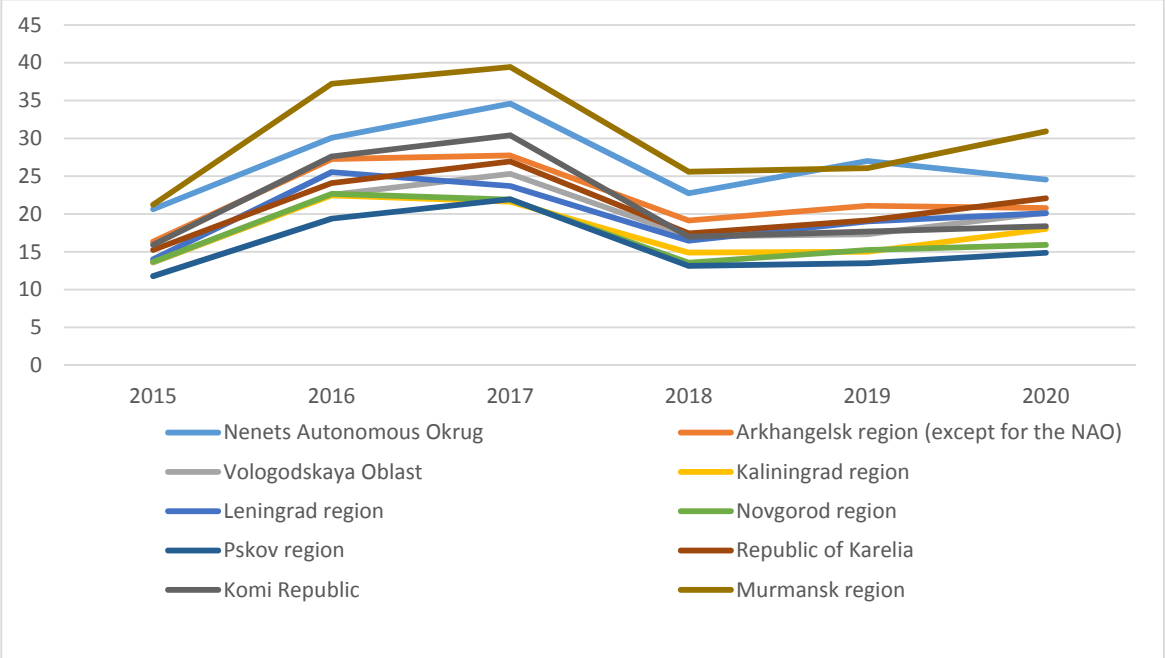


Fig. 2. **Analysis of the impact of the pandemic on the consumption functions of residents of particular regions of the Northwestern Federal District. Consumption per capita, thousand rubles**

Source: compiled by the authors. Data Regions of Russia. Socio-economic indicators. URL: http://www/https://rosstat.gov.ru/folder/210/document/13204 (accessed on 23.08.2021).

structural shifts in the data in individual subjects of the Northwestern Federal District. This allows us to conclude that there is no need to include dummy shift variables when modeling the district's regional domestic product.

To study the impact of the pandemic-related crisis on the level of consumption, a graph of the consumption function was plotted, which is shown in *Fig. 2*. We see that in 2020 there was no sharp drop in consumption compared to 2019. On the contrary, in some areas, such as the Murmansk region, there was an increase in consumption. This means that the pandemic has not made significant adjustments to the dynamics of the process.

A new set of variables included in the equations of the Menges model is of research interest since it allows us to quantify the feedback on the measures taken to support the population and business.

For the consumption function and the function of the regional domestic product, the included additional factors, on the one hand, allow us to assess the degree of influence on consumption levels and regional domestic product, and on the other hand, they can be associated with measures of government support measures for the population. As a result, the following hypotheses were put forward:

Hypothesis 2. The value of the regional domestic product can be explained by the amount of investment, the level of consumption, the amount of taxes collected, the income of the population and business, and the volume of industrial production.

Hypothesis 3. The amount of investment in the economy of the Northwestern Federal District can be explained by factors within the framework of the classical Menges model, namely, the current values of the regional domestic product and the profits of organizations.

Hypothesis 4. In addition to the factors of the classical Menges model, the variables that influence per capita consumption levels are wages, social benefits, and interest rates on deposits and loans.

Hypothesis 5. The current value of the profit of organizations in the Northwestern Federal

District can be explained by the profit in the previous period and the volume of production, i.e. the assertion of the classical Menges model is valid.

Indeed, rising household incomes should increase the cost of acquiring additional goods that cannot be provided at low-income levels. The size of the interest rate on deposits should influence the decision on the advisability of saving on deposits for the purpose of deferred consumption. Interest rates on credit products, on the contrary, should influence the decision-making on spending in the current period. Thus, it can be assumed that an increase in interest rates on deposits should lead to a decrease in consumption in the current period since individuals will prefer to save. A decrease in interest rates on loans, on the contrary, should lead to an increase in consumption, since low-interest rates on loans can contribute to an increase in the processing of credit products, the funds from which are directed to consumption.

Hypothesis 3 refers to household income only as a part of income, the impact of which on consumption level is of research interest.

The specification of the model, taking into account the new notation and Hypotheses 2 and 3, has the form:

$$GRP_t = a_0 + a_2 * I_t + a_3 * C_t + a_4 * Tax_t + a_5 * SP_t + \\ + a_6 * Dep_t + a_7 * Cr_t + a_8 * Vol_t + \\ + a_9 * Bis_t + a_{10} * \Pr ofit_t + \mu_{1t}, \quad (5)$$

$$I_t = b_0 + b_1 * GRP_t + b_2 * \Pr ofit_t + \vartheta_{1t}, \quad (6)$$

$$C_t = c_0 + c_1 * S_t + c_2 * SP_t + c_3 * Dep\_Rate_t + \\ + c_4 * Cr\_Rate_t + c_5 * GRP_t + c_6 * CPI_t + c_7 * C_{t-1} + \varphi_{1t}, \quad (7)$$

$$\Pr ofit_t = d_0 + d_1 * \Pr ofit_{t-1} + d_2 * Vol_t + \varepsilon_t. \quad (8)$$

The new variables of the modified Menges model are presented in *Table 3*.

Tables 4 and *5* present the results of estimating the coefficients of equations (5–8) for the pooled regression model and the fixed effects (FE) model.

For the regional domestic product dynamics represented by equation (5), the significant variables in the pooled regression model were investments, consumption expenditures per capita, industrial production, household income, taxes, and business profits. Within the framework of this model, changes in the regional domestic product are explained by changes in the specified set of variables by 97%. The quality of the model specification based on the results of the F-test is high, and the prerequisites of the Gauss-Markov theorem are fulfilled.

The simulation of the regional domestic product represented by equation (5) using the fixed effects (FE) model includes fewer significant factors (variables in *Table 4* with asterisks after the coefficient). The explanatory power of the model is also lower, only 79%.

The Hausman test to choose between the Pooled and FE models concludes that at a significance level of 0.01 for the regional domestic product represented by equation (5), a pooled regression model can be applied.

It can be argued that Hypothesis 2 is confirmed. A pooled regression model can be used to predict the economic development of the Northwestern Federal District and analyze the effectiveness of managerial decisions.

The evaluation of the parameters of the investment function represented by equation (6) showed that the variable GRP_t turned out to be insignificant at any reasonable level of significance. As a result, it was decided to replace the current value of the regional domestic product with its lagged value. The output statistics of the simulation results are presented in *Table 5*. The explanatory power of the pooled regression model is 82%, the regression is generally significant, and the residuals of the model are heteroscedastic and non-autocorrelated. The model is adequate and suitable for forecasting. Hypothesis 3 was partially confirmed.

For the model of consumption expenditure per capita, represented by equation (7), the results are presented in *Table 4*. Then pooled regression model and a fixed effects model for

Table 3

Predefined variables of the modified model

Regional Domestic Product Function. Equation 5		
I	Investments	Billion rubles
C	Consumption expenditure per capita	Thousand rubles
Tax	Taxes	Billion rubles
Sp	Social payments	Million rubles
Dep	Household deposits	Million rubles
Cr	Personal loans	Million rubles
Vol	Volume of industrial production	Million rubles
Bis	Business income	Million rubles
Profit	Profits of organizations	Billion rubles
Consumption Function. Equation 7		
S	Salary per worker	Thousand rubles
Sp	Social payments	Million rubles
Dep_Rate	Interest rate on household deposits	%
Cr_Rate	Interest rate on consumer loans	%
GRP	Regional domestic product	Billion rubles
CPI	Consumer price index	%
C_{t-1}	Consumption per capita in the previous time period	Thousand rubles

Source: compiled by the authors.

the consumption function gave similar results. Both models explain consumption levels by a set of variables included in the equation by 66%, which is significantly higher than in the classical Menges model (13%). The quality of model specifications based on the F-test results is high. The residuals of the models are homoscedastic and non-autocorrelated. There is no multicollinearity.

Thus, hypothesis 4 is confirmed, which allows us to analyze the simulation results and draw new important conclusions, which are given in the section "Interpretation of the results of the empirical study".

Testing Hypothesis 5 on the significance of the coefficients of the profit function equation of organizations in the Northwestern Federal District showed that for the pooled regression model and the model with fixed effects, the output variable is not significant. In addition, there is the autocorrelation of residuals in the model. To be able to use this model to analyze the economic situation in the area, it is necessary to choose exogenous variables that will improve the output of statistics. However, the explanatory power of the pooled model is quite high (85%) and the regression is generally significant. Thus, we can conclude that Hypothesis 5 cannot be confirmed.

INTERPRETATION OF THE RESULTS OF THE EMPIRICAL STUDY

Regional domestic product function

The modified model of the regional domestic product is represented by equation (5). The hypothesis put forward was confirmed, which allows us to draw important and new conclusions regarding the identification

Table 4

Results of regional domestic product and consumption modeling

Model Variables	GRP_{it}		C_{it}	
	Pooled	*FE*	*Pooled*	*FE*
GRP_{it}	–	–	–0.0075	–0.02
I_{it}	0.39*	0.089	–	–
C_{it}	1.79**	0.27	–	–
C_{it-1}	–	–	0.16**	0.004
Vol_{it}	0.00017*	4.53e-05	–	–
CPI_{it}	–	–	–29.7*	–103.23**
S_{it}	–	–	0.27***	0.17
Sp_{it}	0.0027***	0.0040***	8.3e-05*	33e-05**
Dep_{it}	0.13***	0.11***	–	–
Cr_{it}	–1.02***	–2.73***	–	–
Dep_Rate_{it}	–	–	–191.19**	–147.5
Cr_Rate_{it}	–	–	251.57***	193.9**
Tax_{it}	0.00095***	0.00043*	–	–
Bis_{it}	–0.00058	0.0038	–	–
$Profit_{it}$	–0.3***	0.001	–	–
Const	8.7	12.9	6.76	96.5
R^2-*adjusted*	0.97	0.79	0.660	0.659

Source: compiled by the authors.

Note: * the coefficient is significant at α = 0.1; ** the coefficient is significant at α = 0.05; *** the coefficient is significant at α = 0.01.

Table 5

Results of the investment function and profit of the district economy modeling

	I_{it}		$Profit_{it}$	
	Pooled	*FE*	*Pooled*	*FE*
GRP_{it-1}	0.47***	0.58***	–	–
Vol_{it}	–	–	–1.3	44.21
$Profit_{it}$	0.25*	2.66	–	–
$Profit_{it-1}$	–	–	0.96***	0.29*
Const	–85.7***	–132.1**	4.15	23.1
R^2-*adjusted*	0.82	0.44	0.85	0.18

Source: compiled by the authors.

Note: * the coefficient is significant at α = 0.1; ** the coefficient is significant at α = 0.05; *** the coefficient is significant at α = 0.01.

and justification of key factors that have a significant impact on the economic development of the region.

First, according to the results of *Table 4* for the model of the pooled regional domestic product function, the increase in consumption has a positive impact on the economic development of the region. Thus, an increase in the expenses of each resident of the region by 1 thousand rubles per year leads to an annual increase in the regional domestic product of the Northwestern Federal District by 1.79 billion rubles.

Second, the tax variable also accounts for consumption indirectly, since it includes, among other things, the value-added tax that each buyer pays when purchasing goods or services. The growth of tax revenues to the regional budget, according to the results of *Table 4* also boosts the county's economy.

Of particular interest is the analysis of the impact of household income on the economic growth of the region. The study of the impact of wages on the regional domestic product showed that, despite the inclusion of the variable *S* in various versions of the equation for the regional domestic product function, in all cases, it turned out to be insignificant and led to a deterioration in output simulation results statistics.

Third, significant variables indirectly related to household income were social payments and the amount of household deposits and loans. As follows from the results of *Table 4*, the funds of the population, concentrated in bank accounts, do not work for the development of the region's economy. Increase in the volume of deposits of the population by 1 million rubles leads to a decrease in the regional domestic product by 191 billion rubles per year. On the other hand, the volume of borrowed funds of the population leads to economic growth. But in this case, attention should be paid to the possible risks for the economy and social well-being of the population associated with the degree of debt burden.

Income from doing business in both variants of the intra-regional product model (Pooled and FE) turned out to be insignificant. The non-exclusion of the variable from the model, on the one hand, is associated with research interest in the degree of influence of government support measures for business on the growth of the region's economy, on the other hand, with ensuring the properties of the residual model that satisfy the prerequisites of the Gauss-Markov theorem.

The volume of industrial production has a positive impact on the economy of the region. An increase in output by 1 million rubles leads to an increase in the regional domestic product by 0.17 million rubles. One billion investments lead to an increase in the regional domestic product by 0.39 billion rubles.

Investment function

The results of modeling the investment function represented by equation (6) are given in *Table 5*. For the consumption function and the regional domestic product, pooled regression model approaches and fixed effects models were applied to the investment function. The pooled model explains the dynamics of the investment function by changes in the exogenous variables of the model by 82%, while the FE model explains only by 44%.

A new result of the study of the economic development of the Northwestern Federal District is the conclusion that, in contrast to the classical Menges model, in which the current value of the regional domestic product is included in equation (2) for the investment function, in the modified model this variable turned out to be insignificant, therefore, in equation (6), we included the lagged value of the regional domestic product. At the same time, the volume of investments depends on the economic development of the region. The regional domestic product increased by 1 billion rubles at present can attract investments in the amount of 0.44 billion rubles in the next period of time. Increasing the profit of organizations by 1 billion rubles leads to the growth of the region's economy by 0.25 billion rubles per year.

Consumption Function

Estimation of the coefficients of equation (7) for the modified consumption function is given in *Table 4*. It can be noted that consumption increases with the growth of wages and the increase in the amount of social benefits received. An increase in the wages of an individual employee by 1 thousand rubles leads to an increase in consumption expenditure per capita by 0.27 thousand rubles, which, taking into account the population of the district, gives an increase in consumption by 3.7 billion rubles per year. At the same time, an increase in social payments provided to individual subjects of the Northwestern Federal District by 1 billion rubles leads to an increase in total consumption expenditure by 1.37 billion rubles.

An increase in the interest rate on loans does not lead to a decrease in consumption, while the size of the interest rate on deposits leads to a decrease in current consumption. Moreover, if the level of expenses at the current moment of time increases by 1 thousand rubles, then we can expect an additional increase in expenses in the next period of time by 0.16 thousand rubles. This can be interpreted as follows: making decisions on consumption or savings by the population of the Northwestern Federal District is based on the propensity to consume. The more a person consumes, the more he will consume in the future, even if, in the absence of his own funds, he will have to use borrowed funds (an increase in the interest rate on loans does not lead to a decrease in consumption). At the same time, with an increase in the interest rate on deposits by 1%, the value of consumption expenditure per capita decreases by 191.19 thousand rubles per year, indicating that, holding all other factors constant, the decision to save prevails over the decision to spend if the deposit rate increases. At the same time, as noted earlier, an increase in consumption expenditure per capita leads to economic growth. Thus, the value of interest rates on loans and deposits can be one of the determining factors in the development of the region's economy. It is worth noting that an increase in consumption can lead to an increase in inflation, which has a negative impact on economic growth.

Fourth, changes in the consumer price index have a negative impact on consumption expenditure. When the cost of goods and services rises, the consumer can buy fewer of them for the same amount of money. An increase in inflation by 1% leads to a decrease in private consumption by 29.7 thousand rubles per year, which in general in the Northwestern Federal District leads to a decrease in consumption by 414 million rubles per year. This result can be explained as follows: if other factors are not taken into account, then when the cost of goods increases, their purchase is postponed for the following periods of time, and current consumption decreases.

Profit of Organizations Function

According to the results presented in Table 5 of the current value of the profit of organizations in equation (8) of the Menges model using a pooled regression model, we can conclude that an increase in current profit by 1 billion rubles in the current year may lead to an increase in profit by 0.96 billion rubles in the next period. It should be noted that due to the non-fulfillment of the third premise of the Gauss-Markov theorem, the standard error of the coefficient estimate can be underestimated, as a result of which the t-statistics, on the basis of which the conclusion about the significance of the linear regression coefficients, can be overestimated, and the conclusion about the significance of the coefficients may be unreasonable. At the same time, the pooled regression model is 85% explained by its lagged value, and the quality of the model specification is high. The fixed effects model for the profit of organizations function gives a significantly worse result. The explanatory power of the model is only 18%. The rest of the statistics are the same.

It should also be emphasized that the profit of organizations is included in the regional domestic product model as a significant

explanatory variable. On the one hand, an increase in the profits of organizations, according to the results obtained, does not lead to the growth of the region's economy, on the other hand, an increase in the profits of organizations leads to an increase in tax revenues, which positively affects the economic development of the region.

CONCLUSIONS

This study made it possible to identify the key factors that have a significant impact on the economic development of the Northwestern Federal District, the quantitative characteristic of which is the function of the regional domestic product. These variables are both factors included in well-known classical macro-econometric models (investment, consumption, taxes, inflation) and a new set of exogenous socio-economic variables that were not previously taken into account in the regional development analysis (household incomes, interest rates on deposits, and loans). The developed model allows us to draw important economic conclusions about the influence of each factor on the growth of the region's economy and the effectiveness of the government support measures.

An important research result is the developed model of the consumption function, which allows estimating the consumption expenditure levels depending on the change in the factors of the model. The study also made it possible to analyze the regulatory and government measures taken to curb inflation and support the population and businesses.

The paper proves that interest rates on loans and deposits can be one of the controlling factors in the development of the region's economy. In this regard, we can conclude that the government policy aimed at reducing interest rates, for example, on mortgage loans, aimed at supporting the industry, will, on the one hand, lead to an improvement in the social situation, an increase in the standard of living, and on the other hand, the debt burden of the population, which may adversely affect the economy of the region in the future.

At the same time, an increase in consumption can lead to an increase in inflation, which negatively affects the growth of the economy. The regulator's measures aimed at curbing inflation, including those associated with an increase in the key interest rate, lead to a change in interest rates on loans and deposits, which also affects the growth of the Northwestern Federal District's economy. The developed model makes it possible to evaluate the impact of the regulator's action on the growth of the economy of the Northwestern Federal District.

The revealed relationship between consumption expenditure per capita and the function of the regional domestic product of the Northwestern Federal District allows us to quantify the effectiveness of the measures taken to curb inflation and support the socio-economic development of the region. The models developed in the study make it possible to quantify the impact of social support measures for the population and business on the economic growth of the region. According to the data received, support for entrepreneurs and businesses did not have a significant impact on the economic growth of the region. Thus, it can be concluded that these support measures are ineffective. At the same time, the increase in social payments to the population and wages led to economic growth. If we compare the influence of each of the socio-economic factors on the growth of consumption and, as a result, the growth of the region's economy, then the return from wage increases is greater than from an increase in social payments. Consequently, measures aimed at raising wages are more effective for the economy of the Northwestern Federal District than measures related to increasing social benefits.

Answers to all the questions posed in the study were given, and all the hypotheses put forward were tested. In this paper, certain results have been achieved that determine the novelty of the research. For the first time, the paper investigates the relationship between socio-economic indicators of the Northwestern Federal District using panel

data and an analysis of the effectiveness of government measures to support the population and business, as well as the impact of the crisis associated with the pandemic on the socio-economic indicators of the region. For the calculations, we used data for the last 6 years (from 2015 to 2020), which made it possible to identify a clear pattern between the analyzed indicators within the region's economy. As a result, we can conclude that the modified Menges model is applicable to the analysis of the economic development of the Northwestern Federal District of the Russian Federation. This makes this paper important from both theoretical and practical perspectives.

ACKNOWLEDGMENTS

The article is based on the results of budgetary-supported research within the framework of the university-wide complex topic "Formation of conditions for long-term sustainable development of Russia: Theory and practice" for 2021–2025. Financial University, Moscow, Russia.

REFERENCES

1. Granberg A.G. Strategy of territorial socio-economic development of Russia: From idea to implementation. *Voprosy ekonomiki*. 2001;(9):15–27. (In Russ.).
2. Minakir P.A., Demyanenko A.N. Spatial economics: Evolution of approaches and methodology. *Ekonomicheskaya nauka sovremennoi Rossii = Economics of Contemporary Russia.* 2010;(3):7–25. (In Russ.).
3. Kleiner G.B., Rybachuk M.A. System balance of the Russian economy: Regional perspective. *Ekonomika regiona = Economy of Regions.* 2019;15(2):309–323. (In Russ.). DOI: 10.17059/2019–2–1
4. Aivazian S.A., Afanasiev M. Yu., Kudrov A.V. Methodology of socio-economic development assessment given the characteristics of regional differentiation. *Model Assisted Statistics and Applications*. 2020;15(4):311–314. DOI: 10.3233/MAS-200502
5. Bereznyatskiy A.N., Brodskiy B.E. Modeling macrodynamics of a Russian economic region. *Aktual'nye problemy ekonomiki i prava = Actual Problems of Economics and Law.* 2019;13(3):1273–1286. (In Russ.). DOI: 10.21202/1993–047X.13.2019.3.1273–1286
6. Egorov N.E. Innovative potential econometric estimation of Far East Federal District regions. *Izvestiya Dal'nevostochnogo federal'nogo universiteta. Ekonomika i upravlenie = The Bulletin of the Far Eastern Federal University. Economics and Management.* 2017;(1):43–50. (In Russ.). DOI: 10.5281/zenodo.386576
7. Tregub I.V., Buffet R.C. Managing methods of investment projects in the field of education on the example of France. In: Proc. 2019 12th Int. conf. "Management of large-scale system development" (MLSD 2019). (Moscow, Oct. 1–3, 2019). Piscataway, NJ: IEEE; 2019. DOI: 10.1109/MLSD.2019.8911089
8. Tregub I.V. Econometric analysis of influence of monetary policy on macroeconomic aggregates in Indian economy. *Journal of Physics: Conference Series.* 2018;1039:012025. DOI: 10.1088/1742–6596/1039/1/012025
9. Barro R.J., Chu A.C., Cozzi G. Intermediate macroeconomics. Andover: Cengage Learning EMEA; 2017. 512 p.
10. Aivazian S.A., Afanasiev M. Yu., Kudrov A.V. Models of productive capacity and technological efficiency evaluation of regions of the Russian Federation concerning the output structure. *Ekonomika i matematicheskie metody = Economics and Mathematical Methods.* 2016;52(1):28–44. (In Russ.).
11. Litvintseva G., Karelin I. Effects of digital transformation of the economy and quality of life in Russia. *Terra Economicus.* 2020;18(3):53–71. (In Russ.). DOI: 10.18522/2073–6606–2020–18–3–53–71
12. Hoang H.T., Huynh L.T.D., Chen G.S. How new economic geography explains provincial wage disparities: Generalised methods of moments approach. *Ekonomika regiona = Economy of Regions.* 2019;15(1):205–215. DOI: 10.17059/2019–1–16
13. Bravok P.S., Pynko L.E. Econometric analysis of the gross regional product of the Far Eastern Federal District. *Sovremennye tekhnologii upravleniya = Modern Management Technology.* 2020;(3):11. URL: https://sovman.ru/article/9311/ (In Russ.).
14. Kuznetsov S.V. The strategy of scientific and technological development of the North-West as an instrument of regional policy. *Ekonomika regiona = Economy of Regions.* 2011;(3):23–29. (In Russ.). DOI: 10.17059/2011–3–3

15. Grinchel' B.M., Nazarova E.A. Assessment of the sustainability of regional innovative development in the Northwestern Federal District. *Ekonomika i upravlenie = Economics and Management.* 2019;(2):20–27. (In Russ.).
16. Tregub I.V. Forecasting demographic indicators of the regions of Russia. In: Proc. 2018 11th Int. conf. "Management of large-scale system development" (MLSD 2018). (Moscow, Oct. 1–3, 2018). Piscataway, NJ: IEEE; 2018. DOI: 10.1109/MLSD.2018.8551856
17. Stroev P.V., Vlasyuk L.I., Makar S.V. Management of the microregion development: The South Pole of growth. *Ekonomika. Biznes. Banki = Economy. Business. Banks.* 2018;(2):109–123. (In Russ.).
18. Kashin V.K., Makar S.V. On the prospects of development of the RF Northwestern Federal District regions. *Regional'naya ekonomika: teoriya i praktika = Regional Economics: Theory and Practice.* 2017;15(1):19–34. (In Russ.). DOI: 10.24891/re.15.1.19
19. Frohn J., ed. Makroökonometrische Modelle für die Bundesrepublik Deutschland. Göttingen: Vandenhoeck & Ruprecht; 1978. 239 p.
20. Tregub I.V., Turkina S.N. Menges model applied to the economy of the USA. *Marketing i menedzhment innovatsii = Marketing and Management of Innovations*. 2014;(4):128–135. URL: https://mmi.fem.sumdu.edu.ua/sites/default/files/mmi2014_4_128_135.pdf
21. Tregub I.V. Econometrics. Model of real system. Moscow: PSTM; 2016, 164 p.
22. Tregub I.V., Dremva K.A. Estimating the consequences of Russia's and the EU's sanctions based on OLS algorithm. *International Journal of Machine Learning and Computing.* 2019;9(4):496–505. DOI: 10.18178/ijmlc.2019.9.4.832

ABOUT THE AUTHORS

Ilona V. Tregub — Dr. Sci. (Econ.), Prof. Department of Mathematics, Financial University, Moscow, Russia
https://orcid.org/0000-0001-7329-8344
Corresponding author
itregub@fa.ru

Margherita De Iaco — postgraduate student, University of Studies Turin, Turin, Italy
https://orcid.org/0000-0002-0231-9598

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 30.11.2021; revised on 15.12.2021 and accepted for publication on 17.12.2021.
The authors read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-146-156
UDC 657(075.8)(045)
JEL M48, H56, D73

Forming Strategic Directions for Development and Reforms of the State Financial Control in the Republic of Tajikistan

D.A. Davlatzoda
Russian-Tajik (Slavonic) University, Dushanbe, Tajikistan

ABSTRACT

This paper considers the problems of state financial control, which are relevant for the government of any country. **The aim** of the study is to develop practical proposals to improve the effectiveness of the state financial and control authorities in the Republic of Tajikistan. The author uses **the methods** of analysis, generalization and comparison to identify shortcomings of the state financial control in the Republic of Tajikistan. The study presents an analysis of international experience in this area and some aspects of the Lima Declaration, which have been applied in practice and showed their effectiveness. **The scientific novelty** of the study lies in the fact that the author proposes measures to improve the state financial control activities of the Republic of Tajikistan and the Accounts Chamber. Among them: increasing the efficiency of using public funds, creating an integral system of state financial control, preventing duplication of powers and functions of state financial control bodies, introducing international experience in the field of state financial control and measures to improve laws and regulations in the field of state financial control, and creating a mobile and efficient system of state financial control. The author **concludes** that it is necessary to adjust the financial control system of Tajikistan to ensure its relevance and compliance with the requirements of the country's current policy objectives. The author's recommendations can be used to improve the regulatory framework of the state audit in the Republic of Tajikistan.

Keywords: Republic of Tajikistan; financial and control activity; state audit; Accounts Chamber; state budget; international standards; Lima Declaration

For citation: **Davlatzoda D.A. Forming strategic directions for development and reforms of the state financial control in the Republic of Tajikistan.** ***Finance: Theory and Practice.*** **2022;26(3):146-156. DOI: 10.26794/2587-5671-2022-26-3-146-156**

INTRODUCTION

State control is one of the main provisions of the general control system. It determines the importance of monitoring the work of business entities located on the territory of the Republic of Tajikistan, including monitoring the performance of official duties by civil servants. The purpose of such control is primarily to improve the work of the management apparatus, in particular the voluminous structure of the financial apparatus, and secondly — to expand and specify the area of responsibility of State officials in the performance of the tasks assigned to them. A key feature of state control is that it should be directly related to those state bodies and their supervised institutions that deal with the formation of the budget, making proposals for its adjustment, approving objects of expenditure, and budgetary allocations to their recipients. Monitoring should be comprehensive not only for all financial stages regarding the assets of the state, but also for a huge list of state bodies and services, for example, customs, tax authorities, banks, etc. This is the particular importance and necessity of State control, which will avoid corruption schemes, ensure accurate and uninterrupted delivery of the full amount of money to the entities to which they were intended.

The State needs more active participation in anti-corruption activities, as well as in the fight against economic offences in the Republic of Tajikistan. Establishment of financial and economic controls that help to ensure a balance between the State and its agents is paramount [1, p. 86]. The function of State control is obligatory and inherent in any developed State, as it not only ensures the development of the country as a whole, but also affects the effectiveness of its functioning [2, p. 35, 36].

In the period of the pandemic in the Republic of Tajikistan, the need for financial control has increased [3, p. 16]. It is essential to adjust the management of financial assets as well as to establish a scheme for their effective use. Therefore, such directions state "control", "independent financial audit" and "management activity" can be called priorities for the development of the country. It is necessary to consolidate and delineate these areas in the legal and regulatory acts of Tajikistan, which will contribute to the objective assessment of the wrongfulness of the actions of the entities, as well as the identification of economic violations leading to consequences at the legislative level [4, p. 6].

MATERIALS AND METHODS

Statement of the problem

The establishment of State control in the Republic coincided with the development of the State and the transition to market relations. Unfortunately, the development of State financial control in the early years of independence was not given due attention, which resulted in many years of delay in the formation of its complete system [5, p. 31, 32].

It is now known that a weakening of State control could have the following consequences:

- budget billions can be spent for other purposes or embezzled;
- hundreds of thousands of depositors or investors can be cheated;
- months of non-payment of wages can become almost the norm.

At the same time, a large number of private enterprises and organizations have appeared in the Republic. They have different sources of financial resources from the state structures, they are guided by their principles of lending, pricing, money settlement, income distribution, etc. Overcoming the former principles of party control and the class approach to it, the so-called worker, the people's control characteristic of the totalitarian society, the elimination of excessive centralization has led to the creation of new public financial supervisory bodies. Their task — on the basis of appropriate control and expert-analytical measures, to focus on the existing problems of the use of public funds and property, to analyze the efficiency of budget expenditures and, as a result, to offer reasonable proposals

to optimize them, eliminate violations and shortcomings that occur in the industry. The need to increase the level of control over the State's finances is due to a number of complicated factors, such as the need to use the Bank's services to serve the purposes of the State, etc.

Effective control in the financial sphere for the coordination of state funds — is a main determinant of the establishment of a trusting relationship between society and the state. A high level of public trust is achievable when there is no pressure from external factors and various bodies.[1]

In reviewing the current system of financial control in Tajikistan, the author concluded that it needed to be improved, as it did not correspond to changes in many aspects and activities.

RESULTS OF THE STUDY

Due to the legislative gap, there is a need to solve this problem through the creation of a core document that is well thought out: concept, objectives, tasks, functions of specific bodies of state financial control, features of their interaction.

It was also important to ensure the adoption of uniform, no contradicting each other laws, which required adjustments in practice [6, p. 177]. Local regulatory legal acts and recommendations contain only layers of information, but do not reflect it in its full, proper form.

However, with regard to the concept of financial control, the existing draft laws do not regulate its adequate implementation with respect to State enterprises, the budget, the expenditure on the maintenance of State-owned joint stock and private enterprises, Legal status of State and local supervisory authorities. Supervisory staff are not given sufficient attention and are not guaranteed independence and protection from interference by government agencies [7, p. 370]. The only thing that is in almost all laws — the unjust desire of their authors to solve departmental problems in their own interests.

The moment regarding when it is necessary to draft and adopt a law " On State Financial Control",[2] which would include the following sections:

1. **General provisions.**

The section should contain a clear definition of State financial control, justification of its necessity, types, forms and basic methods of work.

2. **Objects of State Control.**

The section should define the objects of control by group.

3. **State financial supervisory bodies, their powers and areas of activity.**

Status is clearly established: the Accounts Chamber; the Agency for State Financial Control and the Fight against Corruption; organizations and institutions containing at the expense from the State budget of the Republic of Tajikistan; local budgets; state and trust funds; state economic organizations; joint-stock and private enterprises.

4. **The principal effects of state financial control.**

General tasks in the context of control bodies are detailed.

5. **Rights and duties of employees of State financial control bodies.**

Defines the powers, rights and duties of the employees of the control services in the exercise of their control functions; guarantees their independence from the intervention of the authorities and State protection in case of unlawful actions.

The Coordinating Council, chaired by the President of the Court of Accounts, should be established or at least considered. The state of control of the agencies should be audited by the Agency for State Financial Control and the Fight against Corruption. In addition, emphasis should be placed on the development

[1] Eurasian fund for Stabilization and Development. Republic of Tajikistan. Recent social and economic trends and short-term prospective. September, 2019. URL: https://efsd.eabr.org/upload/iblock/c1e/Eng_RT_macro_2019Q2.pdf (accessed on 26 June 2021).

[2] Law of the Republic of Tajikistan from 02 December 2002 No. 66 "On the State Financial Control in the Republic of Tajikistan". URL: http://minfin.tj/downloads/ogfkvr.pdf (accessed on 24 June 2021).

of a monitoring and audit plan. It is necessary to divide the control into departments with delimitation of their specifics:

1. Department of Budget Control, which will define tasks and requirements, as well as verification methodology and auditing techniques for operations and the use of funds from the budget.

2. Department of Material Assets Control, which provides details on the conduct of features of inspections and audits of companies.

3. Department of Audit to preserve values of a material nature, as well as cash, which will consider the need for inspections, their ways and methods.

4. Department of Audit Output. It should set out the formalities required for certain facts and violations.

Time should be devoted to actual disclosure of inspections and audit information. The first step is to establish a number of requirements for materials that are the result of or are part of the checks, including how they are implemented by the financial control authorities. Specific requirements should be established for the heads of all types and levels of organizations regarding their actions and reactions in obtaining instructions from the monitoring bodies.

It is important to pay attention to the establishment of criteria for evaluating the implementation of decisions taken on the basis of the measures to compensate for losses or to bring the subject of the inspection into the required state, while distinguishing the functions of the police, courts, prosecutor's office, etc.

It would be better for the supervisory authorities to hold the heads of services accountable on a material basis. Inspectors should have the authority to dismiss a leader from their posts in extreme cases.

The question arises: how to be, if the verification results do not satisfy the subject? Answer is simple: provide for appeals against decisions or even specific actions by monitoring bodies. Prosecution of the verifiers is possible for errors indicated in the documents, illegal actions, lack of objectivity and other actions to the detriment of the organization or employees. And it should be noted that all employees of the monitoring bodies are obliged to have higher economic education, as well as timely to take advanced training courses

Financial motivation of monitoring bodies needs to be adjusted. In particular, the question of the reimbursement of business travel costs, which has not yet been regulated, can be highlighted.

It is also important to note that the Republic of Tajikistan lacks adequate control over the revenues of the State and local budgets. The fact that it is connected exclusively with the tax committee — is inappropriate. This service ensures that payments are correct and that taxes are paid on time. But the economic side of the taxpayer's activity is not properly monitored by the Ministry of Revenue and Charges, as it has other tasks and functions. This was repeatedly brought to the attention of specialists of the Court of Accounts of the Republic, experts of the World Bank, but, contrary to logic, with the adoption of the Tax Code,[3] the verification of tax payments became the exclusive right of the tax committee, and customs duties — customs authorities, etc. In essence, certain areas of State financial control have been monopolized in the absence of a socially transparent control procedure [8, p. 365–366].

The Law of the Republic "On the Court of Accounts"[4] generally complies with the Lima Declaration on the guidelines for financial control,[5] adopted at the IX INTOSAI

[3] The Tax Code of the Republic of Tajikistan from 17 September 2012 No. 901 (with changes and additions on 17 December 2020). URL: https://online.zakon.kz/Document/?doc_id=31270626 (accessed on 26.06.2021).

[4] Law of the Republic of Tajikistan from 28 June 2011 No. 749 «On the Accounting Chamber of the Republic of Tajikistan». URL: https://www.sai.tj/index.php/ru/zakonodatelstva/zakonodatelstvo-o-schetnoj-palate (accessed on 24.06.2021).

[5] International Standards of Supreme Audit Bodies (ISSAI) issued by the International Organization of Supreme Audit Bodies (INTOSAI). URL: https://www.intosai.org/fileadmin/downloads/documents/open_access/ INT_P_1_u_P_10/issai_1_ru.pdf (accessed on 24.06.2021).

Congress in October 1977, although some of its provisions are not reflected in the law. However, there is still no legal regulation of independent financial control by the Court of Accounts of the formation of income from the State budget of the Republic and the use of budgetary funds [9, p. 157]. The resolution of this issue stretched for 7 years due to the interest of individual officials, which caused significant economic losses to the State. The reasons for this attitude to the crucial issue should be identified separately.

"Out of the game" and the state audit, although the need to carry out such control is confirmed by statistics, which shows that every second enterprise of the Republic is unprofitable.

A special problem is the control over the use of state and municipal property, the finances of enterprises operating on its basis. On the part of the owner of the property — the State — the activities of such enterprises in terms of generating profit and using it are not managed, huge financial flows are not controlled. Neither the Audit Chamber, nor the bodies of the State Financial Inspectorate of the Republic, nor the financial authorities do not carry out the State audit on the use of the specified property. The Committee of State Property of the Republic has partially such functions, it has the right to exercise the powers of the owner with respect to the shares of joint-stock companies that have not been realized in the process of privatization. However, he does not use this right, because he is too busy with other tasks, and he lacks specialists in financial matters. Such analysis can be carried out by audit firms, but their services are very expensive and audit reports are confidential. Therefore, without appropriate legislative changes, reports cannot be provided neither the Ministry of Finance nor the Government nor the Majlisi namoyandagon of Majlisi Oli of the Republic of Tajikistan [10, p. 1].

In my opinion, all companies and individual enterprises (IE) in the country need regular inspections. It is important to give legal status to the auditor's opinion as well as to define the scope of application of such a document. This approach is based on the division of State control into two parts: control by the Government and control by Parliament.

Both types of control should have their own powers, although this has been a controversial issue for many years, as the central supervisory authority has not yet been defined and therefore this must be resolved.

Tajikistan does not currently have a single competent module of State financial control bodies. Based on local legislation, we can only conclude that the Audit Chamber is such a body. However, in the absence of well-thought-out laws and State financial control strategies may be associated with this function by some group of bodies, but not by one. Given the separation of the functions of the above-mentioned group, it is worth clarifying that most of the bodies themselves demand control and do not relate to those carrying it out.

The Audit Chamber is confused about the content of some draft laws that include parliamentary control as a minor procedure. The development of the Charter of the Audit Chamber, possibly taking into account the principles of INTOSAI, would provide a stable position. If it reflected the possibilities, tasks and functions of the Audit Chamber, questions concerning the supremacy of the organization over others would cease to exist.

The work of the Audit Chamber has a vector for tracking the usefulness, quality of cash flow and monitoring the use of funds received from the budget. This leads to the conclusion that the authority does not check the funds, but the appropriateness of their use, including the property of the State.

There is now a need to legislate for new functions, for example:

- risk assessment applicable to financial and economic security;
- carrying out initial analysis and further control of the state programs to guarantee the security of the finances and economy;
- supervision over the financial transactions;

• development of methods of carrying out control measures, and ways of interpreting the received information [11, p. 100];

• facilitating the exchange of experience across countries on the issue of financial implications;

• verification prior to the sale of property of significance to the State;

• leading the establishment of standards in other countries and controlling the Republic's finances.[6]

The Audit Chamber should regularly review and verify the implementation of its decisions by the chief budget managers.

Most budget expenditures for a given year are already regulated by the laws and financial obligations in force, which limits the ability of the Majlisi namoyandagon of Majlisi Oli to make decisions on budgetary issuesIt is therefore important that the Majlisi namoyandagon of Majlisi Oli receive advice and the results of inspections from the Audit Chamber on long-term financial obligations. Provision of relevant recommendations to the Majlisi namoyandagon of Majlisi Oli is also important for medium-term financial planning purposes.

The Audit Chamber takes part in the preparation of the State budget of the Republic, usually by means of comments and proposals, which are advisory in nature. However, the Audit Chamber does not have the authority to change the budget estimates made by the executive branch. The proposals of the Audit Chamber are not binding. The involvement of the Audit Chamber in the budget process should not compromise its independence in the subsequent monitoring of budget implementation. The Audit Chamber will gain more independence when advising the Majlisi namoyandagon of Majlisi Oli if its budget funding is outside the influence of the executive power.

The Audit Chamber should be more actively involved in the financial examination of draft laws related to the State budget and extrabudgetary State funds (institutions), international treaties to which the Republic of Tajikistan is a party, as they are ratified by the Majlisi namoyandagon of Majlisi Oli.

The State's financial control sector is not provided programmes and information at the proper level, communication is not concentrated in one particular body, but is divided among different actors. In addition, the content of the work of the financial supervisory authorities and the range of options prescribed in the various instruments have not been adapted to the new criteria for financial control.

In order to eliminate the problem of information exchange, it was decided to create an electronic platform in which different programmes that are not interconnected can be used. The problem has not been solved altogether, the exchange of information is still difficult, but the prerequisites for overcoming this shortcoming have emerged.

The next problem is that all budgetary procedures are not covered by the permanent state financial control: from the beginning of the formation of budget requests, approval of relevant budgets and estimates to the financing of expenditures, payment of bills and actual use of funds.

It is necessary to establish an effective system of internal financial control of the State. Currently, it is carried out by the Audit Chamber of the Republic of Tajikistan, the Agency for State Financial Control and Combating Corruption, the Tax Committee, etc. [12, p. 37].

Priority in internal financial control is process of establishing and using budget allocations.

Control of the efficiency of the application of the revenue and expenditure parts of the budget and the use of the property of the State involves the following factors:

• organizational and supervisory activities in terms of the speed of implementation of the income and cost parts of budgets, and State off-budget funds, based on the volume, structure and purpose [13, p. 212];

[6] Centre for Financial Reporting Reform. Internal audit key performance indicators. URL: https://documents1.worldbank.org/curated/en/840271546848703532/pdf/Internal-Audit-Key-Performance-Indicators.pdf (accessed on 26.06.2021).

• determination of rationality of costs and use of state property;
• assessment of the reasonableness of income and cost component of budgets.

The Audit Chamber found a significant number of irregularities during the internal control process, of which the key ones are:

a) costs incurred in contravention of current Tajikistan legislation, including failure to comply with the conditions for writing off cash against cash expenditures of budgetary institutions or spending in excess of specified period limits;

b) unreasonable calculation of cost rates, which budget was overstated to its recipients;

c) reduction of the group of "fixed assets" and other inventory items in accounting;

d) payroll in excess of number of services rendered or work performed;

e) no capitalized transactions of material or cash value;

f) Illegal payment of bonuses, pensions, benefits and other payment;

g) cash shortfalls by State-owned companies in the disposal of assets arising from undervalued purchases and sales;

h) other illegal expenses.

The problem of control over budget funds and public expenditures remains today for various reasons quite complex and requires an accelerated solution.

Results of the auditor's check, which is carried out as control of budget implementation and use of budget funds, suggests that the unsatisfactory pre-monitoring and ongoing monitoring of budgetary resources and demonstrates the need for a coherent system of financial and budgetary control. The introduction of new State institutions and changes in budgetary technologies require the development of an adequate system of financial and budgetary control, ensure a coherent system of control over the management of public resources, to coordinate the activities of all competent authorities of the State in the organization and implementation of public expenditures and to ensure the mobilization of budgetary resources [14, p. 210].

The principles of control set out in the comments to the Law on the State Budget of the Republic are the basis for the formation of a coherent system of control over State resources, funds and property in the preparation of a system legislative act on the organization and implementation of State financial and budgetary control.

Another legal problem of the system of state financial control is the uncertainty of its impact on the management of resources at the municipal level.

Many heads of local authorities are of the view that the State is not responsible for their financial obligations and therefore some State financial control should not be discussed. However, this is not the case, because it does not mean that the territorial collectivity carries out economic activities on equal terms with other economic entities.

It is logical that the leadership of the Republic should ensure reliable external and internal financial control, compliance of regional bodies with similar restrictions on financial activities. However, there is currently no legislative definition of who should exercise this control and to what extent.

Another problem of state financial control is the uncertainty of its forms and methods, the unpreparedness to act in the context of global transformations in the global economic mechanism. In order to bring the system of state financial control in the Republic in accordance with the norms and rules of EU legislation, it is necessary to adopt "Concept of reform of the State financial control system" and make significant changes to the procedures and mechanisms of the entire system.

The modern system of state financial control and audit has many components, the organizational structures of such a system, existing in different countries of Europe, may differ. However, the system of mechanisms, procedures, approaches and standards is common to all EU countries.

The main parts of the European Union's internal public financial controls are:

• cash management and control;

• internal audits;
• harmonizing the central apparatus.

There is an opinion on the need to develop one system and external control of finance. Control is now fragmented into organs, but it is not one. This allows each entity to act within its own competencies, carrying out the tasks assigned to it, which has a number of disadvantages. Among them — similarity of actions, poorly built communications, inconsistencies in different areas, etc. It reduces productivity, takes a long time to discover the truth and fix problems. As a result, each can pass the test once, and one — not once.

Control of finances requires updating, getting rid of contradictions, harmonization of activities with laws. These measures must be taken now. Developed countries have long used experience, where not only organizations, but also the disposal of State funds is subject to control.

State control of the financial sphere is essential for regulation, because it touches aspects of the sustainability of the economy in a country, so that the object of control increases and acquires gets new paint [15, p. 137, 138].

It should be noted that today it must be recognized that the inconsistency of actions and elementary errors have not yet allowed the Republic to establish an effective, market-based State system of financial control [16, p. 456]. The official position on the system has not yet been clearly established. Although, as a rule, the Constitution of each democratic country determines the procedure for the establishment of the supreme control body and the necessary degree of its independence and rights, the powers of control and other details are established by the relevant law. According to the Constitution,[7] the Audit Chamber of the Republic performs State control functions and manages through independent information. The concept of a system of public financial control has not yet been adopted, nor does it exist in existing laws. These circumstances have a negative impact on the nation-building process. In addition, much of the control of public funds is now concentrated in the hands of the executive authorities, which are responsible for the administration of the funds. This limits the rights and oversight powers of the financial control bodies, including the Audit Chamber, and prevents control of various areas of public resource management.

The situation cannot meet the needs of the State or the expectations of society. Therefore, one of the ways to reform the system of public financial control is to further develop external independent control in the management of public resources.

According to the "Lima Declaration of Control Guidelines" as a strategic direction of development of the state financial control, the transition from the so-called regulatory audit (verification of compliance with established standards) to audit of effectiveness is considered. It is well known that the cost-effectiveness assessment is based on two main elements: result, the extent to which a particular objective is achieved and the resources required to achieve it. The ratio determines the effectiveness of an activity.

CONCLUSION

Thus, the most important condition for the organization of state financial control to determine the effectiveness of the use of public resources is the presence of an appropriate target, or, in other words, a clear understanding of what goals and objectives we monitor. Unless the goals and objectives are understood and the means to achieve them are understood, it is impossible to determine the effectiveness of the formation and use of public resources [17, p. 10].

The system of State financial control should be holistic only because it should be based on uniform principles, norms and standards of operation and performance, with clear definition of control bodies and allocation of their functions and duties. Such a system

[7] The Constitution of the Republic of Tajikistan (at the edition referendum from 26 September 1999, from 22 June 2003, from 22 May 2016). URL: https://www.mfa.tj/ru/main/tadzhikistan/konstitutsiya (accessed on 24.06.2021).

should not be hierarchically subordinate and hierarchical. Each body should carry out its specific tasks, have its own auditees and act according to the same methodological and organizational principles [18].

The next important condition for public financial control is its unity and integrity. Interaction of the bodies of the system, coherence of their actions will guarantee maximum transparency of budget movements, objective and comprehensive information, as well as proper monitoring of the elimination of violations and implementation of proposals.

The control of finance throughout the system should be oriented towards the management of the bodies exercising this type of control. Here it is necessary to provide for the delimitation of authorities, making the function "individual". Interestingly, the tasks that arise involve the use of various powers, which obliges the delineation of functions.

If there is no consistency in actions, the duties are multiplied and carried out by different bodies repeatedly, although the peer work on the projects increases the quality, freshness of views, depth of analysis.

The development of one system would, a priori, add value to the activities of the bodies that are part of it. To do this, you need to define the hierarchy, clarity of actions, responsibility, functions. But do not forget that the leading role in the control of finances belongs to the legislator.

Since financial control laws are imperfect, it should be noted that:

1. There is a need to adjust the information showing the current status of cash amounts, their targeted spending during the execution of the budget, as well as the results of actions and performance analysis.

2. It is important to introduce new principles and change the previously adopted framework regarding internal communications in the bodies, which contributes to their development.

3. It is necessary to separate external communications as "special".

4. International experience should be taken into account and attempts should be made to incorporate it into their own monitoring system.[8]

The main objective of State control is to ensure the stability and economic security of the State. To achieve this, it is necessary:

- first, to increase the revenue of the state budget;
- second, save budget expenditure;
- third, reduce the number of financial offences;
- fourth, reduce the level of corruption in the State.

All this should be the basis for defining the main tasks of the state financial control bodies. Deleting at least one of them would not achieve the main objective of public financial control.

Public financial supervision must become comprehensive. The more budget funds come under control, the more complete and quality information will be provided to the public authorities and the public; unaudited matters by control officers will be reduced and, therefore, budget violators will not go unnoticed.

Also, the system of state control should be mobile, i.e. operational, adequate and professionally responsive to any changes in the economy and legislation; ensuring the fullest possible prior, ongoing and follow-up monitoring.

Attention should be given to the first warning of financial discipline violators in order to avoid undesirable consequences in the future.

Taking into account the above-mentioned prerequisites for reforming State control, it is advisable to address its problems in the following areas:

- formation of a common conceptual framework for the organization of State financial control, taking into account the modern conditions of development of the economy of the Republic;

[8] Implementing the International Standards for Supreme Audit Institutions (ISSAIs): Strategic considerations URL: https://docplayer.net/19107167-Implementing-the-international-standards-for-supreme-audit-institutions-issais-strategic-considerations.html (accessed on 26.06.2021).

• creation of a single legal field for the development of the system of State financial control and legislative consolidation of its subjects' control functions;

• establishment of new and improved public financial control procedures;

• improvement of methods of state financial control;

• optimization of State financial control organizational structures;

• improvement of staffing of the State financial control system, creation of a modern research and training base on international standards;[9]

• creation of an adequate modern information and communication infrastructure;

• improved logistics and financial management of the State financial control system.

The use of these recommendations will create a strong management system, which minimizes lawlessness in the budget, modernizes the management of the state's finances, will lead to the optimization and implementation of those standards, which have long been accepted by developed countries and can form the basis of a successful system.

[9] Performance Audit of Guidelines: ISSAI 3000–3100. Standards and guidelines for performance auditing based on INTOSAI's Auditing Standards and practical experience. URL: https://www.auditorgeneral.gov.tt/sites/default/files/Performance%20Audit%20Guidelines.pdf (accessed on 26.06.2021).

REFERENCES

1. Zdyrko N. Problem aspects of the internal state audit development. *Ekonomichnii chasopis-XXI = Economic Annals-XXI*. 2016;161(9–10):85–90. DOI: 10.21003/ea.V161–19
2. Dikan L., Shevchenko I. The development of a typology of state audit in the international and national practice. *Ekonomika rozvitku = Economics of Development*. 2016;(3):34–40.
3. Chernyakov M., Muhtorzada S., Akberov K., Shuraev I., Mammadli F., Rzyaeva H., Huseinova K. The state of theory of financial stability in the Republic of Tajikistan. *Deutsche Internationale Zeitschrift für Zeitgenössische Wissenschaft/German International Journal of Modern Science*. 2021;(10–2):13–17. DOI: 10.24412/2701–8369–2021–10–2–13–17
4. Igibaeva Z., Beiseinova L., Kazhmuhtametova A., Nikiforova E. Modern trends of Kazakhstan's internal state audit: The US and UK experience. *Problems and Perspectives of Management*. 2020;18(2):1–12. DOI: 10.21511/ppm.18(2).2020.01
5. Loseva N.A., Zabolotina A.V. State audit as a form of financial control budgetary funds. *Contemporary Problems of Social Work*. 2017;3(1):31–38. DOI: 10.17922/2412–5466–2017–3–1–31–38
6. Tsarenko O.V., Karmazina N.V. Methodical techniques of system research of the state financial audit process of budgetary establishments and organizations. *Naukovii visnik Polissya = Scientific Bulletin of Polissia*. 2016;(3):176–180.
7. Rustamov D. Yu. Constitutional and legal regulation of state financial control in the field of budgetary relations of the Republic of Tajikistan. In: Scientific works. Russian Academy of Legal Sciences. Moscow: Yurist; 2019:368–371. (In Russ.).
8. Issataeva K.B., Adambekova A.A. The Republic of Kazakhstan budget system development and the increase of its transparency. *American Journal of Applied Sciences*. 2016;13(4):364–371. DOI: 10.3844/ajassp.2016.364.371
9. Davlatzoda D.A., Sakhibov E.N. Features of the establishment and development of the institutions of financial control on the territory of modern Tajikistan. *Trud i sotsial'nye otnosheniya = Labour and Social Relations Journal*. 2019;30(4):150–162. (In Russ.). DOI: 10.20410/2073–7815–2019–30–4–150–162
10. Endrawes M., Feng Z., Lu M., Shan Y. Audit committee characteristics and financial statement comparability. *Accounting & Finance*. 2020;60(3):2361–2395. DOI: 10.1111/acfi.12354
11. Mamutova K., Maulenkulova G., Toktibaeva M. Improving the application of the risk management system in the field of state audit. In: Industrial technologies and engineering ICITE-2020. Proc. 7th Int. conf. (Shymkent, Nov. 12–13, 2020). 2020:100–105.
12. Zokirova F.D. Agencies for state financial control and combating corruption in the system of state financial control of the Republic of Tajikistan. In: Topical issues of economic theory: Development and application

in practice of Russian reforms. Proc. 7th Int. sci.-pract. conf. (Ufa, May 25–26, 2018). Ufa: Ufa State Aviation Technical University; 2018:36–40. (In Russ.).
13. Sembiyeva L.M., Zhagyparova A.O., Kozhubekova A.A. Analysis of the state of the current state audit model in Kazakhstan at the present stage. *Ekonomicheskaya seriya Vestnika ENU im. L.N. Gumileva = Economic Series of the Bulletin of the L.N. Gumilyov ENU*. 2020;(3):208–215. DOI: 10.32523/2079–620X-2020–3–208–215
14. Shokhin S.O. Legal organization of budget control. In: Zapol'skii S.V., ed. Management of budgetary resources and revenues (legal aspects). Moscow: Prometheus; 2017:206–245. (In Russ.).
15. Schevchenko I. Factors influence on the condition of the development of the state financial audit of budget programs. *Veles*. 2016;(9–1):136–139.
16. Davlatzoda D.A. Formation and legal basis of financial control institutions' activities in the Republic of Tajikistan. *Aktual'nye voprosy sovremennoi ekonomiki = Topical Issues of the Modern Economy.* 2020;(9):455–459. (In Russ.). DOI: 10.34755/IROK.2020.50.67.063
17. Trofimov M.V. On the problems of improving financial control in the sphere of functioning of public authorities of the Russian Federation. *Sovremennoe pravo = Modern Law.* 2004;(7):6–10. (In Russ.).
18. Tlipiev A.T. Constitutional and legal foundations of state financial control in Russia (issues of theory and practice). Cand. legal sci. diss. Rostov-on-Don: Rostov State University; 2003. 181 p. (In Russ.).

ABOUT THE AUTHOR

Dilmurod A. Davlatzoda — Cand. Sci. (Econ.), Assoc. Prof., Department of Management and Marketing, Russian-Tajik (Slavonic) University, Dushanbe, Tajikistan
https://orcid.org/0000-0003-0115-7731
d.a.davlatov@mail.ru

Conflicts of Interest Statement: The author has no conflicts of interest to declare.

The article was submitted on 08.08.2021; revised on 20.08.2021 and accepted for publication on 27.12.2021.
The author read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-157-168
UDC 336.012.23(045)
JEL G41

Financial Behavior as a Result of People's Interaction in the Socio-Economic Space

S.V. Makar[a], A.V. Yarasheva[b], D.I. Markov[c]
[a,c] Financial University, Moscow, Russia; [b] ISESP FCTAS RAS, Moscow, Russia

ABSTRACT

Financial behavior is a category and its relevance for Russia is growing rapidly along with the activity of people who carry out daily interactions in a dynamic socio-economic space at all its territorial levels (local, regional, national, and supranational). In this context, the authors emphasize the scientific significance of studying the process of behavior transformation leading to financial results. From the point of view of the authors, this is due to the understanding of the primary and secondary interdependencies, the coherence of the concepts used, their sequence, priority, and significance. **The aim** of the article is to show the connectedness of elements in the behavior of people who are actors simultaneously in three subspaces: social, economic, and financial, which are organized hierarchically. The goal is due to the diversity of people's financial behavior and the need for a counter study of both the factors influencing it and the analysis of monitoring the intertwining forms of financial behavior. The authors use **methods** of spatial analysis, content analysis, and sociological factor analysis. The factors of influence are placed in the "pyramid" of subspaces, which contributes to their prioritization, the vectors of people's interaction, the economic, socio-demographic determinism of their actions, and the territorial differentiation of the financial infrastructure, and other aspects are also highlighted. The article systematizes scientific approaches to the study of the economic and financial activities of the population from the point of view of the relationship of incentive motives which simultaneously affect decision-making. The study identifies structural elements of each of the three subspaces (relationships-resources-choice-experience), within which a person (social group) forms and implements a particular model of behavior. The main theoretical concepts used in the study of financial behavior are singled out: social relations, economic resources, financial strategy, and their relationship. The authors **conclude** that the financial behavior of people has a spatial structure. Its result is determined by the categories of three subspaces — social, economic, and financial, which together represent the "matryoshka" model. The proposed methodology is significant for identifying the characteristics of the financial behavior of young people and the factors influencing the transformation of their financial literacy.

Keywords: financial behavior; social subspace; economic subspace; financial space; economic resources; income and expenses; financial infrastructure; financial literacy

For citation: **Makar S.V., Yarasheva A.V., Markov D.I. Financial behavior as a result of the interaction of people in the socio-economic space.** ***Finance: Theory and Practice.*** **2022;26(3):157-168. DOI: 10.26794/2587-5671-2022-26-3-157-168**

INTRODUCTION

Based on the three most important spheres in which people simultaneously handle their daily activities, we distinguish three subspaces: social, economic, and financial. The first of them is the widest in terms of scale (volume) and points of application of people's actions, and the second and third (the narrowest) — in the form of an inverted pyramid "enter" the first subspace. Financial behavior, according to the authors, is the most concentrated result of the active interaction of people — in terms of justifying choices and making complex decisions. In this context, the study of the transformation process under the influence of factors in the behavior of the population (economic and social), leading to a certain financial result obtained under the influence of a combination of both rational decision-making and irrational motives for choices, acquires scientific significance.

RESEARCH TOPIC DEVELOPMENT

The study of financial behavior over the past few decades has been relevant in Russia and abroad due to the constantly (and relatively quickly over the past 2 years) changing economic and social conditions. The determinism of personal financial activity by belonging to different social income and demographic groups, territories of residence (with different levels of economic development [1] and the availability of financial infrastructure [2]), as well as the level of culture, including financial [3], is of particular interest [3].

Scientific approaches to aspects of people's financial behavior are reflected in publications of various levels: from Nobel Laureates in Economics to young professionals. Thus, D. Kahneman [4], G. Akerlof [5], R. Shiller [6], and other well-known scientists who analyzed the features of financial behavior considered the conditions and motives of decision-making as a complex of simultaneously acting factors, including emotions.

R. Thaler, whose works on behavioral economics are recognized as classics, carried out theoretical developments at the intersection of disciplines: economics, psychology, sociology, culture, and other social sciences [7]. The nudge concept popularized by R. Thaler and C. Sunstein [8], showing the role of external (social, economic, environmental) and internal (psychological, emotional) incentives in human decision-making (the so-called architecture of choice), found its supporters and critics [9].

The well-known psychologist D. Ariely, a specialist in behavioral economics, in his writings provided the results of experiments to identify the mutual influence of market and social norms in relations between people, the influence of the "delay" effect on all types of human activity [10].

In connection with the new trends in the behavior of individuals (especially at the present stage of the economic downturn under the influence of the coronavirus pandemic [11]), it turned out to be in demand to adapt the economic person model to changing conditions (decrease in income levels; instability in the labor market; an increase in the amount of information and a change in its sources, as well as confidence in it; the personification of the economy; request for emotions as a new type of economic goods, etc.) [12].

Annual monitoring studies (2019–2021) of the behavior of social income groups of the population, conducted by scientists from the at the Institute of Socio-Economic Studies of Population of the Federal Center of Theoretical and Applied Sociology of the Russian Academy of Sciences (hereinafter referred to as "ISESP FCTAS RAS") [13], demonstrate the dynamics of the activity of Russians in the consumption, savings, credit and investment sectors. Gender differences in the financial behavior of the population are of interest [14].

The annual international conferences "Incomes, Expenses and Savings of the Russian Population: Trends and Prospects",

held at the ISESP FCTAS RAS (2014–2021) [15], gather for scientific discussion scientists involved in behavioral economics not only from Moscow and regions of Russia but also from Belarus, Tajikistan, Poland, Japan, etc.

Scientific projects of the Financial University under the Government of the Russian Federation (R&D: "Institutionalization of Financial Literacy of the Population of the Russian Federation"; Grant "Financial Culture of Capital Youth" [16]) show the role and importance of increasing the level of financial literacy of Russians for making decisions in the economic sphere; the results of systematization of conceptual approaches to improving the efficiency of financial education of the country's residents are presented and a structural-functional model within which this can be implemented is described [17].

Despite a large number of published research results in the field of behavioral economics, in particular, the financial strategies of the population [18], there is still a demand for scientific research that reflects the relationship between the factors of human activity "within" the social and economic subspaces.

RESEARCH METHODOLOGY

The article uses scientific methods of analysis (content analysis, structural-logical analysis) and synthesis (systematization) of existing theoretical approaches. In the course of processing materials on the dynamics of the activity of Russians in the financial sector, based on the results of monitoring the behavior of population groups, an expert approach was used. The authors prefer the methodology of spatial analysis as a scientific methodology for studying the sustainability of the development of an integrated space based on the search for relationships and modeling potential interactions (spatial organization) of the structural components of various subspaces [19, p. 28, 29, 33, 34].

RESEARCH RESULTS

The social subspace, characterized as an area of social processes functionally interconnected by various forms of social relations and practices, determines a wide range of human activity. The economic subspace, which is part of the social one, involves: firstly, the use of limited economic values (resources) that differ in function and purpose; secondly, the orientation of a person to receive benefits/profits. The financial subspace as part of the economic one is associated with the manifestation of the activity of an individual/group of people in the use of money (monetary behavior), this is the stage on which the struggle is unfolding to meet the needs and improve the quality of life through the use of financial resources through the interaction of people with each other and social institutions of the financial sphere.

Financial behavior, according to the authors, is the most concentrated result of the active interaction of people – in terms of justifying choices and making complex decisions.

An individual, as well as each household and/or social group, live and interact simultaneously in three subspaces. For scientific analysis, they can be "separated" for the purpose of a detailed study (consideration) of the essence of the phenomena included in these three "volumes" and the factors influencing human behavior.

Fig. 1 shows the relationship of subspaces in the form of a pyramid "recumbent matryoshka doll". In each subspace, accents are highlighted: the social subspace includes relations that arise in the society in which the individual constantly lives; the economic subspace includes the resources

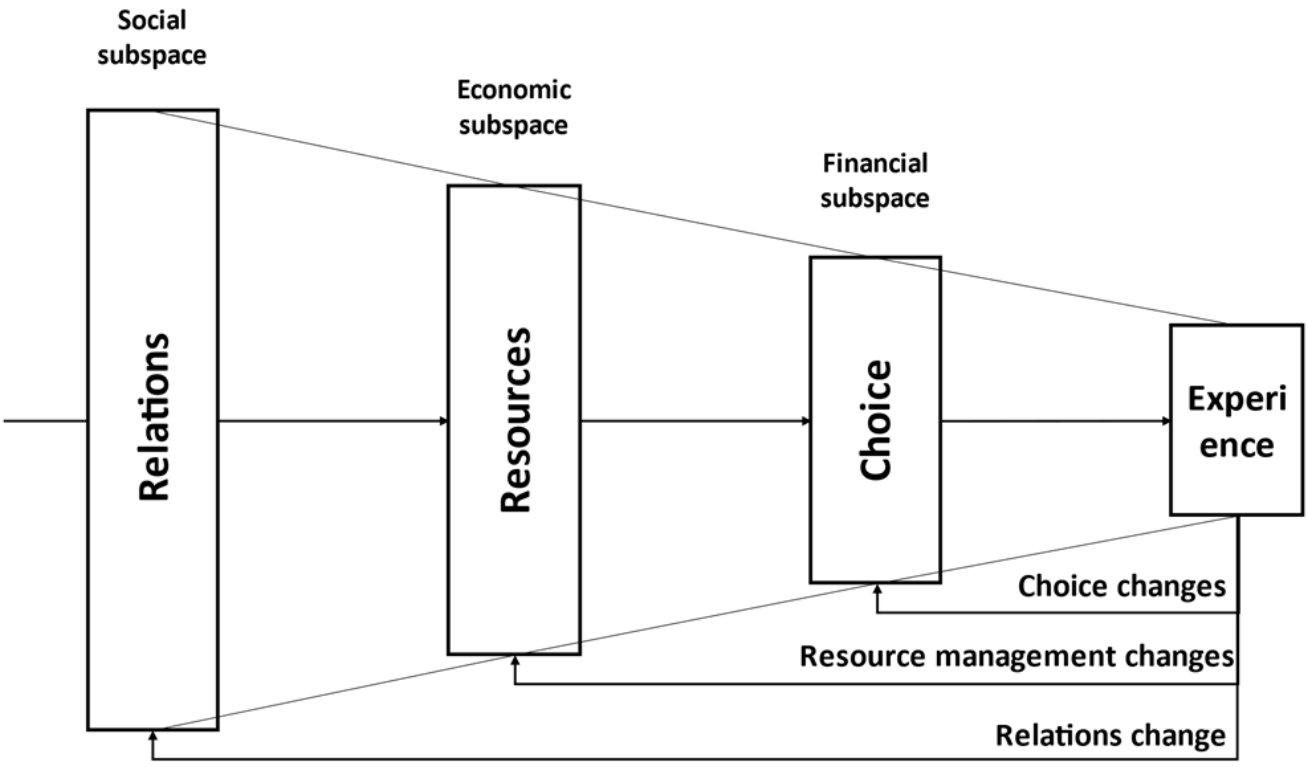


Fig 1. **Relationship of key elements in the hierarchy of three subspaces: A functional approach**
Source: developed by the authors.

available to the population are formed and then directed to the needs of consumption; the financial subspace — the conditions and motives for choosing a behavior strategy. Functioning in all three subspaces, an individual gains experience (positive and negative), which affects to some extent (including, depending on the mentality and cognitive abilities) a change in the choice of financial decisions, different disposal of resources, the transformation of relations.

Let's consider each of the subspaces and their interconnected elements (*Fig. 2*).

In the social subspace (*Fig. 2*), an important role is played by the degree of trust of a person (a group of people) in the state, its individual structures and institutions at all levels (federal, regional, municipal), and confidence that it is the state that guards the interests of the population as a consumer economic benefits and services (ensuring economic security and protection of economic rights, the availability of high-quality educational services and the possibility of successful employment and further realization of personal potential). The influence of the social environment is manifested in a person's choice of a certain strategy of behavior under the "pressure" of the reference group (family, friends, colleagues, etc.). An important element of the social subspace is the interaction under the "Man and man" principle, which implies a certain level of mutual trust between individuals. In particular, microblogging on social networks and the use of recommendations (advice from professional and non-professional bloggers and influencers) that teach various financial practices to people are becoming increasingly popular.

The economic subspace (*Fig. 2*) includes a set of resources: labor, capital, and property. At the same time, labor (employment) and its results (wages, entrepreneurial income) are the basis for the formation of monetary resources that an individual/social group uses to meet their needs.

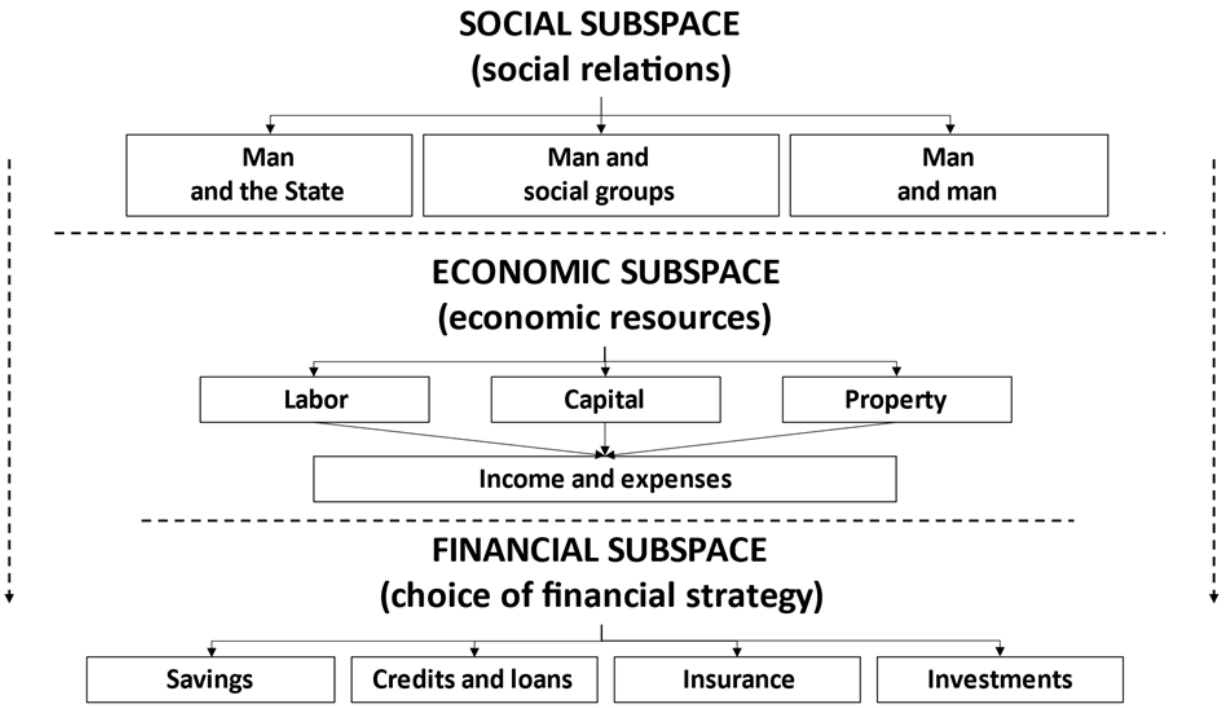


Fig 2. **Key categories of social, economic and financial subspaces**

Source: developed by the authors.

Capital is understood as cash income from savings (interest on deposits) and investments (dividends), as well as human capital as the ability to work, the possibility of professional growth, and realization of personal potential to replenish resources. The property of individuals (movable and immovable assets) is considered as a source of income from its lease. From the economic perspective, there are not only incomes but also expenses, their structure, as well as the possibility of transformation (changes in percentage terms) of this structure (item of expenditure).

As mentioned above, the social and economic subspaces have a financial subspace "inside" which includes the possibility of a person implementing certain financial strategies: savings [20], the lending [21], investment [22], insurance [23].

In each case, researchers have a specific task, so priorities often (and inevitably) change places: how to consider and interpret certain types of financial behavior. Thus, savings are often inseparable from investments [investments in conservative securities or instruments are actually savings; buying real estate or jewelry is considered an investment, but it is also a way to save (keep, accumulate) depreciating money]. Consumer behavior is the purchase of goods and services, but banks also offer services (financial) for the population. Therefore, depending on the task, scientists separately study food consumption (as consumer behavior in the broader *economic* sense, i.e. basic daily expenses) and consumption of financial (banking) services.

In this regard, there is an obvious need to study the totality and priority of factors influencing the behavior of people in their economic life. Depending on which of the three subspaces they directly occupy, they can be grouped (*Fig. 3*). A set of systematized influence factors was used in practice when performing a monitoring study at the ISESP FCTAS RAS in 2021 [13], as well as when developing tools for a sociological survey of young people of Moscow as part of a joint project of two academic institutions (2021–2022) — the Institute of Socio-Economic

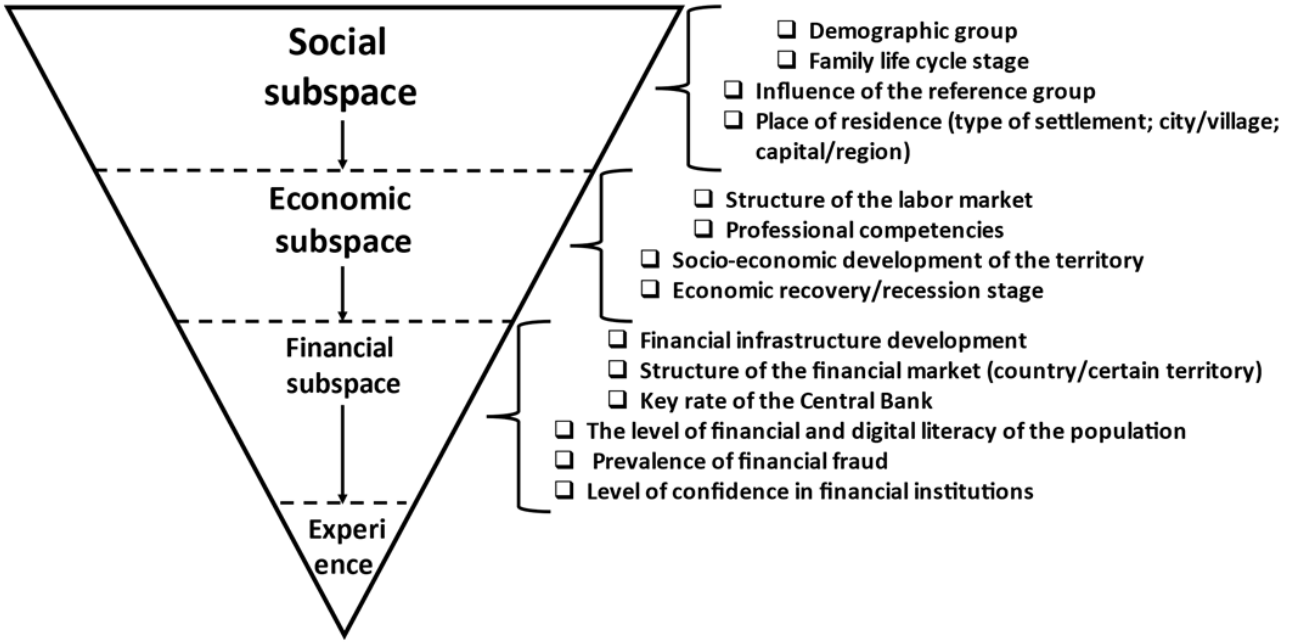


Fig 3. **Factors of influence on the behavior of a person who is in the "pyramid" of subspaces — social, economic and financial**

Source: developed by the authors.

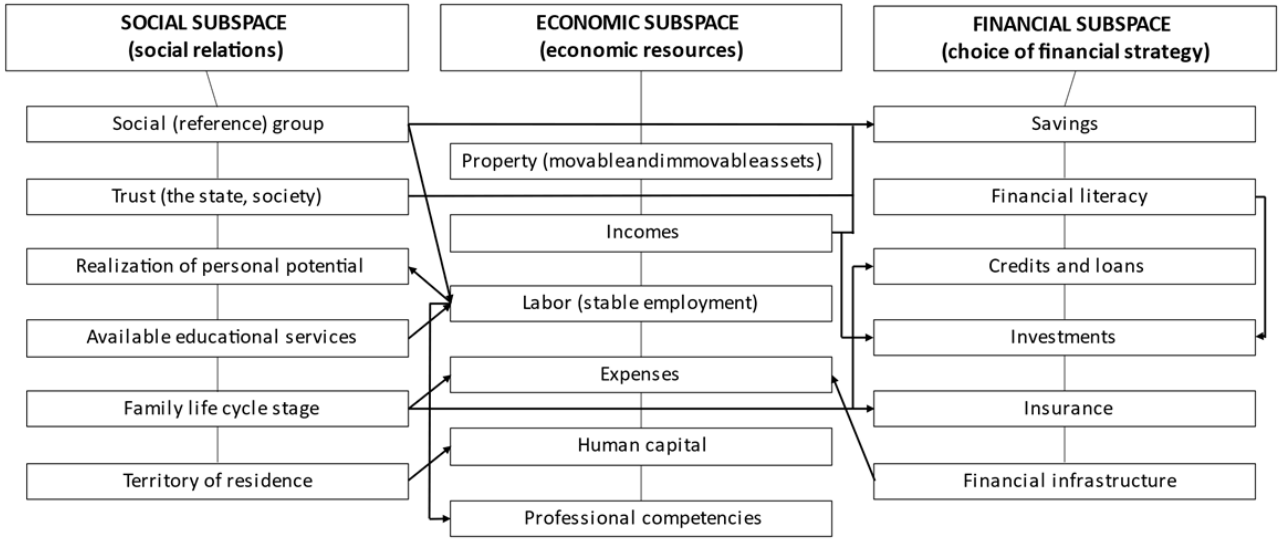


Fig. 4. **Basic theoretical concepts in the study of people's financial behavior and their relationship**

Source: developed by the authors.

Strengths	Weaknesses
Social networks are an established platform for obtaining information, which gives us the opportunity to assess the preferences of the population, in particular, young people, who are a strategic resource of our country, on which all financial market participants rely, these are future active users of the financial market	Lack of legislation regulating the activities of financial advisors in social networks, namely: 1) the law does not fix who has the right to provide such educational services, there are no requirements for consultants; 2) what content can be promoted to users and under what conditions, according to what parameters the customer should choose a financial consultant, and how to distinguish a fake expert from a knowledgeable, experienced consultant? Therefore, the protection of the interests and rights of customers using the services of financial consultants comes to the fore; 3) the low level of literacy as well as culture. Thousands of people believe in rituals to attract wealth; 4) inability to navigate in a huge flow of information
Threats	**Opportunities**
1. A large stage is open for scammers, fake consultants, magicians, and other charlatans. 2. A huge uncontrolled information flow, in which it is extremely difficult for an average person with a low level of financial literacy to identify conscientious, competent financial teachers. The activity of unscrupulous consultants casts a shadow over the entire field of financial consulting. 3. The scale of potential «victims». It is not necessary that the victim literally lose money. It is important to focus on something else, rather, what will the victim get for this money and what damage can this cause the victim? 1) loss of money; 2) loss of time; 3) disorientation; 4) work with incorrect information and, as a result, failure to achieve the goal; 5) an incorrect understanding of the meaning of the stock market (as one of the most sought-after topics), which leads to the fact that novice investors seek short-term profit based on short-term trading, which, as a rule, leads to a loss of capital (rather than weighted long-term investments), which leads to demotivation and loss of confidence in the stock market	1. Opportunity to refocus financial education, rethink approaches and methodologies, and analyze audience needs 2. Improvement of the financial market by identifying strong, high-quality players and screening out unscrupulous participants in terms of financial education for the population 3. Through social networks, the process of increasing financial literacy will be faster due to the large coverage

Fig 5. **SWOT analysis of the impact of social media on the financial behavior of young people**

Source: Alikperova N.V., Vinogradova K.V. The influence of social media on the formation of financially literate behavior of youth.

Studies of Population and the Institute of Psychology.

The population carries out its activities in the social subspace depending on the influence of the environment around: family (and the stage of its development), colleagues, neighbors-residents of the metropolis/city/village/district, etc.; and also depending on belonging to a certain gender and age (demographic) group. Within the framework of the economic subspace, a huge role is played by the possibility of displaying personal, primarily professional, characteristics of a person in the labor market. This can be determined by the individual's professional competencies and their successful application in the territory of residence (however, remote employment and exemption from local residence are becoming increasingly important). The formation of the monetary resources of the population is facilitated (or hindered) by the level of development and the stage (recovery/recession) of the region's economy (place of residence).

In the financial subspace, for the implementation of financial behavior strategies by a person, the following factors are equally important:

- availability of banking (and, in general, financial) infrastructure;
- variety of banking services (opportunity to choose the type of deposit, the form of loan, investment, and insurance instruments);
- dynamics of changes in the monetary policy of the Central Bank;
- protecting a person as a consumer of financial services from fraud;
- the possibility of increasing the level of financial literacy as a result of organized activities in the country and regions for the financial education of residents.

The level of financial [24] and digital literacy is, along with trust in banking institutions and all participants in the financial market, an essential factor in gaining experience in implementing people's monetary strategies. Both positive and negative experience allows changing the strategy of financial behavior in order to minimize losses and risks, and most importantly, not be caught in the financial scammers' net.

Summing up, it is possible to single out and "link" the main theoretical concepts used in the study of a person's financial behavior (*Fig. 4*). All concepts used in the three subspaces under consideration are directly or indirectly interconnected.

When studying the basics of the credit and investment behavior of the population, it is important to take into account and analyze the level of financial literacy of people, and the consideration of savings strategies in financial behavior involves tracking changes in the level of confidence in the state of the banking system and the level of income of citizens.

CONCLUSIONS

Personal financial behavior (family/household, social group), carried out simultaneously in three subspaces — social, economic, and financial — is determined by a complex of factors that influence (motivate and/or demotivate) the population in daily activities. The categories used in financial strategy design conditionally belong to three subspaces: social, economic, and financial.

The integral space for making financial decisions is a hierarchical model of a "recumbent matryoshka doll", where the financial subspace is included in the economic one, and the social one, in turn, includes the first two. This theory is confirmed, in particular, by the results of the development of connections between the labor behavior of certain categories of people and the possibilities of building successful (effective) financial strategies [25]. The problems of distribution of resources (incomes of the population) between consumption, savings, and investments are reflected in publications related to the dynamics of the index

of expected changes in the economic situation [26].

Studies based on the methodology proposed by the authors[1] are of practical importance for determining the value system of young people as the most important demographic stratum for the future economic development of the country. For example, identifying the characteristics of the financial behavior of a particular social group and factors influencing the formation of a strategy was conducted in the context of the role of social media in transforming the financial literacy of young people.

The data obtained by the authors made it possible to compile a SWOT analysis of the impact of social media on the financial behavior of young people (*Fig. 5*), which characterizes the main problems and potential opportunities in the implementation of financial activity.

The findings of the study can be applied in the development and implementation of state policy, which allows creating conditions for the formation of financially literate behavior of young people and bringing it to a more effective level, in particular, involving young people in real market economic processes to effectively use their savings and subsequent investments. For the population using various financial strategies, the conclusions of the authors are of practical importance for increasing economic well-being and expanding opportunities during periods of adaptation to unstable conditions by building more flexible strategies.

Thus, the authors show that not only an interdisciplinary approach to the study of people's financial behavior, familiar to scientists, is in demand, but also a scientific approach that takes into account the factors of mutual influence on the formation of people's financial strategies in three spatial dimensions.

[1] The study (2021) was conducted by the survey method (total youth aged 18–35 years (88%), the rest of the participants — aged 14–17 years (12%) — 405 respondents, including respondents from Moscow — 384) with limit error sampling at the level of 5%; 2⁄3 are female respondents. Every second participant of the survey was employed (studies and works — 23% of respondents, and only works — 29%). At the same time, every 3 out of 5 respondents who have a job indicated that they work in areas indirectly or directly related to finance.

REFERENCES

1. Belekhova G.V., Gordievskaya A.N. Population's financial behavior: Demographic features. *Problemy razvitiya territorii = Problems of Territory's Development.* 2018;(1):133–150. (In Russ.).
2. Makar S.V., Yarasheva A.V., Alikperova N.V. Financial infrastructure: Public access points. *Ekonomika. Nalogi. Pravo = Economics, Taxes & Law.* 2021;14(3):56–66. (In Russ.). DOI: 10.26794/1999–849X-2021–14–3–56–66
3. Aleksandrova O.A., Alikperova N.V., Markov D.I., Nenakhova Yu.S. Classification of factors of influence on the formation of financial and economic culture of Russians. *Narodonaselenie = Population.* 2021;24(3):18–31. (In Russ.). DOI: 10.19181/population.2021.24.3.2
4. Kahneman D. Thinking, fast and slow. New York: Farrar, Straus & Giroux; 2013. 499 p. (Russ. ed.: Kahneman D. Dumai medlenno... reshai bystro. Moscow: AST; 2014. 653 p.).
5. Akerlof G.A., Shiller R.J. Phishing for phools: The economics of manipulation and deception. Princeton, NJ, Woodstock: Princeton University Press; 2015. 288 p. (Russ. ed.: Akerlof G., Shiller R. Okhota na prostaka. Ekonomika manipulyatsii i obmana. Moscow: Mann, Ivanov and Ferber; 2017. 320 p.).
6. Akerlof G.A., Shiller R.J. Animal spirits: How human psychology drives the economy, and why it matters for global capitalism. Princeton, NJ: Princeton University Press; 2009. 248 p. (Russ. ed.: Akerlof G., Shiller R. Spiritus Animalis, ili kak chelovecheskaya psikhologiya upravlyaet ekonomikoi i pochemu eto vazhno dlya mirovogo kapitalizma. Moscow: United Press; 2010. 273 p.).
7. Thaler R. Misbehaving: The making of behavioral economics. New York: W.W. Norton & Co., Inc.; 2016. 432 p. (Russ. ed.: Thaler R. Novaya povedencheskaya ekonomika. Pochemu lyudi narushayut pravila traditsionnoi ekonomiki i kak na etom zarabotat'. Moscow: Eksmo; 2018. 384 p.).

8. Thaler R.H., Sunstein C.R. Nudge: Improving decisions about health, wealth, and happiness. New Haven, CT: Yale University Press; 2008. 304 p. (Russ. ed.: Thaler R., Sunstein C. Nudge. Arkhitektura vybora. Kak uluchshit' nashi resheniya o zdorov'e, blagosostoyanii i schast'e. 2nd ed. Moscow: Mann, Ivanov and Ferber; 2018. 240 p.).
9. Grebenkin A.V. The concept of "nudging" in behavioral economics: Critical notes and new ideas. *Zhurnal ekonomicheskoi teorii = Russian Journal of the Economic Theory*. 2018;15(3):377–384. (In Russ.). DOI: 10.31063/2073–6517/2018.15–3.2
10. Ariely D. Predictably irrational: The hidden forces that shape our decisions. New York, London: Harper Perennial; 2010. 380 p. (Russ. ed.: Ariely D. Predskazuemaya irratsional'nost': skrytye sily, opredelyayushchie nashi resheniya. Moscow: Mann, Ivanov and Ferber; 2010. 296 p.).
11. Makar S.V., Yarasheva A.V. Economic expectations of the population: What has changed during the pandemic? In: Income, expenses and savings of the population of Russia: Trends and prospects. Proc. 6th Int. sci.-pract. conf. (Moscow, Dec. 08, 2020). Moscow: FCTAS RAS; 2021:166–172. (In Russ.). DOI: 10.19181/konf.978–5–4465–3137–0.2021.26
12. Kirsanov I.A., Parfyonova L.B. Behavioral economics and behavioral finance as a necessary stage in the development of the homo economicus model. *Vestnik Altaiskoi akademii ekonomiki i prava = Journal of Altai Academy of Economics and Law.* 2020;(9–1):48–57. (In Russ.). DOI: 10.17513/vaael.1302
13. Aleksandrova O.A., Alikperova N.V., Borkovskaya E.I. et al. Financial behavior of the population (monitoring study). Moscow: FCTAS RAS; 2021. 113 p. (In Russ.). DOI: 10.19181/monogr.978–5–89697–359–1.2021
14. Yarasheva A.V., Alikperova N.V. Gender differences in financial behavior of population. *Narodonaselenie = Population.* 2020;23(2):51–60. (In Russ.). DOI: 10.19181/population.2020.23.2.5
15. Income, expenses and savings of the population of Russia: Trends and prospects. Proc. 6th Int. sci.-pract. conf. (Moscow, Dec. 08, 2020). Moscow: FCTAS RAS; 2021. 294 p. (In Russ.). DOI: 10.19181/konf.978–5–4465–3137–0.2021
16. Alikperova N.V., Vinogradova K.V. The influence of social media on the formation of financially literate behavior of youth. *Narodonaselenie = Population*. 2021;24(1):33–43. (In Russ.). DOI: 10.19181/population.2021.24.1.4
17. Aleksandrova O.A., Alikperova N.V., Vinogradova K.V., Nenakhova Yu.S. Conceptual approaches to creating the preconditions for effective financial education of the Russian population. *Economic and Social Changes: Facts, Trends, Forecast*. 2020;13(4):169–185. DOI: 10.15838/esc.2020.4.70.10 (In Russ.: *Ekonomicheskie i sotsial'nye peremeny: fakty, tendentsii, prognoz.* 2020;13(4):169–185. DOI: 10.15838/esc.2020.4.70.10).
18. Bakhtina O. Yu. The genesis of scientific approaches to the study of financial behavior. *Ekonomika i predprinimatel'stvo = Journal of Economy and Entrepreneurship.* 2019;(5):260–265. (In Russ.).
19. Makar S.V. Application of the methodology of spatial analysis to the study of the forest potential of Russia. Moscow: Economika; 2012. 367 p. (In Russ.).
20. Gamukin V.V. Savings and loan behaviour of the population in the Russian regions. *Ekonomika regiona = Economy of Regions.* 2020;16(3):1003–1017. (In Russ.). DOI: 10.17059/ekon.reg.2020–3–25
21. Yarasheva A.V., Makar S.V., Reshetnikov S.B. The credit strategies of the Russians as the reflection of the model of financial Behaviour. *Finansy: teoriya i praktika = Finance: Theory and Practice*. 2017;21(6):138–153. (In Russ.). DOI: 10.26794/2587–5671–2017–21–6–138–153
22. Baganov V. Yu. Stimulating long-term investment of the population in the stock market. *Ekonomika, predprinimatel'stvo i pravo = Journal of Economics, Entrepreneurship and Law.* 2020;10(5):1547–1564. (In Russ.). DOI: 10.18334/epp.10.5.110171
23. Sinyavskaya T.G., Tregubova A.A. Who chooses different types of insurance: Econometric analysis. *Intellekt. Innovatsii. Investitsii = Intellect. Innovation. Investments.* 2019;(8):58–68. (In Russ.). DOI: 10.25198/2077–7175–2019–8–58

24. Kuznetsov O.V., Ivanov A.V., Vorovskii N.V., Shevalkin I.S. Institutionalization of financial literacy of the adult population of Russia. *Finance: Theory and Practice*. 2020;24(1):34–45. DOI: 10.26794/2587–5671–2020–24–1–34–45
25. Alikperova N.V., Yarasheva A.V., Vinogradova K.V. Motivating young people's labor behavior as an opportunity for implementing financial strategies. *Economic and Social Changes: Facts, Trends, Forecast*. 2019;12(1):226–240. DOI: 10.15838/esc.2019.1.61.14 (In Russ.: *Ekonomicheskie i sotsial'nye peremeny: fakty, tendentsii, prognoz*. 2019;12(1):226–240. DOI: 10.15838/esc.2019.1.61.14)
26. Maslennikov V.V., Larionov A.V. Impact of the behavioral cycle on cash flow formation. *Finance: Theory and Practice*. 2020;24(5):100–111. DOI: 10.26794/2587–5671–2020–24–5–100–111

ABOUT THE AUTHORS

Svetlana V. Makar — Dr. Sci. (Econ.), Assoc. Prof., Leading Researcher, Institute of Regional Economy and Interbudgetary Relations, Financial University, Moscow, Russia; Prof., Department of Physical and Socio-Economic Geography, National Research Ogarev Mordovia State University, Saransk, Russia
http://orcid.org/0000-0002-1681-8814
Corresponding author
svetwn@mail.ru

Aziza V. Yarasheva — Dr. Sci. (Econ.), Prof., Head of laboratory, Institute of Socio-Economic Studies of Population, Federal Center of Theoretical and Applied Sociology RAS, Moscow, Russia
http://orcid.org/0000-0002-6041-7700
baktriana@rambler.ru

Dmitrii I. Markov — Deputy Head of the Educational and Scientific Sociological Laboratory of the Department of Sociology, Faculty of Social Sciences and Mass Communications, Financial University, Moscow, Russia
http://orcid.org/0000-0001-9380-0506
dimarkov@fa.ru

Authors' declared contribution:
S. V. Makar — research concept and design, scientific editing of the text, responsibility for the integrity of the article.
A. V. Yarasheva — research concept, material collection, text design and creation, scientific editing.
D. I. Markov — assembly and graphic processing of the material of parts of the article.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 17.01.2022; revised on 31.01.2022 and accepted for publication on 20.02.2022.
The authors read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-169-185
UDC 349;336.1;369.06(045)
JEL K32, G22, I18

Organizational, Legal and Financial Aspects of Digitalization and Implementation of Artificial Intelligence Technologies in Healthcare

M.A. Lapina
Financial University, Moscow, Russia

ABSTRACT

The paper aims to substantiate the main development directions of legal regulation of artificial intelligence in healthcare. The main hypothesis of the study is the assumption that artificial intelligence should not be a subject of law. The author formulates the postulates necessary for the introduction of modern technologies in the context of the digitalization of medicine. General and special scientific methods are used: the dialectical method of cognition of reality, synthesis and deduction. The comparative and formal legal method of scientific cognition made it possible to analyze the laws and other documents of a number of states in the field of digitalization and the mechanism for financing the provision of medical services and medical care. The article analyzes the directions proposed by scientists and practitioners with the participation of the largest IT companies to improve the provision of medical care and medical services and optimize healthcare management. The author draws attention to financial mechanisms to stimulate the introduction of digital technologies in the healthcare system, directly to the provision of medical care. Structuring the main directions of applicability of digital technologies in healthcare allowed us to formulate proposals for improving their legal support. The analysis of foreign and domestic legislation has revealed the importance of using such a financial and legal mechanism as health insurance. Based on the results of the study, the author makes a conclusion about the need for a systematic approach to digitalization in healthcare and proposes an institutional and legal model for the development of patient-centered medicine based on artificial intelligence technologies.

Keywords: healthcare; artificial intelligence; health insurance; medical ethics; financing; digital technologies

For citation: **Lapina M.A. Organizational, legal and financial aspects of digitalization and implementation of artificial intelligence technologies in healthcare. *Finance: Theory and Practice.* 2022;26(3):169-185. DOI: 10.26794/2587-5671-2022-26-3-169-185**

INTRODUCTION

Russia has adopted a policy of developing the information society and increasing the effectiveness of State authorities in all spheres of life by integrating certain achievements in the development of digital technologies and artificial intelligence (hereinafter — AI).

Among domestic researchers the most systemically revealed the legal basis of artificial intelligence technologies P.M. Morkhat in the monograph "Artificial intelligence: a legal view" in 2017 [1]. One of the first attempts to consider the issues of legal personality of cyberphysical systems, the basis, content and legal prospects of the concept of "electronic person" was made by Professor O.A. Yastrebov [2]. The team of authors of Financial University is engaged in research of legal regulation of AI for the third year. After analyzing many approaches to creating a legal and regulatory framework, the author's team concluded that it is important to consider the risks and possible negative consequences of using robotic systems and AI [3].

It should be noted that the topic of legal regulation of AI, including the narrow application of this technology in medicine, has been repeatedly raised in foreign scientific literature. Both the broad theoretical issues of the application of AI in various health applications and the more practical and specialized developments, which can be used in the applied activities of medical organizations and in the management of public health bodies.

The problem of integration of modern IT-technologies, as well as technologies using artificial intelligence in the field of health care, is one of the most fascinating and at the same time difficult questions of actual legal theory. The initial postulates for further legal research will be as follows:

- interaction with intelligent systems is possible not as subjects, but as mechanisms, "sources of increased danger";
- final decision and responsibility will be for the person, for the subjects of law (natural or legal persons).

We understand that for artificial intellectual systems such concepts as "human factor", "subjective experiences", "intention", "negligence", "presumption of innocence" and others, practically do not take into account when making decisions. Accordingly, responsibility cannot be shifted from the subject of law (in traditional legal doctrine) to artificial intelligence. The use of artificial intelligence should be embedded in a special legal regime. The institutions of law related to their use should be provided for and regulated in the basic legal acts, are in demand by various branches of law and legislation. Earlier in our research, we formulated conceptual conclusions that should form the basis of further development [3–7]. The issue of health law is particularly acute in the context of the digitalization of health care, where only part of the issue has been developed regarding the safety of medical products using AI technology.

On the other hand, the question arises: how much automated system is "responsible" for decision-making. The subject of administration should be responsible for its actions as its own, because the legal consequences are real, and not "virtually". Incorrect diagnosis of the patient on the basis of the mechanism "telemedicine" should entail the same legal consequences as for actions in the normal environment. Otherwise, we get "fiction of a legal entity" and "fiction of a natural person", "fiction of an official, specialist" etc. Where there is no or incomplete liability in such cases, the very basis of the right, the criteria of permissible and impermissible. "Foreign researchers come to the conclusion that there is a sign of equality between the construction of the legal entity and the autonomous system: the robot can operate the company without any human substrate. This problem goes beyond the study of the possibility of applying the legal entity fiction to such systems" [8, p. 357]. Thus, the actions and decisions of any automated system are equivalent to the activities of the legal entity that administers

it. The publication also deals with the management rights system of the company. The author concludes that it is necessary to improve the corporate legislation in terms of developing a special legal regime to determine the nature of the rights and obligations of the legal entity transferred through the automated system. But there is also a contrary opinion, for example, E. A. Sukhanov believes that the basic definitions in the Civil Code of the Russian Federation are quite applicable for the digital environment and there is no need to allocate digital rights and assets in a special group. This could disrupt the regulation of civil rights objects. Both real and virtual objects of civil rights should have the same legal regulation [9, p. 296]. This approach also deserves attention. This requires extensive legal regulation, enshrined not only at the level of civil legislation, but also specialized medical legislation to detail digital rights, digital objects, etc.

A separate problem is the transmission of data constituting a secret, including medical confidentiality. Who is responsible for the transmission of data by an automated system if the person has not given consent to the transmission and processing of the data? This also raises the question of the reliability of the verification system for electronic signature certificates. "Cybersecurity analysts are doing everything possible — quickly studying new threats, releasing tools to protect against them, which are quickly spread. But the fact remains that between the first attack and the advent of the latest countermeasures, users remain defenseless" [10].

Problems of intellectual property protection are connected with the circulation of objects of intellectual rights on virtual platforms. "The complexity of regulating Internet processes affecting intellectual property rights is primarily due to the extraterritoriality of the Internet environment, which sometimes makes it difficult to define the boundaries of national legislation" [11].

The transition to digital governance should not only remove barriers to entry, ensure equal rights of access to information, but also control the process of using the data transferred to the system by attackers. There are opportunities to transmit distorted information, to anticipate threats and risks of cyberattacks, to ensure the protection of information. In the automatic data interchange system "necessary to legislatively work out the issues of validation of the legal validity and significance of the transmitted digital information. To begin with, a special legal regime must be established and elaborated to take certain actions in the digital economy" [12, p. 114]. These include decentralized registry management, certification of rights and automatic data exchange.

In the Russian Federation, according to V. S. Cherenkova, "for full protection of personal medical data, they should be qualified by the court as medical confidentiality" [13, p. 134].

In the European Union, a broader interpretation of medical personal data, as pointed out in particular by the European Court of Human Rights, which decided, "that the expression "health data" should be interpreted broadly to include information relating to all aspects of both physical and mental health. Analyzing Deal T-343/13 — CN v. Parliament from 3 December 2015, E. B. Luparev notes that this concept cannot be expanded to include provisions that do not lead to the disclosure of any data related to a person's health or medical indications" [14]. The author of the article draws attention to the role of the operator of the information system through which information containing medical confidentiality is exchanged. If the medical service is rendered poor-quality due to the fault of the operator of the information system, he is a subject of civil liability [14].

Worldwide, as Professor I. V. Ponkin argued in 2018, is the problem "virtually no regulatory legal regulation and regulatory technical regulation of the bases, conditions

and features of development, start-up, operation and activity, integration into other systems and monitoring of artificial intelligence technologies" [15, p. 91].

Over the past two years, this problem has been resolved intensively, and numerous regulations have been adopted, which we will elaborate on in this publication.

The main hypothesis of our research results is that AI should not end up being a legal subject in all areas of public relations, including health. It is advisable to classify all actions of automated systems as high-risk facilities for the operation of which subjects of their administration are responsible. However, a special legal regime should be developed that defined the legal consequences of the actions and decisions of intelligent machines, and provided for legal liability. The object of the right should never be a subject of law and should enjoy the rights and protection of such a subject. Neither the presumption of innocence nor other human rights gains, nor the guarantees of rights established by law, must cease to exist. The results of the use of automatic means cannot be initially uncontroversial and accepted without evidence to bring a person to justice, etc. Such practices, as well as other negative decisions on the application and transfer of responsibility to intellectual systems from the standpoint of human rights, violate existing standards and norms of international law.

WHAT MODERN MEDICINE REPRESENTS BASED ON THE USE OF MODERN TECHNOLOGIES, INCLUDING AI?

On the basis of the introduction of information technologies, the formation and development of patient-oriented medicine is being accelerated throughout the world, including in Russia. In order to implement the Strategy for the Development of the Information Society in the Russian Federation for 2017–2030, Strategy for the development of medical science in the Russian Federation for the period up to 2025, a priority project "Improvement of the processes of organization of medical care on the basis of introduction of information technologies", approved by the protocol of the Council under the President of the Russian Federation on strategic development and priority projects from 25 November 2016 No. 9; and Order of the Government of the Russian Federation from 28 July 2017 No. 1632 "Digital Economy of the Russian Federation" (no longer valid); and Decree of the President of the Russian Federation from 21 July 2020 No. 474 "National development objectives of the Russian Federation for the period up to 2030", where one of the national goals is "Digital Transformation", and other documents provides for the development of **patient-oriented medicine** using modern innovative technologies, including information technologies.

Development of legal regulation of patient-oriented medicine and transformation on its basis of the health system is impossible without legal enforcement of artificial intelligence technologies in medicine.

In the Address of the President of the Russian Federation to the Federal Assembly of the Russian Federation dated from March 2018, V. V. Putin noted that "the role, position of the state in the modern world determine not only and not so much natural resources, production capacity, (...) and above all people, conditions for development, self-realization, creativity of each person. Therefore, the basis of everything is the saving of the people of Russia and the well-being of our citizens. This is where we need to make a decisive breakthrough", "in the shortest possible time we need to create an advanced legislative framework, remove all barriers to the development and wide application of robotics, artificial intelligence, unmanned transport, e-commerce, big data processing technologies. And such a regulatory framework should be constantly updated,

based on a flexible approach to each field and technology".[1]

In accordance with the implementation of the President's Address to the Federal Assembly from 20 February 2019, large-scale national level programs in the field of AI have been launched.[2] For example, a national AI strategy was developed in 2019[3] and in the last year or a year and a half a number of other normative legal acts significant for the development of AI technologies in the health system.

According to experts, a modern health system and a traditional approach to treatment cannot be considered effective for a number of reasons, and not only for economic reasons. There is evidence that about half of patients do not have the desired result [16]. Recourse to **patient-oriented** rather than **client-oriented** medicine will be in accordance with the principles of humanism and the values proclaimed by the Constitution of the Russian Federation, human rights and freedoms, foremost among which are the right to life and health.[4] The application of artificial intelligence technologies in patient-oriented medicine is based on interconnection "information technology, clinical therapy, marketing approach, legal support aimed at both improving health and meeting the needs of patients" [16].

Problems of health, its protection and guarantees by the State are the focus of attention of scientists. The search for new socio-innovative approaches based on the principles of evidence-based medicine, to solving the problems of medical care from the perspective of intersectoral interaction between the State, business and civil society becomes the most important national task.

An important aspect of the creation and development of patient-oriented medical information systems is interaction with the professional community. The scientific literature pays some attention to the digitalization of health care. So, individual authors specializing in medicine, reveal the features of patient-oriented medicine [17–22].

Patient-oriented medicine is a medical approach that focuses on the needs of the patient. This requires refocusing and reconfiguring health information systems in accordance with the principles of personalized medicine. Treatment analysis can inform the implementation of information technology by revealing roles, actions and other important data, such as information transfer and situational awareness requirements. Such data can be obtained through electronic workflows, including clinical observations [23] and database analysis [24, 25]. Nevertheless, research on health work processes tends to focus on medical staff — doctors or nurses. These clinician-oriented models often describe a series of discrete actions of a particular type of doctor and the amount of time spent on each type of activity. Less common, but no less important, are patient-oriented studies that describe the sequence of actions of all involved health-care personnel to assist a particular patient [26].

K. Blakemore was working on the assessment the role of the State through the system of medical institutions in the field of client-oriented medicine [27].

The organization of client-oriented medicine is within the scope of attention in the works of L.M. Mukharyamova, I.B. Kuznetsova, G. G. Vafina [28], L. A. Goncharov [29] et al. Authors S. P. Troshin, I. N. Legkova, K.V. Gavrilenko reveal the features of patient-

[1] In the Address of the President of the Russian Federation to the Federal Assembly from 01 March 2018. Rossiyskaya Gazeta. No. 46. 02 March 2018.

[2] In the Address of the President of the Russian Federation to the Federal Assembly from 20 February 2019. Rossiyskaya Gazeta. No. 38. 21 February 2019.

[3] Decree of the President of the Russian Federation from 10 October 2019 No. 490 "On the development of artificial intelligence in the Russian Federation" (with the «National Strategy for the Development of Artificial Intelligence until 2030»). Legislative assembly of the Russian Federation, 14 October 2019, No. 41, art. 5700.

[4] The Constitution of the Russian Federation (adopted by popular vote on 12 December 1993 with the amendments approved by the All-Russian vote on 01 July 2020). Official Internet Legal Information Portal. URL: http://www.pravo.gov.ru, 04 July 2020.

oriented medicine on the basis of the use of modern information technologies [30].

AI technologies are used intensively in health care. Doctors already have medical robots of various types (for example, Da Vinci and Hospice surgical systems). In the field of health, AI technologies are used in disease diagnosis and forecasting, data collection, identification of patients at high risk of disease, drug development. The most promising use of AI in telemedicine and diagnostics [19]. According to some researchers, e-health has become one of the main "drivers" of the entire industry [31].

The improvement of medical care and optimization of health care management are engaged not only directly health institutions, but also many major IT-companies. So, IBM® Watson Health™ proposes three stages in the evolution of the use of information technology in health care:

1) digitization of routine processes in the health and medical care system (for example, MRI, CT scan, automatic management of payment for services, digitalization of data storage, access to their turnover);

2) rejection of old models and applying such breakthrough innovations as artificial intelligence, mobile technology, analytics and cloud[5];

3) transformation of the health-care system, which will integrate and integrate the digital functions and processes of health-care institutions and other organizations in the health-care system.

By implementing the stages of evolution formulated by the company IBM® Watson Health™, healthcare from a technological point of view will turn "from scattered fragments into an integrated ecosystem, allowing the medical professionals to successfully solve problems on a larger scale, keeping the focus on the patient, and provide value-oriented medical care".[6]

The necessity of introduction of AI in the health system is stated in the Strategy of development of the information society in the Russian Federation for 2017–2030,[7] where AI is considered one of the main directions of development of Russian information and communication technologies, along with convergence of communication networks, biotechnologies and cloud computing. According to S. G. Vasin, it is necessary to implement the system of management at the state level using AI [32]. As a confirmation of this thesis, the author cites examples from the National Artificial Intelligence Strategic Plan, from the UK Digital Economy Strategy. These documents give considerable attention to the application of AI in the field of social relations, and also raise issues of public investment in

[5] 1. The cloud provides flexibility and scalability to run and manage a set of analytical capabilities that support current and next-generation applications, all in a secure environment. Ultimately, the cloud analytics platform can share data to potentially help physicians, patients, caregivers and clinicians make timely and effective decisions about patients' treatment. 2. AI has already been used in health care to construct disease-specific progression models and to analyses genetic information to determine treatment effectiveness. AI can also be used for the design and development of clinical trials, previously labor-intensive and manual process. Already now, AI-automated trial matching can integrate data from electronic medical records, medical literature and eligibility criteria from legislators and learn to interpret test requirements based on patient cases. 3. Internet of Things. Using this capability requires a robust and integrated approach that allows multiple micro-services and devices to operate on the same analytical platform. Reduced sensor costs and growth of protected cloud platforms can have a profound impact on the healthcare industry. 4. Blockchain may offer a solution for safe access to health data. The data added to the blockchain can be transmitted almost in real time to a group of authorized persons and/or institutions. Each event or transaction is timed and becomes part of an object's immutable record, be it a drug record or a patient record. The transparency provided by blockchains moves data from individual, isolated ownership to a shared, secure record for permitted stakeholders. This public record becomes the only source of truth for the patient's history (or thing), freely carried by the patient in the form of a digital record, regardless of the location or health system. 5. Quantum will enable quantum computing to identify patterns and obtain information in artificial intelligence systems so complex that the world simply lacks classical computer resources to model them, including drug research and development.

[6] Technologies open the way for health transformation. URL: https://www.ibm.com/ru-ru/watson-health/learn/healthcare-transformation (accessed on 17 March 2021).

[7] Decree of the President of the Russian Federation from 09 May 2017 No. 203 "On the Strategy for the Development of the Information Society in the Russian Federation for 2017–2030". Legislative assembly of the Russian Federation. 2017. No. 20. Art. 2901.

these spheres of public administration. As the author correctly points out, foreign countries have been paying attention to this issue for quite a long time: as an example, the comprehensive report of the National Council for Science and Technology of the USA can be cited "Preparing For the Future of Artificial Intelligence. Executive Office of the President National Science and Technology Council Committee on Technology" for 2016.[8] A significant part of the report is devoted to the need for public investment in the development of AI-solutions in the field of social relations, including health, which should lead to an increase in productivity, reducing working hours, increasing wages, and ultimately the prosperity of American companies and workers, as well as continuing the leadership of the American nation in the creation and use of AI.[9]

When analyzing scientific publications on the current state of digitalization of medicine, business process researchers Sascha Kraus, Francesco Schiavone, Anna Pluzhnikova, Anna Chiara Invernizzi, five main areas of focus in health management: development and introduction of effective medical products using AI; patient-centred approach; organizational factors and improvement of health management; personnel policy; socio-economic aspects [33]. In particular, the authors recommend scientists to address in the future the reliability of solutions using AI technologies in the field of health care.

The direct applicability of information technologies in healthcare was researched by scientists I. C. Marques and J. J. Ferreira. After reviewing the scientific literature on digitalization in the field of health, they propose to highlight seven main directions: 1) integrated information technology management in health care; 2) medical images; 3) electronic health records; 4) information technology and portable devices in health care; 5) access to e-health; 6) telemedicine; 7) medical confidentiality [34].

The collection, processing and circulation of big data is an essential part of the digitalization of health care. They are based on clinical research results, electronic health records (EHR), as well as personal data of patients or users of medical products obtained from self-testing devices, such as wearable devices for monitoring work or sports [35]. Patient records usually include all treatment documentation: written and visual medical records, doctors' letters, e-prescriptions, and insurance payments. Siemens Medical Solutions USA, Inc. considers that the main producers of such data are medical products suppliers, support service providers (e.g., pharmaceutical companies), public and private institutions and patients.[10] In order to establish a legal regime for data trafficking in health care, it is important to define the entire list of entities possessing information that could potentially constitute medical confidentiality.

Big data-based, artificial intelligence-processed analytics will help develop personalized individual help in predictive models for large populations.

The correct use of big health data requires the collection of reliable data, including medical records, genomics and information obtained from various applications.

Electronic Health Cards (EHR) represent the «patient data repository in digital form, secure storage and exchange accessible to many users» [36].

[8] Preparing For the Future of Artificial Intelligence. Executive Office of the President National Science and Technology Council Committee on Technology, October 2016. URL: https://obamawhitehouse.archives.gov/sites/ default/files/ whitehouse_files/microsites/ostp/NSTC/preparing_for_the_ future_of_ai.pdf (accessed on 17.03.2021).

[9] Artificial intelligence, automation, and the economy. Executive Office of the President Washington, D.C. 20502. December 20, 2016. 55 p. URL: https://obamawhitehouse.archives.gov/ sites/whitehouse.gov/files/ documents/Artificial-Intelligence-Automation-Economy.pdf. (accessed on 28.03.2021).

[10] Knight M. Healthcare Dives into Big Data Increasingly used data-driven care protocols will change healthcare delivery systems globally. August 5, 2015. URL: https://www.siemens-healthineers.com/en-us/news-and-events/mso-big-data-and-healthcare-1 (accessed on 28.03.2021).

In addition to EHR, electronic medical records (EMR) are used in the digitization of health care, representing "digital systems that functionally provide patient history, patient demographic data and registration data" for the use of professionals, often based on telemedicine approaches [37]. With EMR systems, it is possible to solve many problems related to the analysis of data in the healthcare system and improvement of operational processes involving AI.

However, not all foreign scientists welcome the concept of digital health care. Some have emphasized that digitalization based on big data, mobile health, e-health, telemedicine and tele-health clearly undermines traditional health systems. A number of scientists cite the destructive trend of further digital transformative technologies, such as artificial intelligence and robotics [38–40].

LEGAL REGULATION

In the EU, a strategy for the development and use of artificial intelligence was adopted in April 2018. From mid-2019, national plans for the development of AI, the creation of funds to finance start-ups for the development of AI technologies were adopted. Financing of the "Digital Europe" program (the duration of the program — from 2021 to 2027) according to the current version approved by the European Parliament and the Council of the European Union, is 9.2 billion euros. Among them, it is planned to allocate 2.7 billion euros for high-performance computing; 2.5 billion euros for artificial intelligence; 2 billion euros to increase cyber security and confidence; 1.3 billion euros to ensure wide use of digital technologies in the economy and society and 700 million euros to develop digital skills.

The European Parliament Resolution on Comprehensive European Industrial Policy on AI and Robotics, published in February 2019, establishes general provisions of a recommendatory nature. For example, the resolution contains recommendations on the application of AI technologies in the field of health.

A digital platform has been established by the joint task force from ITU (International Telecommunication Union) and WHO (World Health Organization) on AI for Health (FG-A14H) to test and compare AI applications for health.[11] The task force itself includes experts and representatives of stakeholders in the field of research, practice, ethics and legal regulation in health, capable of developing guidance documents related to ethics, evaluation, legal regulation of AI for application in various fields of health care — in ophthalmology, histopathology, dentistry, radiology, etc. All documentation of this group is publicly available.

The EU Ethics Guide for AI is an important tool for the application of AI technologies in different areas of society. The management has established a common three AI principles that can be summed up as doing well, doing no harm, and operating transparently. These principles are expected to lead to sustainable development of AI technologies: they must be safe, accountable, non-discriminatory, and human.

These EU documents are, in fact, the basic legal basis for the legal regulation of AI in the field of health care. Many developed States have rules for the collection of patient data. These include the Health Insurance Portability and Accountability Act (HIPPA) and the European General Data Protection Regulation (GDPR). In the US, HIPPA also protects patients' medical data. Such data collection is necessary for machine learning and the use of AI in health care.

The proposed path of legal regulation development in the EU states is reasonable, but does not take into account the risks arising from such activities [41]. Insurance is one way to cover health risks and to recover costs for medical care. The Federal Republic of Germany adopted the German Digital Health Act in November 2019 (Digitale

[11] ITU-WHO Task Force on Artificial Intelligence. URL: https://ru.abcdef.wiki/wiki/ITU-WHO_Focus_Group_on_Artificial_Intelligence_for_Health (accessed on 28.03.2021).

Versorgung Gesetz — DVG), which entitles all persons covered by statutory health insurance to reimbursement for certain digital medical applications (i.e. insurers will pay for their use).[12] It is known that the German compulsory health insurance system is one of the largest in the world. Approximately 90% of the population (about 75 million people) in Germany is covered by public health insurance and the remaining 10% by private insurance. The DVG (Digital Health Care) Act entitles those who are insured by an independent provider of compulsory health insurance in Germany to receive insurance payments for certain digital medical applications.

As a rule, insured persons are entitled to insurance payments on digital medical applications if they meet the following criteria:

1. Are medical products with lower risk.

2. Their main function is based on digital technology.

3. They are designed to support the monitoring, detection, relief or treatment of diseases, injuries, care provided by service providers.

4. They were included in the newly established official register of digital medical applications maintained by the Federal Institute of Medicines and Medical Products of Germany (Bundesinstitut für Arzneimittel und Medizinprodukte — BfArM).

5. They are used either with the consent of the medical insurer or by appointment of the attending physician or psychotherapist [SGB V, § 33a(1)].

In addition, DVG also aims to accelerate the introduction and use of telemedicine. In particular, within the framework of DVG the patient can easily take advantage of video consultations. During such consultation, the patient may be informed of the circumstances essential to consent to medical care, including its nature, scope, implementation, anticipated risks and consequences (SGB V, § 291g (4); BGB 14, § 630e). DVG also contains provisions to make demographic data from health insurers more suitable for research purposes (SGB V, § § 303a-303f). In particular, under the General Data Protection Regulation 2016/679 (GDPR), DVG allows certain beneficiaries, such as universities and publicly funded research institutions (e.g., the Max Planck Society) are process certain demographic data from health insurers for specific research purposes, especially for the analysis of treatment or care processes over longer periods (SGB V, § § 303b, 303e(1) and (2).

In the US, the federal Medicare Program provides coverage for more than 60 million Americans.[13] Since 2019, Medicare Part B (outpatient health insurance) provides insurance coverage for some telemedicine services such as office visits, psychotherapy and other consultations. In 2020, during the COVID-19 pandemic, Medicare significantly expanded its telemedicine coverage policy.

Health insurance, which stimulates the accelerated adoption of AI technologies, provides both Germany and the US with cost recovery for innovative technologies and useful digital health solutions.

Over 318,000 mobile health apps are available worldwide, and about 200 apps are added to app stores every day [42].

Researchers of the problems of legal regulation of digitalization of healthcare Sara Gerke, Ariel D. Stern and Timo Minssen come to quite legitimate conclusions about the importance of ensuring the safety of patients and consumers when using medical applications. In doing so, safety, functionality, quality, data protection, data security and positive impact on care should be the criteria for their evaluation. The authors of the article thought, «risk-oriented approach can serve as a useful starting point for the development of

[12] Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz — DVG) [Digital Healthcare Act] of 9 December, BGBl I at 2562 (Germany, 2019).

[13] CMS. CMS fast facts. URL: https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/CMS-Fast-Facts/index (2020).

standardized, evidence-based assessment processes and legal requirements for digital health solutions. However, legislative bodies will also need to ensure that they do not over-regulate the sector in a way that would make it too burdensome for producers to comply with qualification requirements and to recover losses, thereby preventing or slowing down innovation» [43].

Independent direction of legal provision of application of AI technologies in the field of health care is development of mandatory requirements for medical products (medical devices). Unlike the development of drugs that are considered dangerous until proven otherwise, the development of medical products should be based on compliance with certain risk requirements.[14] For example, in the United States, the Food and Drug Administration is the regulator of medical products (FDA). The FDA Department of Industry and Consumer Education (DICE) determines whether a digital «device» is a medical product or a product intended for recreational purposes.[15] Medical products can presumably be classified according to the risk profile currently defined by the FDA as Class I, II or III in order of increasing risk. Class I devices require little security testing. Today, about 50% of medical products fall into this category and 95% of them are exempt from the regulatory process.[16]

There are international organizations to streamline the standardization of medical products in digital health care. For example, the Society of Digital Medicine develops a resource on its website (www.dimesociety.org) to track and monitor various standards, documents and requirements.[17]

In a review article "Digital Medicine: A Measurement Textbook", the authors reveal three basic principles of biomedical ethics, as described in the Belmont Report (1978),[18] which are applicable for research of digital technologies:

1. Respect for the individual. This principle is demonstrated through an informed consent process that occurs when a person is given the information necessary to make an informed decision on whether to volunteer. How this information is communicated is important because voluntary participation in the study differs from, say, accepting the Terms of Service (ToS) for accessing the application or signing the consent form for receiving medical care.

2. Charity. This is when the assessment of the likelihood and magnitude of potential harm is compared with the possible benefits to the participant, the people it represents and society.

3. Equity. This principle focuses on the equitable sharing of the benefits and burdens of research and development [44].

In scientific articles on digitalization of health care authors pay attention to the problems of information security and protection of information, both personal and personal medical information.

Healthcare has witnessed the spread of vulnerabilities, especially in related technologies, many of which are vital: Johnson & Johnson insulin pumps, implantable St. Jude Medical heart devices, and WannaCry extortion attacks, which infected 200 000 computers, many of which

[14] U.S. Food and Drug Administration. Information Sheet Guidance For IRBs, Clinical Investigators, and Sponsors. Silver Spring, MD: U.S. Food and Drug Administration. January 2006. 15 p. URL: https://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM126418.pdf (accessed on 28.03.2021).

[15] U.S. Food and Drug Administration. Multiple Function Device Products. Policy and Considerations. Silver Spring: U.S. Food and Drug Administration. 2018 Apr 27. 18 p. URL: https://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm605683.pdf (accessed on 28.03.2021).

[16] U.S. Food and Drug Administration. Reclassification [Internet] [updated 2018 Dec 13]. URL: https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/ucm080412.htm (accessed on 28.03.2021).

[17] The Digital Medicine Society. The Digital Medicine (DiMe) Society. 2019. URL: www.dimesociety.org (accessed on 28.03.2021).

[18] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, Department of Health, Education and Welfare (DHEW). The Belmont Report. Washington: United States Government Printing Office. 1978 September 30. URL: https://videocast.nih.gov/pdf/ohrp_belmont_report.pdf (accessed on 28.03.2021).

are part of the critical infrastructure of hospitals, in 150 countries [45].

In addition to hacking attacks, medical professionals can also illegally disseminate information. A number of researchers developed the project "Hippocratic oaths for connected medical devices" (HOCMD).[19] The Oath sets out a number of principles of safety and ethics, including "safety by design" and "sustainability and containment" [46].

Describing the issues of the use of AI in healthcare and its legal regulation in the Russian Federation, it is necessary to mention the national project «Healthcare», which identified the key task of digitalization of the healthcare system.[20] Within the framework of the project, the federal project "Creation of a unified digital circuit in health care on the basis of the unified state information system in the field of health care was adopted (USHIS)",[21] which will categorize all health organizations with information systems, information and telecommunication networks, automated management systems as critical information infrastructure (KII).[22] Will be subject to categorization AI, ITS, ICS, which provide management, technological, production, financial-economic and (or) other critical business processes within the framework of the performance of functions (powers) or the implementation of activities of a health organization. As a result, a list of potentially significant KII facilities for health care organizations will be established and a single digital circuit in health care will be created based on a unified State information system in the health sector (USHIS). The overall objective of digitalization is to ensure that reliable and structured information is readily available to doctors and managers.

For citizens, the main patient-oriented service is "My Health", allowing you to make an appointment with a doctor; attach to a medical organization; submit an application on the choice of an insurance medical organization; to obtain information on the provided medical services and their cost; to obtain a number of documents, for example, a medical certificate on admission to the driving vehicle, electronic prescriptions, a certificate of vaccination from COVID-19, etc. The service is actively developing.

With the completion of the formation of the Unified State Health Information System (USHIS) within the next four years, electronic patient records Central Medical Imaging Archives and Unified Laboratory Systems to be accessible to all medical institutions in the country. All state medical institutions will be required to provide citizens with access to their electronic medical documents through the Single Portal of State Services.

Telemedicine is being actively developed to make medicine accessible to the geography of our country,[23] which implemented by many health organizations.[24]

Russia is actively working to standardize the use of AI in the field of health care. Thus, in 2020, the first draft of a national

[19] I Am the Cavalry. Hippocratic Oath for Connected Medical Devices. URL: https://www.iamthecavalry.org/domains/medical/oath/ (accessed on 28.03.2021).

[20] Passport of the national project "Healthcare" (approved by the Presidium of the Council under the President of the Russian Federation on strategic development and national projects, protocol of 24.12.2018 No. 16). The document was not published (ConsultantPluse).

[21] National project "Healthcare". The federal project "Creation of a unified digital circuit in healthcare on the basis of the unified state information system in healthcare sphere (USHIS)". Methodological recommendations for the categorization of critical health information infrastructure facilities (Version 1.0) (approved by the Ministry of Health of Russia 05 April 2021). The document was not published (ConsultantPluse).

[22] Federal Law from 26 July 2017 No. 187 "On the Security of the Critical Information Infrastructure of the Russian Federation", 31.07.2017, No. 31 (Part I), art. 4736.

[23] Order of the Ministry of Health of the Russian Federation from 30.11.2017 No. 965 "On approval of the procedure of organization and provision of medical assistance with the use of telemedicine technologies". Official Internet Legal Information Portal. URL: http://www.pravo.gov.ru, 10.01.2018 (accessed on 28.03.2021).

[24] It should be noted that, according to the law on telemedicine, a physician was not entitled to make any recommendations without an initial appointment, and remote diagnosis was prohibited. There are no admission restrictions.

standard for AI in healthcare in Russia was developed, which regulates the conduct of clinical trials of medical AI-systems (AIS). In order to coordinate the work on unification and standardization of requirements for AI systems in healthcare, as well as the establishment of certification requirements for medical products, a subcommittee "Artificial Intelligence in Healthcare" was established as part of TC 164 (PC 01/TC 164).[25]

In accordance with the long-term work plan of the Technical Committee for Standardization "Artificial Intelligence" (TC 164), created on the basis of the Joint-Stock Company "Russian Venture Company" (RVC), by 2027, it is planned to develop about 50 standards in the field of AI in health care in individual areas, including general requirements and classification of AI systems in clinical medicine, big data in health care, functional diagnostics, radiation diagnostics, remote monitoring systems, histology, medical decision support systems, image reconstruction in diagnosis and treatment, medical systems of analysis and forecasting, as well as educational programs in health care.

A task force of the International Organization for Standardization Technical Committee No. 215 "Informatization of Health" has been established to develop standards at the international level (ISO) — TC 215 ISO Health informatics. Its task is to form the requirements to confirm the compliance of high-tech AI systems for their application in health care. In particular, risk assessment of the impact on the health of patients, determination of consistency of medical databases, correctness of withdrawal of decisions and organization of information exchange between AI systems and electronic medical cards, medical products and health information systems.

Based on the analysis provided, it follows that the Russian Federation is actively working on the digitalization of health care, standardizing the use of AI, domestic experts are involved in a number of international organizations, including: the digitalization of health care, there is an exchange of experience both within the state, and with foreign colleagues, telemedicine is being actively developed, AI technologies are being introduced into the high-tech medical care sector, and a single health information circuit is being created in the country on the basis of USHIS.

Taking into account the analysis of the existing legal regulation in Russia and abroad, we propose to build the institutional-legal model of the development of patient-oriented medicine on the basis of the use of AI as follows.

1. The model is based on a paradigm shift from client-oriented to patient- oriented medicine.

2. The main elements of the model are:

- digitalization of management processes in the health system (management of medical services, digitization of the formation and circulation of medical data, digitization of the organization's access to them, electronic medical cards, electronic medical records);
- digitization of medical care (e.g., MRI, CT, high-tech robotic operations, database creation and turnover);
- integration processes in the health care system, unification of digital functions of management of medical institutions and digital processes of organization and provision of medical care and medical services in a unified system of health organizations, independently of organizational andlegal forms of specific medical and pharmaceutical institutions;
- development of mandatory requirements for medical products (medical devices) using risk-based AI and their implementation through pilot legal regimes (regulatory sandboxes). The legal framework

[25] Order of Rosstandart from 31 December 2019 No. 3471 «On making amendments to the order of the Federal Agency for Technical Regulation and Metrology from 25 July 2019 No. 1732 «On creation of technical committee for standardization Artificial Intelligence». The document was not published. ConsultantPluse.

for their regulation is the basic AI principles applicable, inter alia, in the field of health;

• intensive spatial development of telemedicine, especially in hard-to-reach areas of our country;

• introduction and promotion of the institution of self-regulation of medical ethics, including the observance of medical confidentiality in the context of the digitization of medical care and medical services;

• Introduction of a financial and legal mechanism for the use of AI in the health-care system, the basis of which should be the institute of health insurance by including in the list of insurance coverage under the CHI the provision of medical care using AI, provision of telemedicine services, etc.;

• at the international level, the Russian Federation should be represented and participate more actively in the work of international organizations dealing with digital health development.

CONCLUSION

1. The article offers an institutional-legal model of development of patient-oriented medicine based on use of innovative technologies, including AI technologies. Rapid changes are taking place in the health-care system based on technologies that need to be managed in a systemic manner. In this approach, separate rules of different branches of law, together with standards (technical regulation), are considered in close connection with each other.

2. It was found that despite the complexity of forecasting the problems and consequences caused by the use of AI technologies in the field of health, scientific publications highlight the following problems of legal regulation:

• problems with the correct determination of legal responsibility for acts committed using artificial intelligence technologies [1–7, 15];

• problems associated with adapting the newly created norms of law aimed at regulating relations in this sphere to the general body of existing legislation [1, 5];

• problems of ensuring the confidentiality of personal information and personal information in the context of the collection and storage of a large amount of personal information in electronic and automated databases (institute of medical confidentiality);

• problems of standardization of AI technologies in medicine;

• problems of licensing medical products using AI technologies;

• financial and legal problems of providing medical services in digital format.

3. Development of legal regulation of AI in the field of health should be carried out in the following main directions:

– depending on the work performed, the services provided in the digital health system and the high-tech products manufactured for medicine (medical products) and pharmacology;

– depending on the application of weak or strong AI in digital health care;

– depending on the risks and regulatory requirements of cybersecurity, information security, cyberphysical systems security;

– depending on the financing of the work performed, the services provided in the digital health system and the high-tech products manufactured for medicine (medical products) and pharmacology (Stimulation of the State, improvement of the legal mechanism for payment of medical services and provision of medical assistance using AI technologies through insurance companies).

4. Foreign experience of developed countries (Germany and the United States) demonstrates that legal regulation of digital medicine significantly ahead of domestic in a number of areas. In particular, the institution of compulsory health insurance needs to be activated more actively in order to foster innovation in the health system.

It is also a positive experience to develop the bioethical principles needed to engage citizens as volunteers in medical research.

It is appropriate for the medical community at the level of self-regulation of medical ethics to modernize the "Hippocratic Oath" for the digital healthcare environment.

5. International standardization and domestic technical standardization demonstrate the approach to technical regulation of medical products as narrow profile. This approach should be evaluated positively. So, in 2020–2021, numerous domestic standards have been adopted in the form of GOSTs and preliminary GOSTs establishing terminology and safety requirements. This unquestionably regulates the requirements for AI technologies in digital healthcare.

The second big part of the legal regulation in the field of technical regulation is special standards in medicine, mainly for certain medical products. In addition, requires improvement and the institute of licensing of medical products (separation with household goods in the field of health), accelerated introduction of innovative technologies AI, robotics, CPS, work in medicine, used in finished products or in the provision of health services and works. Finished products require more careful control and supervision by the State. In particular, in addition to compliance with standards, medical products are subject to compulsory licensing (this is the cost, effort and most importantly — time).

6. It is proposed to apply more regulatory sandboxes — experimental legal regimes for the following chain: new technologies — high-tech medicine — experimental treatments.

ACKNOWLEDGEMENTS

The article is based on the results of the III stage of the basic research work of the University of Finance on "The theory of legal regulation of artificial intelligence, robots and robotics objects in the Russian Federation". Financial University, Moscow, Russia.

REFERENCES

1. Morkhat P.M. Artificial intelligence: Legal view. Moscow: Buki Vedi; 2017. 257 p. (In Russ.).
2. Yastrebov O.A. Discussion about the prerequisites for assigning robots the legal status of "electronic persons". *Voprosy pravovedeniya*. 2017;(1):189–202. (In Russ.).
3. Ruchkina G.F., ed. The theory of legal regulation of artificial intelligence, robots and robotics objects in the Russian Federation. Moscow: Prometheus; 2020. 296 p. (In Russ.).
4. Popova A.V. Legal aspects of artificial intelligence ontology. *Gosudarstvo i pravo = State and Law*. 2020;(11):115–127. (In Russ.). DOI: 10.31857/S 102694520012531–5
5. Popova A.V. Ethical principles of interaction with artificial intelligence as a basis of legal regulation. *Pravovoe gosudarstvo: teoriya i praktika = The Rule of Law State: Theory and Practice*. 2020;(3):34–43. (In Russ.). DOI: 10.33184/pravgos-2020.3.4
6. Gorokhova S.S. Artificial intelligence-based technologies: Prospects and liability in the legal environment. *Yurist = Jurist*. 2021;(6):60–67. (In Russ.). DOI: 10.18572/1812–3929–2021–6–60–67
7. Sviridova E.A., Rahmatulina R.S., Shaydullina V.K., Gorohova S.S., Lapina M.A. Issues of economic and legal responsibility in the application of artificial intelligence technologies in the coal industry. *Ugol'*. 2020;(7):57–61. (In Russ.). DOI: 10.18796/0041–5790–2020–7–57–61
8. Serova O.A. Problems of development of the methodology of civil law research in the digital age. *Metodologicheskie problemy tsivilisticheskikh issledovanii = Methodological Problems of the Civil Law Researches*. 2019;1(1):351–362. (In Russ.).
9. Vaipan V.A., Egorova M.A., eds. Legal regulation of economic relations in modern conditions of digital economy development. Moscow: Yustitsinform; 2019. 376 p. (In Russ.).
10. Gnidenko I.G., Egorova I.V., Merdina O.D. Intelligent systems as a vector of enterprise information security development. *Peterburgskii ekonomicheskii zhurnal = Saint-Petersburg Economic Journal*. 2020;(1):147–154. (In Russ.). DOI: 10.25631/PEJ.2020.1.147.154

11. Elchaninova N.B. The problems of legal protection of intellectual property on the Internet. *Obshchestvo: politika, ekonomika, pravo = Society: Politics, Economics, Law*. 2016;(12):118–120. (In Russ.).
12. Ruchkina G.F., Lapina M.A., eds. Public-legal means of economic and financial development efficiency. Pt. 1. Public-legal means of digitalization of management, economic and social space. Moscow: RuScience; 2020. 246 p. (In Russ.).
13. Lazarev V.V., Gadzhiev Kh.I., Alimov E.V. et al. Data protection: Scientific and practical comment on judicial practice. Moscow: Kontrakt; 2020. 176 p. (In Russ.).
14. Luparev E.B. Legal regulation of the export of telemedicine services: Russian and European experience. *Kubanskoe agentstvo sudebnoi informatsii PRO-SUD-123.RU: Yuridicheskii setevoi elektronnyi nauchnyi zhurnal*. 2020;(1):69–78. (In Russ.).
15. Ponkin I.V., Redkina A.I. Artificial intelligence from the point of view of law. *Vestnik Rossiiskogo universiteta druzhby narodov. Seriya: Yuridicheskie nauki = RUDN Journal of Law*. 2018;22(1):91–109. (In Russ.). DOI: 10.22363/2313–2337–2018–22–1–91–109
16. Dudnik V. Yu. A model of personalized medical care. *Akademicheskii zhurnal Zapadnoi Sibiri = Academic Journal of West Siberia*. 2016;12(1):98–100. (In Russ.).
17. Nikolaev N. Patient-oriented antihypertensive therapy: Clinical guidelines for practitioner. *Vrach = The Doctor*. 2016;(4):82–85. (In Russ.).
18. Menshikov A.N. Development of a model of patient-oriented healthcare in the Republic of Bashkortostan. *Obyazatel'noe meditsinskoe strakhovanie v Rossiiskoi Federatsii*. 2016;(3):34–37. (In Russ.).
19. Troshin S.P., Legkova I.N., Gavrilenko K.V. Creation of a patient-oriented models of obligatory medical insurance based on the application of current information technologies. *Obyazatel'noe meditsinskoe strakhovanie v Rossiiskoi Federatsii*. 2016;(6):34–41. (In Russ.).
20. Averin A.V., Shvardakova O.A. Patient-oriented technologies in the practice of a multidisciplinary medical organization. *Glavnaya meditsinskaya sestra*. 2015;(10):39–48. (In Russ.).
21. Kitaeva E.A., Kitaev M.R., Suetina T.A., Salyahova L. Ya., Vafin A. Yu. Implementation of patient oriented program aiming to develop compliance to drug therapy in rural area patients (in terms of Rybnaya Sloboda district of the Republic of Tatarstan). *Vestnik sovremennoi klinicheskoi meditsiny = The Bulletin of Contemporary Clinical Medicine*. 2017;10(2):64–71. DOI: 10.20969/VSKM.2017.10(2).64–71
22. Schunk M., Stark R., Reitmeir P., Meisinger C., Holle R. Towards patient-oriented diabetes care: Results from two KORA surveys in Southern Germany. *Journal of Diabetes Research*. 2015;2015:368570. DOI: 10.1155/2015/368570
23. Zheng K., Haftel H.M., Hirschl R.B., O'Reilly M., Hanauer D.A. Quantifying the impact of health IT implementations on clinical workflow: A new methodological perspective. *Journal of the American Medical Informatics Association*. 2010;17(4):454–461. DOI: 10.1136/jamia.2010.004440
24. Huang Z., Lu X., Duan H. On mining clinical pathway patterns from medical behaviors. *Artificial Intelligence in Medicine*. 2012;56(1):35–50. DOI: 10.1016/j.artmed.2012.06.002
25. Bouarfa L., Dankelman J. Workflow mining and outlier detection from clinical activity logs. *Journal of Biomedical Informatics*. 2012;45(6):1185–1190. DOI: 10.1016/j.jbi.2012.08.003
26. Ozkaynak M. Characterizing workflow in hospital emergency departments. Madison, WI: University of Wisconsin-Madison; 2011.
27. Blakemore K. Are professionals good for you? The example of health policy and health professionals. In: Blakemore K. Social policy: An introduction. Buckingham, Philadelphia: Open University Press; 1998.
28. Mukharyamova L.M., Kuznetsova I.B., Vafina G.G. Sick, patient, client: The positions of labor migrants in the Russian health care system (evidence from the Republic of Tatarstan). *Vestnik sovremennoi klinicheskoi meditsiny = The Bulletin of Contemporary Clinical Medicine*. 2014;7(1):43–49. (In Russ.).
29. Goncharov L.A. From state paternalism to client orientation? Russian government institutions in search of effective models for the formation and training of the corps of state civil servants. *Vestnik Voronezhskogo instituta ekonomiki i sotsial'nogo upravleniya*. 2017;(2):68–74. (In Russ.).

30. Ponkin I. V., Ponkina A. A., Laptev V. S. Concept of digital public health and e-health. *Narkologiya = Narcology*. 2014;13(6):34–40. (In Russ.).
31. Grishchenko G. A. Artificial intelligence in public administration. *Rossiiskii yuridicheskii zhurnal = Russian Juridical Journal*. 2018;(6):27–31. (In Russ.).
32. Vasin S. G. Artificial intelligence in state management. *Upravlenie*. 2017;5(3):5–10. (In Russ.).
33. Kraus S., Schiavone F., Pluzhnikova A., Invernizzi A. C. Digital transformation in healthcare: Analyzing the current state-of-research. *Journal of Business Research*. 2021;123:557–567. DOI: 10.1016/j.jbusres.2020.10.030
34. Marques I. C.P., Ferreira J. J.M. Digital transformation in the area of health: systematic review of 45 years of evolution. *Health and Technology*. 2020;10(3):575–586. DOI: 10.1007/s12553–019–00402–8
35. North K., Maier R., Haas O., eds. Knowledge management in digital change. Cham: Springer-Verlag; 2018. 394 p. (Progress in IS). DOI: 10.1007/978–3–319–73546–7
36. Häyrinen K., Saranto K., Nykänen P. Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics*. 2008;77(5):291–304. DOI: 10.1016/j.ijmedinf.2007.09.001
37. Chakravorty T., Jha K., Barthwal S. Digital technologies as enablers of care-quality and performance: A conceptual review of hospital supply chain network. *IUP Journal of Supply Chain Management*. 2018;15(3):7–25.
38. Agnihothri S., Cui L., Delasay M., Rajan B. The value of mHealth for managing chronic conditions. *Health Care Management Science*. 2020;23(2):185–202. DOI: 10.1007/s10729–018–9458–2
39. Patrício L., Teixeira J. G., Vink J. A service design approach to healthcare innovation: From decision-making to sense-making and institutional change. *AMS Review*. 2019;9(8):115–120. DOI: 10.1007/s13162–019–00138–8
40. Yousaf K., Mehmood Z., Awan I. A., Saba T., Alharbey R., Qadah T., Alrige M. A. A comprehensive study of mobile-health based assistive technology for the healthcare of dementia and Alzheimer's disease (AD). *Health Care Management Science*. 2020;23(2):287–309. DOI: 10.1007/s10729–019–09486–0
41. Lapina M. A. Theoretical and legal aspects of risk management. *Gosudarstvo i pravo = State and Law*. 2015;(2):35–44. (In Russ.).
42. The rise of mHealth apps: A market snapshot. Liquid State. Mar. 26, 2018. URL: https://liquid-state.com/mhealth-apps-market-snapshot/
43. Gerke S., Stern A. D., Minssen T. Germany's digital health reforms in the COVID-19 era: Lessons and opportunities for other countries. *npj Digital Medicine*. 2020;3(1):94. DOI: 10.1038/s41746–020–0306–7
44. Coravos A., Goldsack J. C., Karlin D. R., Nebeker C., Perakslis E., Zimmerman N., Erb M. K. Digital medicine: A primer on measurement. *Digital Biomarkers*. 2019;3:31–71. DOI: 10.1159/000500413
45. Brewster T. Medical devices hit by ransomware for the first time in US hospitals. Forbes. May 17, 2017. URL: https://www.forbes.com/sites/thomasbrewster/2017/05/17/wannacry-ransomware-hit-real-medical-devices/#dad2b18425cf
46. Woods B., Coravos A., Corman J. D. The case for a Hippocratic Oath for connected medical devices: Viewpoint. *Journal of Medical Internet Research*. 2019;21(3):e12568. DOI: 10.2196/12568

ABOUT THE AUTHOR

Marina A. Lapina — Dr. Sci. (Law.), Professor, Chief Researcher, Professor, Department of International and Public Law, Financial University, Moscow, Russia
https://orcid.org/0000-0003-0320-161X
MALapina@fa.ru

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 02.09.2021; revised on 20.09.2021 and accepted for publication on 27.12.2021.
The author read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-186-195
UDC 336.767.3:51(045)
JEL G120, Y80

Interest Rate Risk of Bonds in the Condition of a Changing Key Rate

N.V. Popova
Plekhanov Russian University of Economics, Moscow, Russia

ABSTRACT

The article is devoted to the analysis of the behavior of the interest rate risk of bonds in the conditions of a changing key interest rate. As known, the key rate is an instrument of monetary regulation of the Central Bank of the Russian Federation. During periods of instability, the key rate may change, which leads to changes in yields in the bond market. The latter, in turn, means that bonds on the Russian market are exposed to interest rate risk. When investing in federal loan bonds (OFZ), interest rate risk becomes the main type of risk, since there is no credit risk in such bonds. **The aim** of the paper is to obtain proof of the dependence of the interest rate risk of bonds on the term to maturity. The author applies **methods** of differential calculus to obtain the proof. **The novelty** of the research is that there is no similar proof in the literature. The instability of interest rates in the market persists at the present time, which allows us to speak about the **relevance** of this work. **Results:** it is established that with fixed values of the coupon rate, initial yield and the amount of interest rate change, the interest risk of bonds increases with an increase in the term to maturity. For long-term bonds sold at a discount, there is a term of a maximum interest rate risk. The formula for the approximate value of the term of maximum is obtained. Proven statements are confirmed by calculations, are consistent with previously performed studies, and are in line with market observations. The author comes to the **conclusion** that the proof obtained in the article of the dependence of the interest rate risk of bonds on the term to maturity can be used to analyze the behavior of the interest rate risk of bonds in the conditions of a changing key interest rate. **The practical significance**: the results of the research can be useful to the issuer and investor, as well as in theory when studying the investment properties of bonds.

Keywords: bonds; interest rate risk; mathematical methods; term to maturity; key interest rate

For citation: **Popova N.V. Interest rate risk of bonds in the condition of a changing key rate. *Finance: Theory and Practice*. 2022;26(3):186-195. DOI: 10.26794/2587-5671-2022-26-3-186-195**

INTRODUCTION

During periods of economic crises or economic difficulties, the risks of investing are exacerbated. The types of risk of investing in bonds is given in the book by F.J. Fabozzi [1]. There are two main types of bond risk: interest rate risk and credit risk [2]. Credit risk, or default risk, is the risk that an issuer will default on its obligation to pay a bond. In fixed income markets without credit risk, the main type of risk is interest rate risk [3]. Similarly, the statement of F. J. Fabozzi [4, p. 1]: "The major risk faced by participants in the bond market is interest rate risk".

F. J. Fabozzi [1, p. 22] defines interest rate risk: "interest rate risk is the risk that an increase in interest rates will lead to lower bond prices". The literature also gives a broader definition of interest rate risk as "the possibility of changing the price of a bond under the influence of changes in interest rates" [3, p. 208]. Publications [2, 5–12] are devoted to the study of various types of investment risk in bonds, including the analysis of factors that contribute to changes in market interest rates. The main macroeconomic factors are inflation, interest rate, political and economic risks. According to A.K. Isaev and V.N. Dem'yanov [5], inflation of the currency in which a given issue is nominated is a fundamental factor determining the yield of corporate bonds in a particular country. According to L. J. Gitman and M. D. Joehnk [13, p. 467], "In fact, investors are most worried about inflation. This not only undermines the purchasing power of the principal amount of the loan, but also greatly affects the dynamics of interest rates." In Russia, the Bank of Russia uses the key interest rate to regulate inflation [14, 15]. The value of the key rate affects the yield of bonds. I.M. Balkoev [16] considered the mechanism of change in the market profitability under the influence of changes in the key rate. As a result, according to [16, p. 35], "the influence of the Central Bank rate on bonds lies in the fact that when it increases, the yield on the bond market also grows, and when it decreases, the yield on bonds proportionally decreases." The impact of the key rate on the growth of market interest rates could be observed in the Russian bond market in 2021. In March 2021, the Bank of Russia began raising the key rate in order to reduce inflation and increased this indicator by 4.25 p.p.— up to 8.5% in December. According to the Moscow Exchange[1] for the period from March to December 2021, the increase in the key rate was accompanied by an increase in yields on the bond market and, as a result, an increase in the number of bonds sold at a discount. At the same time, bonds purchased on the Russian market prior to March 2021 were subject to significant interest rate risk.

A number of studies have noted that bond parameters such as coupon rate, term to maturity, and yield to maturity affect the interest rate risk of a bond. According to F. J. Fabozzi [1, p. 22] "the sensitivity of the price of each particular issue to changes in market interest rates depends on the parameters of the bond, namely the coupon rate and term to maturity". The influence of yield to maturity or, which is the same, the level of market interest rates on the percentage change in the price of a bond was proved in [17] and is formulated as follows: for a given term to maturity and coupon rate, the higher the initial yield level, the smaller the relative (percentage) price change bonds when its yield changes by a given value. Thus, "with a given change in returns, price volatility is higher in a market where returns are low, and vice versa: with high returns, volatility is low" [1, p. 95]. The effect of the coupon rate on the interest rate risk of a bond was proved in [18] and is formulated as follows: for a given term to maturity and initial yield, the relative (percentage) change in the price of a bond when its yield changes by a given value is greater, the lower the coupon rate. Studies of the effect of the coupon rate in [6, 10]

[1] Moscow Exchange. URL: http://www.moex.com/ (accessed on 10.05.2022).

confirm this statement. For example, [6, p. 254]: "Compared to coupon bonds, zero-coupon bonds are more susceptible to price volatility due to fluctuations in the market interest rate". Based on the results of studies in [6], it was concluded that the coupon rate is an important characteristic of the issue, but its value is not as great as the term to maturity of the bond.

According to [1], term to maturity is a key characteristic of a bond and is of paramount importance when evaluating any bond. The authors of papers [2, 5, 6, 10, 11] consider term to maturity to be a significant risk factor when investing in bonds, primarily due to the exposure of bonds to interest rate risk. According to [10, p. 626], bonds with longer maturities directly indicate that these bonds are subject to risks. First of all, the risks associated with interest rates, reflecting the dynamics of the economy and business. According to [6, p. 248] the probability of large risks may arise with long-term investments. This higher potential risk is related to the bonds' exposure to interest rate and price risk. As a result, bonds with longer maturities require guarantees, which typically have high yield-to-maturity requirements. According to [5, p. 139], "as a rule, investors are interested in a large risk premium when buying "long" bonds, since the uncertainty is higher with a long maturity". The authors of a number of papers note the relationship between the amount of interest rate risk and the term to maturity. According to [10, p. 621] as a result of changes in interest rates, "greater price changes will occur in bonds with a longer maturity". This conclusion is consistent with the statement in [6, p. 255, 256]: "bonds with a longer maturity tend to be subject to greater price volatility in the bond market", as well as in [2, p. 18]: "Longer bonds should provide the investor with an additional risk premium associated with higher duration and volatility". It should be noted that all of the above statements about the relationship between the amount of interest rate risk and maturity are formulated on the basis of market observations.

The interest rate risk of a bond is estimated by the value of the relative (percentage) change in the price of the bond when the market interest rate changes by a given value $\Delta P/P$. In practice, the interest rate risk of bonds is assessed using duration and convexity. According to F. J. Fabozzi [4, p. 1] "a typical way to measure interest rate risk is to approximate the effect of changes in interest rates on a bond or portfolio of bonds using duration and convexity". This statement is confirmed by works devoted to assessing the interest rate risk of bonds [19–24].

Due to the accepted method of assessing interest rate risk, the effect of the term to maturity on the interest rate risk of a bond is established on the basis of the dependence of the duration of the bond on the term to maturity. There are mathematical proofs in the literature of dependence bond duration on the term to maturity [20, 25–27]. At the same time, in the literature much less attention has been paid to the problem of proof of dependence on the term to maturity of the directly interest rate risk of the bond, and therefore the theory of investing in financial instruments with fixed income seems to be incomplete. In [28], to solve the problem of the effect of term to maturity on the percentage change in the price of a bond when the yield changes by a given value, numerical sequences of the form $\{\Delta P_n/P_n\}$, were used, where the number of the sequence member n coincided with the number of coupon periods remaining until the bond is redeemed. In this article, methods of differential calculus were used to solve the problem.

According to F. J. Fabozzi [1, p. 95] the formulation of the dependence of the interest rate risk of a bond on the term to maturity is as follows: "For a given coupon rate and initial yield, the longer the term to maturity, the higher the price volatility". This statement is clarified in this paper.

MATERIALS AND METHODS

The problem of the dependence of the interest rate risk of a coupon bond on the

term to maturity is considered. As noted, the interest rate risk of a bond is estimated by the value of the relative (percentage) change in the bond price $\Delta P/P$ when the market interest rate changes by a given value. Let r and $P(r)$ — be the annual yield and the bond price at the initial time. We will consider a bond price change for an instantaneous change in the market interest rate, similarly to F. J. Fabozzi [1, p. 93]. Let $\tilde{r}$ — be the yield on a bond as a result of an instantaneous change in the interest rate by $\Delta r > 0$. If $\tilde{r} = r - \Delta r$, then $\tilde{r} < r$ — the rate instantly decreased; if $\tilde{r} = r + \Delta r$, then $\tilde{r} > r$ — the rate instantly increased. The bond price will be equal to $P(\tilde{r})$.

Since the yield and the bond price change in opposite directions, then $P(\tilde{r}) > P(r)$, if $\tilde{r} < r$ and $P(\tilde{r}) < P(r)$, if $\tilde{r} > r$. The value

$$\frac{P(\tilde{r}) - P(r)}{P(r)}, \text{ where } \tilde{r} < r, \qquad (1)$$

is called relative (percentage) growth, and the value

$$\frac{P(r) - P(\tilde{r})}{P(r)}, \text{ where } \tilde{r} > r, \qquad (2)$$

— relative (percentage) decrease in the bond price.[2] These are positive values.

To establish the dependence of the relative change in the bond price $\Delta P/P$ on the term to maturity n, it is sufficient to consider bonds for which coupon payments are paid once a year. In expressions (1) and (2), we will consider prices that do not contain accumulated coupon income, i.e. the quoted bond price immediately after the coupon payment, when n coupon periods remain to maturity.

Let us study the effect of term to maturity n on the value of the relative change in the bond price by differentiating with respect to the variable n function (1) for $\tilde{r} < r$ and function (2) for $\tilde{r} > r$.[3] For $\tilde{r} < r$ we get:

$$\left(\frac{P(\tilde{r}) - P(r)}{P(r)}\right)'_n = \frac{P(\tilde{r})}{P(r)}\left(\frac{P'_n(\tilde{r})}{P(\tilde{r})} - \frac{P'_n(r)}{P(r)}\right). \qquad (3)$$

If $\tilde{r} > r$ we get:

$$\left(\frac{P(r) - P(\tilde{r})}{P(r)}\right)'_n = \frac{P(\tilde{r})}{P(r)}\left(\frac{P'_n(r)}{P(r)} - \frac{P'_n(\tilde{r})}{P(\tilde{r})}\right). \qquad (4)$$

To set the signs of derivatives in expressions (3) and (4), it is necessary to set the sign of the difference $\left(\frac{P'_n(\tilde{r})}{P(\tilde{r})} - \frac{P'_n(r)}{P(r)}\right)$ in expression (3) and the sign of the difference $\left(\frac{P'_n(r)}{P(r)} - \frac{P'_n(\tilde{r})}{P(\tilde{r})}\right)$ in expression (4). To do this, the function $\frac{P'_n(r)}{P(r)}$ must be examined for monotonicity with respect to the variable r.

RESULTS AND DISCUSSIONS

Theorem. For given values of the coupon rate f, the initial yield to maturity r, and the amount of change in the interest rate $\Delta r > 0$ the following statements are true:

1) $\lim\limits_{n \to \infty} \frac{\Delta P}{P} = \frac{\Delta r}{\tilde{r}}$;

2) the interest rate risk of a bond sold at par or at a premium increases with maturity;

3) for bonds sold at a discount, there is a term of maximum interest rate risk.

Proof. By convention, r is the bond initial yield to maturity. Then the bond price at the initial moment is equal to:

[2] Encyclopedia of financial risk management. Ed. A. A. Lobanov and A. V. Chugunov. 4th ed. M.: Alpina Business. Books. 2005. 53 p.

[3] Differentiation with respect to an integer variable is used in studying the investment properties of bonds. For example, in the works of G. A. Hawawini [25], and B. Malkiel [29].

$$P(r)=A(1+r)^{-n}\left(1-\frac{f}{r}\right)+A\frac{f}{r}, \qquad (5)$$

where A — the face value of the bond; n — the term to maturity.

If $\tilde{r}$ — the bond yield as a result of an instantaneous change in the interest rate by an amount $\Delta r>0$, then the bond price will be equal to:

$$P(\tilde{r})=A(1+\tilde{r})^{-n}\left(1-\frac{f}{\tilde{r}}\right)+A\frac{f}{\tilde{r}}. \qquad (6)$$

1. Using formulas (5) and (6), we find the limit of expressions (1) and (2) at $n\to\infty$. We get:

$$\lim_{n\to\infty}\frac{\Delta P}{P}=\frac{\Delta r}{\tilde{r}}, \text{ where } \tilde{r}=r\pm\Delta r.$$

To prove the following statements of the theorem, it is necessary to investigate the monotonicity of the function $\frac{P_n'(r)}{P(r)}$ with respect to the variable r. This function looks like:

$$\frac{P_n'(r)}{P(r)}=\frac{(f-r)\ln(1+r)}{r+f\left((1+r)^n-1\right)}. \qquad (7)$$

We differentiate this function with respect to the variable r. Then we use approximate equalities:

$$(1+r)^n\approx 1+rn,\ \ln(1+r)\approx r.$$

We get:

$$\left(\frac{P_n'(r)}{P(r)}\right)_r'\approx$$
$$\approx\frac{r^2}{B^2(1+r)}\left[(r-f)fn^2-f(1+r)n-(1+f)\right], \qquad (8)$$

where B^2 — the square of the denominator of the right side of equality (7).

Let us prove the second and third statements of the theorem.

2. Under condition $f\geq r$, i.e. for bonds sold at par or at a premium, expression (8) is negative. Hence, for these bonds $\left(\frac{P_n'(r)}{P(r)}\right)_r'<0$ — the ratio $\frac{P_n'(r)}{P(r)}$ is a decreasing function of the variable r. If $\tilde{r}<r$, then $\frac{P_n'(\tilde{r})}{P(\tilde{r})}>\frac{P_n'(r)}{P(r)}$ and expression (3) is positive, which means $\left(\frac{P(\tilde{r})-P(r)}{P(r)}\right)_n'>0$. If $\tilde{r}>r$, then $\frac{P_n'(\tilde{r})}{P(\tilde{r})}<\frac{P_n'(r)}{P(r)}$ and then expression (4) is also positive, which means $\left(\frac{P(r)-P(\tilde{r})}{P(r)}\right)_n'>0$. Thus, for any change in interest rates, the interest rate risk of bonds sold at par or at a premium increases with maturity. The second statement of the theorem is proved.

3. Under the condition $f<r$, i.e. for bonds sold at a discount, expression (8) with an increase in $n>0$ changes sign at a certain value of the term n_0 from minus to plus. This follows from the expression in square brackets, where there is a square trinomial with respect to n with a positive coefficient at n^2 and a negative value of one of the roots.

If $n<n_0$, then $\left(\frac{P_n'(r)}{P(r)}\right)_r'<0$ — the ratio $\frac{P_n'(r)}{P(r)}$ is a decreasing function of the variable r. Then, for any change in interest rates, expressions (3) and (4) are positive, which means

$$\left(\frac{P(\tilde{r})-P(r)}{P(r)}\right)_n'>0 \text{ at } \tilde{r}<r$$

and $\left(\frac{P(r)-P(\tilde{r})}{P(r)}\right)_n'>0$ at $\tilde{r}>r$. Therefore,

Table 1

Dependence $\Delta P/P$ on *n* (*f* > *r*)
f = 10%, r = 8%, Δr = 0,1%, $\tilde{r} = 8{,}1\%$

n	$\Delta P/P$
1	0.000925
2	0.001767
3	0.002535
4	0.003237
5	0.003882
8	0.005525
10	0.006421
15	0.008152
20	0.009360
25	0.010215
30	0.010823
35	0.011258
40	0.011570
50	0.011952
60	0.012147
$\lim\limits_{n\to\infty} \frac{\Delta P}{P} = \frac{\Delta r}{\tilde{r}}$	0.012346

Table 2

Dependence $\Delta P/P$ on *n* (*f* < *r*)
f = 10%, r = 13%, Δr = 0,1%, $\tilde{r} = 13{,}1\%$

n	$\Delta P/P$
1	0.00088417
3	0.00240759
5	0.00363531
10	0.00568087
20	0.00725052
30	0.00758621
40	0.00763938
43	0.00764171
44	0.00764198
45	0.00764208
46	0.00764205
47	0.00764192
50	0.00764109
55	0.00763923
60	0.00763749
$\lim\limits_{n\to\infty} \frac{\Delta P}{P} = \frac{\Delta r}{\tilde{r}}$	0.00763359

Source: compiled by the author.

at maturity $n < n_0$ the interest rate risk of bonds sold at a discount increases with maturity.

If $n > n_0$, then $\left(\frac{P_n'(r)}{P(r)}\right)_r' > 0$ — the ratio $\frac{P_n'(r)}{P(r)}$ is an increasing function of the variable *r*. Then, for any change in interest rates, expressions (3) and (4) are negative, which means

$$\left(\frac{P(\tilde{r}) - P(r)}{P(r)}\right)_n' < 0 \text{ at } \tilde{r} < r$$

and $\left(\frac{P(r) - P(\tilde{r})}{P(r)}\right)_n' < 0$ at $\tilde{r} > r$. Therefore, at maturity $n > n_0$ the interest rate risk of bonds sold at a discount decreases as the maturity increases.

From the conditions $\left(\frac{\Delta P}{P}\right)_n' > 0$ at $n < n_0$ and $\left(\frac{\Delta P}{P}\right)_n' < 0$ at $n > n_0$ it follows that n_0 — is the term of maximum interest rate risk of bonds sold at a discount. Equating expression (8) to zero, we find the approximate value of the term n_0:

Table 3

Exact and approximate values of the term according to formulas (9), (10)
f = 10%, Δr = 0,1% ($f < r$)

r %	11	12	13	15	20	25
n_0 exact, $\tilde{r} < r$, (1)	127	66	47	30	18	13
n_0 exact, $\tilde{r} > r$, (2)	115	63	45	30	18	13
n_0 (9)	120	64.5	45.7	30.3	18.1	13.7
n_0 (10)	121.1	65.3	46.4	30.7	18.0	13.3

Source: compiled by the author.

$$n_0 \approx \frac{f(1+r)+\sqrt{f^2(1+r)^2+4(r-f)f(1+f)}}{2(r-f)f}. \quad (9)$$

The theorem has been proven.

The statements about the behavior of interest rate risk proved in the theorem are confirmed by calculations using formulas (1) and (2) of the relative changes in the prices of bonds sold at a premium or at a discount for various maturities. Examples of such calculations are shown in *Table 1* and *2*. It should be noted that in these tables, as n increases, the values $\Delta P/P$ approach their limiting values obtained by the formula from statement 1 of the theorem. *Table 2* shows that the term of the maximum value $\Delta P/P$ with an increase in yield by Δr = 0.1% of a bond with a coupon rate f = 10% and an initial yield r = 13% is $n_0 = 45$ years.

In [28], for bonds sold at a discount, the following expression was obtained for the approximate value of the term of the maximum value $\Delta P/P$:

$$n_0 \approx 1+\frac{(r+1)(2r-f)+\sqrt{f^2(r-1)^2+4r(r-f+rf)}}{2r(r-f)}. \quad (10)$$

Table 3 shows the exact and approximate values of the term n_0 for given values of the coupon rate f and the amount of change in the interest rate $\Delta r > 0$ for various values of the initial yield r. The exact values of n_0 are obtained from direct calculations using formulas (1) and (2) of the relative price changes of bonds sold at a discount for various maturities n. Approximate values of the term n_0 are obtained by formulas (9) and (10).

As we can see, the calculation results agree with each other. As the difference $(r-f)$ increases the exact and approximate values of the term n_0 converge. As follows from formulas (9), (10), and calculations, the the term of maximum of interest rate risk increases with decreasing difference $(r-f)$ and is practically absent at close values of the coupon rate and yield.

CONCLUSIONS

Macroeconomic interest rate risk factors such as inflation and the key interest rate are considered. The role of bond parameters in the level of interest risk is noted. Maturity is a significant risk factor for bond investments,

primarily because of the bond's exposure to interest rate risk. The proof obtained in the article of the dependence of the interest rate risk of bonds on term to maturity is confirmed by calculations and can be used to analyze the behavior of the interest rate risk of bonds in the context of a changing key interest rate.

The results are obtained under the condition that the market yield curve is horizontal and its movements are parallel. In reality, the yield curve is not horizontal, and its shifts are not necessarily parallel. However, it is known that the investment properties of bonds are studied under given conditions. The results of the study can be useful to the issuer and investor, as well as in theory when studying the investment properties of bonds.

REFERENCES

1. Fabozzi F. J. Bond markets, analysis and strategies. Harlow: Pearson Education Ltd.; 2012. 744 p. (Russ. ed.: Fabozzi F. J. Rynok obligatsii: Analiz i strategii. 2nd ed. Moscow: Alpina Business Books; 2007. 950 p.).
2. Rossokhin V. V. Analysis of risk factors in bond yields. Moscow: Internauka; 2019. 112 p. (In Russ.).
3. Buklemishev O. V. Eurobond market. Moscow: Delo; 1999. 232 p. (In Russ.).
4. Fabozzi F. J. Duration estimation for bonds and bond portfolios. In: Fabozzi F. J., ed. Handbook of finance (in 3 vols.). Vol. III. Valuation, financial modeling, and qualitative tools. Hoboken, NJ: John Wiley & Sons; 2008. URL: https://doi.org/10.1002/9780470404324.hof003015
5. Isaev A. K., Dem'yanov V. N. Analysis of factors affecting the yield of corporate bonds. In: Research and design of intelligent systems in automotive, aircraft and mechanical engineering (ISMCA' 2018). (Taganrog, Apr. 19–20, 2018). Proc. All-Russ. sci-pract. conf. with int. particip. Taganrog: ElDirect; 2018:136–140. (In Russ.).
6. Che-Yahya N., Abdul-Rahim R., Mohd-Rashid R. Determinants of corporate bond yield: The case of Malaysian Bond Market. *International Journal of Business and Society*. 2016;17(2):245–258.
7. Yanto D., Darmansyah D. Determination yield to maturity bonds, audit quality as moderators. *Jurnal Ecodemica: Jurnal Ekonomi, Manajemen, dan Bisnis*. 2021;5(1):1–14. DOI: 10.31294/jeco.v5i1.8604
8. Dhar S. Determinants of corporate bond's yields in economy. *SSRN Electronic Journal*. 2016. DOI: 10.2139/ssrn.2761308
9. Listiawati L. N., Paramita V. S. Pengaruh tingkat suku bunga, inflasi, Debt to equity ratio, dan ukuran perusahaan terhadap yield obligasi pada perusahaan yang terdaftar di bursa efek Indonesia tahun 2010–2016. *Jurnal Manajemen*. 2018;15(1):33–51. DOI: 10.25170/jm.v15i1.97
10. Simu N. Determinants of Indonesia corporate bond yield. *Business and Economic Horizons*. 2017;13(5):619–629. DOI: 10.15208/beh.2017.42
11. Sintami A. A., Marsoem B. S. Analysis of factors affecting yield to maturity of corporate bonds traded on Indonesia Stock Exchange 2016–2018. *International Journal of Innovative Science and Research Technology*. 2020;5(7):1443–1451. DOI: 10.38124/IJISRT20JUL733
12. Suryaningprang A., Suteja J. Analysis of interest rate, capital structure and information risk on yield to maturity and its application on company value: A case study of bond companies listed in the Indonesia Stock Exchange (IDX) from 2009 to 2013. In: The 1st Inter-University forum for strengthening academic competency (IFSAC). (Bandung, Oct. 2–3, 2018). Bandung: Research Institute of Universitas Pasundan Press; 2018:2–9.
13. Gitman L. J., Joehnk M. D. Fundamentals of investing. Reading, MA: Addison-Wesley; 1998. 720 p. (Russ. ed.: Gitman L. J., Joehnk M. D. Osnovy investirovaniya. Moscow: Delo; 1999. 1008 p.).
14. Yakimchuk A. Yu., Teplenko A. I., Konyagina M. N. The impact of key rates on inflation in modern Russia. *Vestnik Akademii znanii = Bulletin of the Academy of Knowledge*. 2020;(2):490–495. (In Russ.). DOI: 10.24411/2304–6139–2020–10217
15. Shalamov G. A., Ageeva N. A. Key bet of the Bank of Russia as a tool for regulating the inflation level. *Finansovaya ekonomika = Financial Economy*. 2019;(4):451–455. (In Russ.).

16. Balkoev I. M. Influence of monetary policy of the Central Bank on the pricing of financial assets. *Nauchnyi al'manakh = Science Almanac*. 2017;(10–1):33–36. (In Russ.). DOI: 10.17117/na.2017.10.01.033
17. Popova N. V. Market theorems and their continuation. *Vestnik Rossiiskogo ekonomicheskogo universiteta imeni G. V. Plekhanova = Vestnik of the Plekhanov Russian University of Economics*. 2013;(7):93–101. (In Russ.).
18. Mel'nikov A.V., Popova N. V., Skornyakova V. S. Mathematical methods of financial analysis. Moscow: Ankil; 2006. 440 p. (In Russ.).
19. Ajlouni M. M. Properties and limitations of duration as a measure of time structure of bond and interest rate risk. *International Journal of Economic Perspectives*. 2012;6(4):46–56.
20. Hsia C.-C., Weston J. F. Price behavior of deep discount bonds. *Journal of Banking & Finance*. 1981;5(3):357–361. DOI: 10.1016/0378–4266(81)90031–5
21. Ivanovski Z., Stojanovski T. D., Ivanovska N. Interest rate risk of bond prices on Macedonian Stock Exchange — Empirical test of the duration, modified duration and convexity and bonds valuation. *Economic Research*. 2013;26(3):47–62. DOI: 10.1080/1331677X.2013.11517621
22. Livingston M., Zhou L. Exponential duration: A more accurate estimation of interest rate risk. *The Journal of Financial Research*. 2005;28(3):343–361. DOI: 10.1111/j.1475–6803.2005.00128.x
23. Nia N. M., Alouj H. A., Pireivatlou A. S. Interest rate risk of zero-coupon bond prices on National Stock Exchange (NSE) — Empirical test of the duration, modified duration, convexity and immunization risk. *Archives des Sciences*. 2012;65(9):168–187.
24. Shirvani H., Wilbratte B. Duration and bond price volatility: Some further results. *Journal of Economics and Finance Education*. 2005;4(1):1–6. URL: https://www.economics-finance.org/jefe/fin/Wilbratepaper2.pdf
25. Hawawini G. A. On the mathematics of Macaulay's duration: A note. INSEAD Research Working Paper. 1981;(82/03). URL: https://flora.insead.edu/fichiersti_wp/inseadwp1982/82–03.pdf
26. Pianca P. Maximum duration of below par bonds: A closed-form formula. *SSRN Electronic Journal*. 2005. DOI: 10.2139/ssrn.738445
27. Popova N. V. Discussing Macaulay duration features. *Vestnik Finansovogo universiteta = Bulletin of the Financial University*. 2011;(1):42–46. (In Russ.).
28. Popova N. V. Maturity impact on bond price volatility. *Vestnik Finansovogo universiteta = Bulletin of the Financial University*. 2013;(3):72–84. (In Russ.).
29. Malkiel B. Expectations, bond prices, and the term structure of interest rates. *Quarterly Journal of Economics*. 1962;76(2):197–218.

ABOUT THE AUTHOR

Natal'ya V. Popova — Cand. Sci. (Phys.-Math.), Assoc. Prof., Department of Higher Mathematics in Plekhanov Russian University of Economics, Moscow, Russia
http://orcid.org/0000-0002-3700-5249
nat_popova_@mail.ru

Conflicts of Interest Statement: The author has no conflicts of interest to declare.

The article was submitted on 18.05.2022; revised on 31.05.2022 and accepted for publication on 01.06.2022.
The author read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-196-225
UDC 330.42(045),51–77(045)
JEL C1, C15, C4, C5, C53

An Algorithm for Restoring a Function from Different Functionals for Predicting Rare Events in the Economy

Yu.A. Korablev
Financial University, Russia, Moscow

ABSTRACT

This paper aims to restore some parameters of functionals using cubic splines to forecast rare events in finance and economics. The article considers the mathematical method for recovering an unknown function from many different functionals, such as the value of a function, the value of its first derivative, second derivative, as well as a definite integral over a certain interval. Moreover, all observations can occur with an error. Therefore, the author uses a method of recovering a function from different functionals observed with an error. The function is restored in the form of a cubic spline, which has a value-second derivative representation. The optimization problem consists in minimizing several sums of squares of the deviation at once, for ordinary values, for the first derivatives, for the second derivatives, for integrals, and for roughness penalty. For greater flexibility, weights have been introduced both for each group of observations and for each individual observation separately. The article shows in detail how the elements of each corresponding matrix are filled in. The appendix provides an implementation of the method as a FunctionalSmoothingSpline function in R language. Examples of using the method for the analysis and forecasting of rare (discrete) events in the economy are given. Formulas for calculating the cross-validation score $CV(\alpha)$ for the automatic procedure for determining the smoothing parameter α from the data in our problem of recovering a function by many functionals are shown. The paper concludes that the presented method makes it possible to analyze and predict rare events, which will allow you to prepare for such future events, get some benefit from this, or reduce possible risks or losses.

Keywords: rare events; forecast; event analysis; recovery by functionals; smoothing spline; roughness penalty; R; FunctionalSmoothingSpline; cross-validation

For citation: **Korablev Yu.A. An algorithm for restoring a function from different functionals for predicting rare events in the economy. *Finance: Theory and Practice.* 2022;26(3):196-225. DOI: 10.26794/2587-5671-2022-26-3-196-225**

1. INTRODUCTION

Rare events in the economy have a special interest. The rarity of events from the point of view of information theory makes such events more significant [1]. At the same time, events can have significant costly consequences and the ability to predict them is an urgent task. In logistics, for example, there is a lot of interest in predicting rare/intermittent demand forecast when demand occurs at a large number of intervals at which there was no demand. Of course, absolutely random events cannot be predicted, but if there is some pattern in the occurrence of events, then prediction is possible. Croston's [2] and Willemaine's [3] most popular methods, as well as many of their modifications, can find statistical patterns. Recent reviews [4, 5] analyse existing publications on such methods of forecasting intermittent demand. But in all methods only statistical analysis takes place to some extent. From the available data (intermittent demand) either the parameters of distributions or the values of transition probabilities of simple models of the Markov process. Such approaches, if they can give the expected number of events per time interval, are not able to give a forecast of the moment of occurrence of a particular event (estimation of probability of an event in the next time period, which is some statistical "shamanism").

The author, unlike other researchers, offers an approach [6] based on the consideration of internal processes leading to the occurrence of events. So, the same rare demand should be analyzed from the point of view of the process of consumption, occurring on the side of a client outside our control. It turns out that it will not be difficult to restore the speed (intensity) of consumption of products at a particular customer. Of course, preliminary data about events (purchases) should be divided into different samples depending on the sources (customers) where they are generated, which may not always be possible due to imperfections in the data collection methodology. Viewing the data in the sample as a sequence of integrals from the consumption rate, there is a recovery of this very unknown consumption rate function by means of functional recovery methods. This approach can be applied not only in the analysis of intermittent demand, but in any area where there are processes similar to devastation (capacity) or the accumulation of some disturbance to a certain level, after which it is reset to the initial level (this method the author called "capacitive method"). In the financial sphere, this approach can also be applied, for example, when the client periodically requests support.

In a similar way, one might wonder what other processes might exist in an economy that generates rare events. Note that we are not interested in completely random events, which in principle is impossible (or at this point is not possible) to predict, we are interested in the events that arise from some process that we could formulate and then reproduce ourselves.

Here it is worth remembering, and what is the randomness itself. Randomness is only a measure of uncertainty, a measure of ignorance, an abstraction introduced to compact all the many unknown factors to the researcher. Of course, because of the universal study of statistics and theory of probability in institutions of higher education and already in schools, the notion of randomness and probability became almost a physical phenomenon. At the same time, there is a philosophical concept of cosmic determinism, when everything in the universe is predestined and a wise man, knowing the position and velocity of all particles at one time, can predict their position at any other time. This concept is opposed by the Heisenberg uncertainty principle, where in quantum mechanics it is impossible to determine both the position and the momentum of particles at the same time. We will not argue with either concept and argue whether there is a true randomness or not. The above was only necessary to ensure that senior colleagues

who have worked for many decades in the field of statistical research look at events or observations not as a statistical sample, and instead they tried to go deep into each event and they thought, what is the cause of this event, what is the mechanism of its formation. They looked at streams of events not as random streams, but as the result of some mechanism of formation of these events, or even the set of mechanisms from which the events are mixed when the implementation of data collection fails.

As mentioned above, many economic events are related to consumption processes (or accumulation of impact). Since we know how they are formed, the streams of events are no longer random, this knowledge brings in the information we need and allows us to better analyze them. If you do not mix events from different sources, in the simplest case it is possible to restore the speed of consumption on the side of the client, then construct a pattern model and calculate the next moment of consumption. For this it is enough to look at the values of volumes of purchases as integral. However, in the more complex case, other available information is required. This study focuses on how to recover unknown process parameters if different types of data are available, such as values at a certain point in time, values of the first and second derivatives at certain times, values of certain integrals at certain periods of time. And these data may be available all at once, and one thing may be available, but the sample sizes of the input data of the different characteristics may not be the same (or are zero if such data are not available).

2. FUNCTIONAL RECOVERY

2.1. Optimization problem

Need to restore dynamically changing value of some process parameter from available data. Assume that this parameter changes continuously as some unknown function $f(t)$, i.e. $f(t)$ — desired function.

The following data are known:

$$y_i = f(t_i) + \varepsilon_i,\ i = 1,\ldots,n_f;$$

$$y_j^{'} = f'(t_j) + \varepsilon_j^{'},\ j = 1,\ldots,n_{df};$$

$$y_l^{''} = f''(t_l) + \varepsilon_l^{''},\ l = 1,\ldots,n_{d^2f};$$

$$Y_u = \int_{t_u^a}^{t_u^b} f(t)dt + \varepsilon_u^{\text{int}},\ u = 1,\ldots,n_{\text{int}},$$

where t_i, t_j, t_l — the moment of observation of the values of the unknown function $f(t)$, its first and second derivatives; t_u^a and t_u^b — lower and upper integration range of the appropriate integral monitoring; ε_i, $\varepsilon_j^{'}, \varepsilon_l^{''}, \varepsilon_u^{\text{int}}$ — errors of observation at values, first derivative, second derivative and integral respectively (with zero mathematical expectation, variances can be different); n_f, n_{df}, n_{d^2f}, n_{int} — correspondingly the number of observations of values, first derivatives, second derivatives, integrals of the desired function.

Of course, it will not be possible to accurately restore the original function, there are an infinite number of ways to chart the function so that the corresponding values of this function comply with predetermined values at the specified points. You can roughly restore the function by imposing certain restrictions. For example, we can say that we will recover bits of polynomials (splines) of a certain degree (third degree is the most common). Will impose restrictions on the flexibility (roughness) of the function.

This class of tasks is called collocation tasks, inverse tasks, approximation tasks (can be considered as synonyms in this context). It is considered (in the English-language literature) that the fundamental work of approximating the restoration of functions is the work G. Kimeldorf and G. Wahba [7]. To date, this work has been quoted 797 times in international citation databases. Studying the whole list of citing works and familiarization with more than 100 full-text publications from this list (based on annotations) showed the lack of a ready solution to the problem, which is

solved in this article (even as presented by the author). For some reason, a lot of work is limited to interpolation splines, when errors are not taken into account, only interpolation conditions are used in the form of exact equations. In some works, it is possible to find a unique solution only on the basis of one functionality, not many, and mainly again with the help of interpolation splines [8, 9].

The decision presented by the author is based on a well written work P. J. Green and B. W. Silverman [10] and is based on cubic smoothing splines, which have a representation in the form of values and second derivatives. To restore the function for many functions, we will minimize several squares of deviation and nonlinearity penalty at once

$$S(f)=\sum_{i=1}^{n_f}\left(y_i-f(t_i)\right)^2+\sum_{j=1}^{n_{df}}\left(y_j^{'}-f'(t_j)\right)^2+\sum_{l=1}^{n_{d^2f}}\left(y_l^{''}-f''(t_l)\right)^2 + \sum_{u=1}^{n_{\text{int}}}\left(Y_u-\int_{t_u^a}^{t_u^b}f(t)dt\right)^2+\alpha\int_{t_{start}}^{t_{end}}\left(f''(t)\right)^2dt, \quad (1)$$

where the last summand — nonlinear penalty; α — smoothing ratio (regularization); t_{start} and t_{end} — the boundary at which the function is restored.

Since the values in which the values, derivatives and integrals are measured can vary greatly, add the appropriate coefficients so that the weight of each group of observations can be increased. In addition, we can change the weight of each individual observation. As a result, the optimization task will take the following form:

$$S(f)=\sum_{i=1}^{n_f}w_i^f\left(y_i-f(t_i)\right)^2+\mu\sum_{j=1}^{n_{df}}w_j^{df}\left(y_j^{'}-f'(t_j)\right)^2+\nu\sum_{l=1}^{n_{d^2f}}w_l^{d^2f}\left(y_l^{''}-f''(t_l)\right)^2 + \psi\sum_{u=1}^{n_{\text{int}}}w_u^{\text{int}}\left(Y_u-\int_{t_u^a}^{t_u^b}f(t)dt\right)^2+\alpha\int_{t_{start}}^{t_{end}}\left(f''(t)\right)^2dt, \quad (2)$$

where w_i^f, w_j^{df}, $w_l^{d^2f}$, w_u^{int} — individual weights of the respective observation groups;

μ — weight of the whole group of observations of the first derivatives; ν — weight of the whole group of observations of second derivatives; ψ — weight of the whole integral group of observations. Note that there is no weight for the whole group of observations of normal values, i.e. it is assumed to be equal to one, and all other coefficients then show weight compared to this first group of observations. Note also that all individual weights could be adjusted proportionally to change the group weight, but this is not very convenient, so we will use both individual and group weights.

Next, we will restore the unknown function $f(t)$ as a cubic spline $g(t)$ (articulation of pieces of third-degree polynomials).

2.2. Spline types

Instead of the usual representation of polynomials with 4 unknown coefficients a_0, a_1, a_2, a_3 for each slice of spline, between nodes s_k and s_{k+1}

$$g(t)=a_0+a_1(t-s_k)+a_2(t-s_k)^2+a_3(t-s_k)^3,\ s_k\le t\le s_{k+1}$$

we will use more convenient performance through spline values $g_k=g(s_k)$ and its second derivative values $\gamma_k=g''(s_k)$ at the ends of each interval (value-second derivative representation) [10, p. 12, 22, 23]

$$g(t)=\frac{(t-s_k)g_{k+1}+(s_{k+1}-t)g_k}{s_{k+1}-s_k}$$
$$-\frac{1}{6}(t-s_k)(s_{k+1}-t)\left[\left(1+\frac{t-s_k}{s_{k+1}-s_k}\right)\gamma_{k+1}+\left(1+\frac{s_{k+1}-t}{s_{k+1}-s_k}\right)\gamma_k\right], \quad (3)$$
$$s_k \le t \le s_{k+1}.$$

As before, each piece of spline is identified by 4 unknown g_k, g_{k+1}, γ_k, γ_{k+1}, but since the end of one piece of spline is the beginning of the next piece, it is enough to define only two parameters g_k, γ_k for each nodes s_k (note that the parameters g_k, γ_k contain more physical meaning than parameters a_0, a_1, a_2, a_3). To define spline in all m nodes $s_1 < s_2 < \ldots < s_m$ (the number of nodes m is usually given a priori by the researcher) it is necessary to specify a vector of values $g=(g_1,\ldots,g_m)^T$ and the vector of second derivatives $\gamma=(\gamma_2,\ldots,\gamma_{m-1})^T$ (the second derivative at the spline ends turns to zero $\gamma_1=\gamma_m=0$ — natural spline conditions).

This form of spline ensures continuity $g(t)$ and its second derivative $g''(t)$ in the articulation points (spline nodes s_k). However, for the continuity of the first derivative at the points of articulation $g'(s_k-0)=g'(s_k+0)$ must be done a system $m-2$ equations

$$\frac{g_{k+1}-g_k}{s_{k+1}-s_k}-\frac{g_k-g_{k-1}}{s_k-s_{k-1}}=\frac{1}{6}(s_{k+1}-s_k)(\gamma_{k+1}+2\gamma_k)+\frac{1}{6}(s_k-s_{k-1})(2\gamma_k+\gamma_{k-1}), \quad (4)$$
$$k=2,\ldots,m-1;$$

or in matrix form

$$Q^T g = R\gamma\,, \quad (5)$$

where Q — tridiagonal matrix of coefficients at unknown g_k dimension $m\times(m-2)$ (column written); R — three-diagonal matrix of coefficients at unknown γ dimension $(m-2)\times(m-2)$ ($h_k=s_{k+1}-s_k$ node distance for $k=1,\ldots,m-1$).

Q	2	3	…	$m-1$
1	h_1^{-1}	0	…	0
2	$-h_1^{-1}-h_2^{-1}$	h_2^{-1}	…	0
3	h_2^{-1}	$-h_2^{-1}-h_3^{-1}$	…	0
4	0	h_3^{-1}	…	0
…	…	…	…	…
$m-2$	0	0	…	h_{m-2}^{-1}
$m-1$	0	0	…	$-h_{m-2}^{-1}-h_{m-1}^{-1}$
m	0	0	…	h_{m-1}^{-1}

R	2	3	4	…	$m-1$
2	$(h_1+h_2)/3$	$h_2/6$	0	…	0
3	$h_2/6$	$(h_2+h_3)/3$	$h_3/6$	…	0
4	0	$h_3/6$	$(h_3+h_4)/3$	…	0
5	0	0	$h_4/6$	…	0
…	…	…		…	…
$m-2$	0	0		…	$h_{m-2}/6$
$m-1$	0	0		…	$(h_{m-2}+h_{m-1})/3$

Instead of including the continuity conditions of the first derivative $Q^T g = R\gamma$ as a constraint system in the optimization problem, from this system of equations express one of the unknown, for example $\gamma = R^{-1}Q^T g$, replace it and solve the optimization problem with only one of the unknowns.

Penalty of smoothness (roughness) $\int_{s_1}^{s_m}\left(g''(t)\right)^2 dt$ is simplified to operations with the same matrices, see [10, p. 24–35]:

$$\int_{s_1}^{s_m}\left(g''(t)\right)^2 dt = \gamma^T Q^T g = \gamma^T R\gamma = g^T\left[QR^{-1}Q^T\right]g = g^T Kg, \tag{6}$$

where $K = QR^{-1}Q^T$ symmetric matrix by dimension $m\times m$.

2.3. Recovery

As a result, we have the following task — to determine the parameters of spline $g(t)$, minimizing

$$\begin{aligned} S(g) = &\sum_{i=1}^{n_f} w_i^f\left(y_i - g(t_i)\right)^2 + \mu\sum_{j=1}^{n_{df}} w_j^{df}\left(y_j' - g'(t_j)\right)^2 + \nu\sum_{l=1}^{n_{d^2f}} w_l^{d^2f}\left(y_l'' - g''(t_l)\right)^2 \\ &+ \psi\sum_{u=1}^{n_{\text{int}}} w_u^{\text{int}}\left(Y_u - \int_{t_u^a}^{t_u^b} g(t)\,dt\right)^2 + \alpha\int_{t_{start}}^{t_{end}}\left(g''(t)\right)^2 dt, \end{aligned} \tag{7}$$

where the spline $g(t)$ has a form (3).

To simplify the record, it is convenient to record the notation $h_k = s_{k+1} - s_k$, $h_k^{-i} = t_i - s_k$, $h_k^{+i} = s_{k+1} - t_i$. The record of related functionalities will be as follows:

$$g(t_i) = \frac{h_k^{-i}}{h_k}g_{k+1} + \frac{h_k^{+i}}{h_k}g_k - \frac{h_k^{-i}h_k^{+i}\left(h_k + h_k^{-i}\right)}{6h_k}\gamma_{k+1} - \frac{h_k^{-i}h_k^{+i}\left(h_k + h_k^{+i}\right)}{6h_k}\gamma_k, \tag{8}$$
$$k : s_k \le t_i < s_{k+1};$$

$$g'(t_j) = \frac{g_{k+1}}{h_k} - \frac{g_k}{h_k} - \left(\frac{h_k}{6} - \frac{\left(h_k^{-j}\right)^2}{2h_k}\right)\gamma_{k+1} + \left(\frac{h_k}{6} - \frac{\left(h_k^{+j}\right)^2}{2h_k}\right)\gamma_k, \tag{9}$$
$$k : s_k \le t_j < s_{k+1};$$

$$g''(t_l)=\frac{h_k^{-l}}{h_k}\gamma_{k+1}+\frac{h_k^{+l}}{h_k}\gamma_k,$$
$$k: s_k \le t_l < s_{k+1}; \tag{10}$$

$$\int_{t_u^a}^{t_u^b} g(t)dt=\sum_{l=0}^{L}\int_{s_{k+l}}^{s_{k+l+1}} g(t)dt-\int_{s_k}^{t_u^a} g(t)dt-\int_{t_u^b}^{s_{k+L+1}} g(t)dt=$$
$$L: s_{k+L} < t_u^b \le s_{k+L+1},$$
$$k: s_k \le t_u^a < s_{k+1}. \tag{11}$$

$$=\sum_{l=0}^{L}\left[\frac{h_{k+l}}{2}g_{k+l+1}+\frac{h_{k+l}}{2}g_{k+l}-\frac{h_{k+l}^3}{24}\gamma_{k+l+1}-\frac{h_{k+l}^3}{24}\gamma_{k+l}\right]$$
$$-\frac{\left(h_k^{-a}\right)^2}{2h_k}g_{k+1}-\frac{h_k^2-\left(h_k^{+a}\right)^2}{2h_k}g_k-\frac{\left(h_k^{-a}\right)^2\left(\left(h_k^{-a}\right)^2-2h_k^2\right)}{24h_k}\gamma_{k+1}+\frac{\left(h_k^{-a}\right)^2\left(h_k^{+a}+h_k\right)^2}{24h_k}\gamma_k$$
$$-\frac{h_{k+L}^2-\left(h_{k+L}^{-b}\right)^2}{2h_{k+L}}g_{k+L+1}-\frac{\left(h_{k+L}^{+b}\right)^2}{2h_{k+L}}g_{k+L} \tag{12}$$
$$+\frac{\left(h_{k+L}^{+b}\right)^2\left(h_{k+L}^{-b}+h_{k+L}\right)^2}{24h_{k+L}}\gamma_{k+L+1}-\frac{\left(h_{k+L}^{+b}\right)^2\left(\left(h_{k+L}^{+b}\right)^2-2h_{k+L}^2\right)}{24h_{k+L}}\gamma_{k+L},$$
$$L: s_{k+L} < t_u^b \le s_{k+L+1},$$
$$k: s_k \le t_u^a < s_{k+1},$$

where in the last expression $h_k^{-a}=t_u^a-s_k$, $h_k^{+a}=s_{k+1}-t_u^a$, $h_k=s_{k+1}-s_k$, $h_{k+L}^{-b}=t_u^b-s_{k+L}$, $h_{k+L}^{+b}=s_{k+L+1}-t_u^b$, $h_{k+L}=s_{k+L+1}-s_{k+L}$.

In all these expressions, at the beginning it is determined at what interval k was observed. In the most recent expression for the integral, it is necessary to define at the beginning, at what interval k dropped out the lower limit of integration t_u^a and at what interval $k+L$ has dropped out the limit of integration t_u^b, where L — number of intervals between them (L can be equal to 0 if both are on the same interval).

All these expressions have a linear form relative to unknown spline parameters g_k and γ_k. Therefore, the optimization problem (7) can be expressed in the following matrix form:

$$S(g)=\left(Y_f-V_f g+P_f\gamma\right)^T W_f\left(Y_f-V_f g+P_f\gamma\right)+$$
$$+\mu\left(Y_{df}-V_{df}g+P_{df}\gamma\right)^T W_{df}\left(Y_{df}-V_{df}g+P_{df}\gamma\right)+$$
$$+\nu\left(Y_{d^2f}-0g+P_{d^2f}\gamma\right)^T W_{d^2f}\left(Y_{d^2f}-0g+P_{d^2f}\gamma\right)+ \tag{13}$$
$$+\psi\left(Y_{\text{int}}-V_{\text{int}}g+P_{\text{int}}\gamma\right)^T W_{\text{int}}\left(Y_{\text{int}}-V_{\text{int}}g+P_{\text{int}}\gamma\right)+$$
$$+\alpha g^T Kg\to\min,$$

where Y_f, Y_{df}, Y_{d^2f}, Y_{int} — observation column; matrix V_f, V_{df}, V_{int} — coefficient matrices at unknown g_k; P_f, P_{df}, P_{d^2f}, P_{int} — coefficient matrices at unknown γ_k; W_f, W_{df}, W_{d^2f}, W_{int} — diagonal weight matrices.

V_f dimension $n_f \times m$, and each i-line looks like

1	…	$k-1$	k	$k+1$	$k+2$	…	m
0	…	0	h_k^{+i}/h_k	h_k^{-i}/h_k	0	…	0

P_f dimension $n_f \times (m-2)$, and each i-line looks like

2	…	$k-1$	k	$k+1$	$k+2$	…	$m-1$
0	…	0	$h_k^{-i}h_k^{+i}\left(h_k+h_k^{+i}\right)/6h_k$	$h_k^{-i}h_k^{+i}\left(h_k+h_k^{-i}\right)/6h_k$	0	…	0

V_{df} dimension $n_{df} \times m$, and each j-line looks like

1	…	$k-1$	k	$k+1$	$k+2$	…	m
0	…	0	$-1/h_k$	$1/h_k$	0	…	0

P_{df} dimension $n_{df} \times (m-2)$, and each j-line looks like

2	…	$k-1$	k	$k+1$	$k+2$	…	$m-1$
0	…	0	$-h_k/6+\left(h_k^{+j}\right)^2/2h_k$	$h_k/6+\left(h_k^{-j}\right)^2/2h_k$	0	…	0

P_{d^2f} dimension $n_{d^2f} \times (m-2)$, and each l-line looks like

2	…	$k-1$	k	$k+1$	$k+2$	…	$m-1$
0	…	0	$-h_k^{+l}/h_k$	$-h_k^{-l}/h_k$	0	…	0

V_{int} dimension $n_{int} \times m$, and each u-line is filled in as follows:

$$V_{u,k}=\frac{\left(h_k^{+a}\right)^2}{2h_k};V_{u,k+l}=\frac{h_{k+l-1}+h_{k+l}}{2},l=1,\ldots,L;V_{u,k+1}=V_{u,k+1}-\frac{\left(h_k^{-a}\right)^2}{2h_k};$$
$$V_{u,k+L}=V_{u,k+L}-\frac{\left(h_{k+L}^{+b}\right)^2}{2h_{k+L}};V_{u,k+L+1}=\frac{\left(h_{k+L}^{-b}\right)^2}{2h_{k+L}}. \quad (14)$$

Note that depending on L, some expressions may change twice (for example, if $L=0$, that k-item is changed by two expressions $V_{i,k}$ and $V_{i,k+L}$). In the case $L>2$ line u will be

$k-1$	k	$k+1$	$k+2$	$\ldots k+l \ldots$	$k+L$	$k+L+1$	$k+L+2$
0	$\frac{\left(h_k^{+a}\right)^2}{2h_k}$	$\left(\frac{h_k+h_{k+1}}{2}-\frac{\left(h_k^{-a}\right)^2}{2h_k}\right)$	$\frac{h_{k+1}+h_{k+2}}{2}$	$\frac{h_{k+l-1}+h_{k+l}}{2}$	$\left(\frac{h_{k+L-1}+h_{k+L}}{2}-\frac{\left(h_{k+L}^{+b}\right)^2}{2h_{k+L}}\right)$	$\frac{\left(h_{k+L}^{-b}\right)^2}{2h_{k+L}}.$	0

In the case $L=0$ line u will be

$k-1$	k	$k+1$	$k+2$
0	$\begin{pmatrix} \frac{\left(h_k^{+a}\right)^2}{2h_k} \\ -\frac{\left(h_{k+L}^{+b}\right)^2}{2h_{k+L}} \end{pmatrix}$	$\begin{pmatrix} \frac{\left(h_{k+L}^{-b}\right)^2}{2h_{k+L}} \\ -\frac{\left(h_k^{-a}\right)^2}{2h_k} \end{pmatrix}$	0

P_{int} dimension $n_{int}\times(m-2)$, each u-line is filled in as follows:

$$P_{u,k}=\frac{h_k^3}{24}-\frac{\left(h_k^{-a}\right)^2\left(h_k^{+a}+h_k\right)^2}{24h_k};P_{u,k+l}=\frac{h_{k+l-1}^3+h_{k+l}^3}{24},l=1,\ldots,L;$$
$$P_{u,k+1}=P_{u,k+1}+\frac{\left(h_k^{-a}\right)^2\left(\left(h_k^{-a}\right)^2-2h_k^2\right)}{24h_k};P_{u,k+L}=P_{u,k+L}+\frac{\left(h_{k+L}^{+b}\right)^2\left(\left(h_{k+L}^{+b}\right)^2-2h_{k+L}^2\right)}{24h_{k+L}};\quad(15)$$
$$P_{u,k+L+1}=\frac{h_{k+L}^3}{24}-\frac{\left(h_{k+L}^{+b}\right)^2\left(h_{k+L}^{-b}+h_{k+L}\right)^2}{24h_{k+L}}.$$

Here too, depending on L, some values may change several times. In the case $L>2$ line u will be

$k-1$	k	$k+1$	$k+2$
0	$\begin{pmatrix} \frac{h_k^3}{24}- \\ \frac{\left(h_k^{-a}\right)^2\left(h_k^{+a}+h_k\right)^2}{24h_k} \end{pmatrix}$	$\begin{pmatrix} \frac{h_k^3+h_{k+1}^3}{24}+ \\ \frac{\left(h_k^{-a}\right)^2\left(\left(h_k^{-a}\right)^2-2h_k^2\right)}{24h_k} \end{pmatrix}$	$\frac{h_{k+1}^3+h_{k+2}^3}{24}$
$\ldots k+l\ldots$	$k+L$	$k+L+1$	$k+L+2$
$\frac{h_{k+l-1}^3+h_{k+l}^3}{24}$	$\begin{pmatrix} \frac{h_{k+L-1}^3+h_{k+L}^3}{24}+ \\ \frac{\left(h_{k+L}^{+b}\right)^2\left(\left(h_{k+L}^{+b}\right)^2-2h_{k+L}^2\right)}{24h_{k+L}} \end{pmatrix}$	$\begin{pmatrix} \frac{h_{k+L}^3}{24}- \\ \frac{\left(h_{k+L}^{+b}\right)^2\left(h_{k+L}^{-b}+h_{k+L}\right)^2}{24h_{k+L}} \end{pmatrix}$	0

Then, due to the continuity conditions of the first derivative, (5) $Q^Tg=R\gamma$, that can express $\gamma=R^{-1}Q^Tg$, optimization task (13) can be written more compactly only through one unknown g:

$$S(g)=\left(Y_f-C_fg\right)^TW_f\left(Y_f-C_fg\right)+\mu\left(Y_{df}-C_{df}g\right)^TW_{df}\left(Y_{df}-C_{df}g\right)+$$
$$+\nu\left(Y_{d^2f}-C_{d^2f}g\right)^TW_{d^2f}\left(Y_{d^2f}-C_{d^2f}g\right)+\quad(16)$$
$$+\psi\left(Y_{\text{int}}-C_{\text{int}}g\right)^TW_{\text{int}}\left(Y_{\text{int}}-C_{\text{int}}g\right)+\alpha g^TKg\to\min,$$

where $C_f=V_f-P_fR^{-1}Q^T$, $C_{df}=V_{df}-P_{df}R^{-1}Q^T$, $C_{d^2f}=0-P_{d^2f}R^{-1}Q^T$, $C_{int}=V_{int}-P_{int}R^{-1}Q^T$.

Finally find the column of unknown parameters g, setting the derivatives equal from $S(g)$ to g (rules for taking matrix derivatives

$\frac{d\left(x^T b\right)}{dx} = b, \frac{d\left(bx\right)}{dx} = b^T, \frac{d\left(x^T Ax\right)}{dx} = \left(A + A^T\right)x = 2Ax$, the latter is true if A — symmetric matrix).

The expression for the parameters g is as follows:

$$g = \left(C_f^T W_f C_f + \mu C_{df}^T W_{df} C_{df} + \nu C_{d^2 f}^T W_{d^2 f} C_{d^2 f} + \psi C_{\text{int}}^T W_{\text{int}} C_{\text{int}} + \alpha K\right)^{-1} \times \\ \times \left(C_f^T W_f Y_f + \mu C_{df}^T W_{df} Y_{df} + \nu C_{d^2 f}^T W_{d^2 f} Y_{d^2 f} + \psi C_{\text{int}}^T W_{\text{int}} Y_{\text{int}}\right). \tag{17}$$

Knowing g, is calculated $\gamma = R^{-1} Q^T g$, after which you can build spline $g(t)$ at any point t on the (3).

2.4. Select a smoothing option

In all similar function recovery (smoothing) tasks, the selection of the smoothing parameter is discussed separately α. The procedure for automatically selecting this option for recovery tasks could not be found across functions. Classic cross-validation or L-curve procedures will not work. No other work to automatically select the smoothing parameter for many different functionalities. The author had to independently obtain modified formulas to estimate cross-validation.

Recall that the basic idea of cross-validation is to choose such an anti-aliasing parameter α, to recover a function $g(t,\alpha)$ was effective in predicting, i.e. that the function has the least variance when predicting the following values. Therefore, to calculate cross-validation exclude one observation, build spline $g^{(-i)}(t,\alpha)$ and observe the quadratic error this excluded observation is determined, and so do all observations. As a result, the cross validation estimate gives some estimate of the variance of observations as if they were predicted for spline sampled with the element-by-element exception of these observations. Formulas for calculating this estimate are as follows (without intermediate calculations):

$$CV(\alpha) = n_f^{-1} \sum_{i=1}^{n_f} w_i^f \left(y_i - g^{(-i)}(t_i,\alpha)\right)^2 + n_{df}^{-1} \mu \sum_{j=1}^{n_{df}} w_j^{df} \left(y_j' - g'^{(-j)}(t_j,\alpha)\right)^2 + \\ + n_{d^2 f}^{-1} \nu \sum_{l=1}^{n_{d^2 f}} w_l^{d^2 f} \left(y_l'' - g''^{(-l)}(t_l,\alpha)\right)^2 + n_{\text{int}}^{-1} \psi \sum_{u=1}^{n_{\text{int}}} w_u^{\text{int}} \left(Y_u - \int_{t_u^a}^{t_u^b} g^{-u}(t,\alpha)\,dt\right)^2 = \tag{18}$$

$$= n_f^{-1} \sum_{i=1}^{n_f} w_i^f \left(\frac{y_i - g(t_i,\alpha)}{1 - \sum_{k=1}^{m} C_{ik}^f A_{ki}^f(\alpha)}\right)^2 + n_{df}^{-1} \sum_{j=1}^{n_{df}} w_j^{df} \left(\frac{y_j' - g'(t_j,\alpha)}{1 - \sum_{k=1}^{m} C_{jk}^{df} A_{kj}^{df}(\alpha)}\right)^2 + \\ + n_{d^2 f}^{-1} \sum_{l=1}^{n_{d^2 f}} w_l^{d^2 f} \left(\frac{y_l'' - g''(t_l,\alpha)}{1 - \sum_{k=1}^{m} C_{lk}^{d^2 f} A_{kl}^{d^2 f}(\alpha)}\right)^2 + n_{\text{int}}^{-1} \sum_{u=1}^{n_{\text{int}}} w_u^{\text{int}} \left(\frac{Y_u - \int_{t_u^a}^{t_u^b} g(t,\alpha)\,dt}{1 - \sum_{k=1}^{m} C_{uk}^{\text{int}} A_{ku}^{\text{int}}(\alpha)}\right)^2 = \tag{19}$$

$$= n_f^{-1} \sum_{i=1}^{n_f} w_i^f \left(\frac{y_i - \sum_{k=1}^{m} C_{ik}^f g_k}{1 - \sum_{k=1}^{m} C_{ik}^f A_{ki}^f(\alpha)}\right)^2 + n_{df}^{-1} \sum_{j=1}^{n_{df}} w_j^{df} \left(\frac{y_j' - \sum_{k=1}^{m} C_{jk}^{df} g_k}{1 - \sum_{k=1}^{m} C_{jk}^{df} A_{kj}^{df}(\alpha)}\right)^2 + \\ + n_{d^2 f}^{-1} \sum_{l=1}^{n_{d^2 f}} w_l^{d^2 f} \left(\frac{y_l'' - \sum_{k=1}^{m} C_{lk}^{d^2 f} g_k}{1 - \sum_{k=1}^{m} C_{lk}^{d^2 f} A_{kl}^{d^2 f}(\alpha)}\right)^2 + n_{\text{int}}^{-1} \sum_{u=1}^{n_{\text{int}}} w_u^{\text{int}} \left(\frac{Y_u - \sum_{k=1}^{m} C_{uk}^{\text{int}} g_k}{1 - \sum_{k=1}^{m} C_{uk}^{\text{int}} A_{ku}^{\text{int}}(\alpha)}\right)^2. \tag{20}$$

Here the top indexes $(-i),(-j),(-l),(-u)$ mean, that spline $g(t,\alpha)$ estimated from data without this specific observation. Matrices $C^f, C^{df}, C^{d^2f}, C^{int}$ — same as in formulas (16), (17). Matrices $A^f(\alpha)=A(\alpha)C_f^TW_f$, $A^{df}(\alpha)=A(\alpha)\mu C_{df}^TW_{df}$, $A^{d^2f}(\alpha)=A(\alpha)\nu C_{d^2f}^TW_{d^2f}$, $A^{\text{int}}(\alpha)=A(\alpha)\psi C_{\text{int}}^TW_{\text{int}}$, where $A(\alpha)=\left(C_f^TW_fC_f+\mu C_{df}^TW_{df}C_{df}+\nu C_{d^2f}^TW_{d^2f}C_{d^2f}+\psi C_{\text{int}}^TW_{\text{int}}C_{\text{int}}+\alpha K\right)^{-1}$. Minimization $CV(\alpha)$ any known manner relative to α gives the desired smoothing parameter value α.

Cross validation works well on conventional splines. However, we have an unusual case, we are recovering also on the first/second derivatives and integrals. As a result, the cross-validation assessment shows not the usual variance of observations, but the variance of observations of both values and derivatives with integral (as you probably understand, mixing of different variances occurs, but in the construction of the spline we also mixed squares of errors of different observations). But the problem is that if you exclude some observations, you can dramatically change the type of the restored function. In the following examples (especially in the second one) it will be obvious that the exclusion of any observation will result in the function being very imprecise, with the forecast error calculated for the excluded observation. When each observation brings much needed information to restore the function, excluding the survey from the sample will result in very, very large errors. In this case cross-validation is not suitable. But if we have very many observations, some observations only statistically supplement the information of other observations, and the exclusion of one of the observations does not lead to significant changes in the restored function, cross validation method can be a good solution to determine the smoothing parameter.

3. R-LANGUAGE SOFTWARE IMPLEMENTATION

The described function recovery method is implemented in R for many different functional as a function FunctionalSmoothingSpline (see *Appendix 1*). Existing ready-made features and packages in R or other languages implementing similar capabilities could not be found.

4. USE CASE TO PREDICT RARE EVENTS IN AN ECONOMY

4.1. Forecasting future customer purchases

Let there be some customer who buys products from us (for example, an ordinary buyer buys in some trading network, or some wholesaler buys products from the manufacturer). We know nothing about the client, except the dates and volumes of his purchases. For simplicity, we model the replenishment process as in classic stock management models. Let the purchase data will be as follows (*table. 1*).

Table 1

Purchases data of a client not controlled by us

Date t_i	Volume y_i	Date t_i	Volume y_i	Date t_i	Volume y_i	Date t_i	Volume y_i	Date t_i	Volume y_i
03.01.2020	2170	06.06.2020	1976	09.10.2020	2093	10.04.2021	2257	07.08.2021	1968
25.01.2020	2281	19.06.2020	2205	24.10.2020	2141	28.04.2021	2189	19.08.2021	2136
22.02.2020	2242	03.07.2020	2096	12.11.2020	2273	12.05.2021	2026	02.09.2021	2145
11.03.2020	2206	17.07.2020	2125	10.12.2020	2217	24.05.2021	2072	20.09.2021	2235
26.03.2020	2142	29.07.2020	2034	31.12.2020	2218	04.06.2021	1983	07.10.2021	2186
12.04.2020	2210	09.08.2020	1980	21.01.2021	2252	16.06.2021	2059	22.10.2021	2141
30.04.2020	2215	21.08.2020	2098	18.02.2021	2211	30.06.2021	2146	09.11.2021	2256
14.05.2020	2102	05.09.2020	2222	09.03.2021	2218	14.07.2021	2082	07.12.2021	2264
26.05.2020	2115	23.09.2020	2191	24.03.2021	2137	27.07.2021	2177	29.12.2021	2241

Source: compiled by the author.

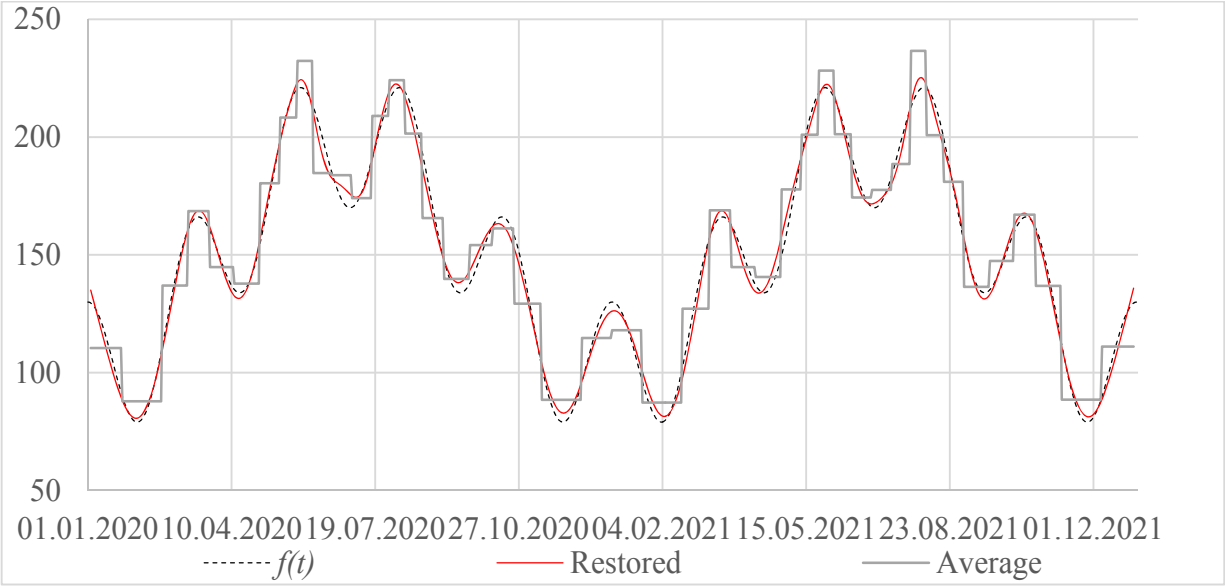


Fig. 1. **The restored intensity of product consumption on the side of the uncontrolled customer**

Source: compiled by the author.

Note: solid line – restored consumption intensity; dashed line – original consumption rate; stepped line – average intensity $y_i / (t_{i+1} - t_i)$ below each step means the purchase volume y_i.

Perceiving this data as a sequence of integrals from an unknown consumption intensity function from the current event to the next, reconstruct this unknown function using the mathematical method described above (*Fig. 1*). *Appendix 2* presents the implementation in R, shows how to prepare the data from the csv file and how to get the function FunctionalSmoothingSpline. For the current data set, the unknown consumption intensity function of the client beyond our control was restored very well.

The next step is to determine the pattern in the identified function, construct the model and extrapolate. Any known method may be used. The very dependence of the restored function can only be built on time, and it is possible to determine the dependence on some known external factors. Here the complete freedom and responsibility of the researcher is assumed. In our case, we will build a model and extrapolate values for the future as the sum of harmonic functions (harmonic function was laid down in the data modelling). Quinn-Fernandes algorithm well-suited to this [11, 12], which represents the sum of a limited number of harmonic functions. The result of this extrapolation is presented in *Fig. 2*. Since the restored function contained some deviations from the true function assumed in the simulation, the parameters of the extrapolated function were determined with some inaccuracy, as a result of which, at some sites, the extrapolated function is noticeably different from the true

The final step is to start the process of determining future events itself. In our case, the consumption process is restarted as in inventory management systems, where extrapolated values are involved as a consumption function. Starting with the most recent observation, it is possible to determine when the stock will run out and thus to give a forecast of the next circulation. If the data determine a further maximum reserve (or build a model for the volume of purchases), it will be possible to predict the entire sequence of future events. *Fig. 2* shows a prediction of future events on the horizontal axis with crosses. And the first few events are determined with an error of only 0–2 days, further the error increases to 6 days, but then again reduced to 2 days (and this is with the time between events 15–28 days). That is, the first events can be defined very well, but when you increase the planning horizon the

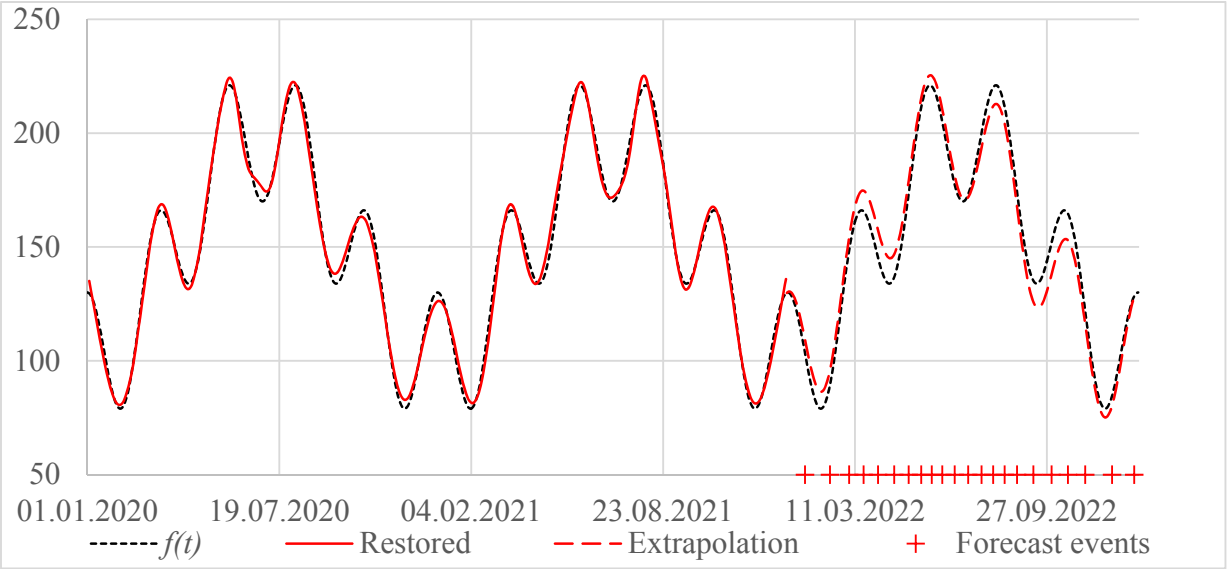


Fig. 2. **Extrapolation of the restored function and forecast of future events**

Source: compiled by the author.

accuracy drops (in this example then improves again), which is logical.

4.2. Determination of the hidden dynamics of the company on communications containing mainly qualitative information

Speculating what other processes of formation of events are in the economy, first of all come to mind those that can be somehow reduced to processes similar to the processes of consumption or the accumulation of disturbances (the depletion of stocks or the accumulation of disturbances to a critical level). Just a reminder that we don't care about random events, where there's no pattern, there's nothing to find. Try to think about what else you think about, whose education process is different from processes of consumption or accumulation of impact. Not immediately, but after a long period of reflection, it was possible to formulate the process of producing rare events in the economy, different from those mentioned, the following will be about one such process. It is worth saying that the process under consideration demonstrates more the possibility of applying the proposed approach to the analysis of rare events than to solve specific practical problems, although there will certainly be (we are interested in the possibility of analyzing rare events formed by processes other than consumption processes).

Let there be some organization beyond our control, from which at certain discrete moments of time there are signals, carrying mainly qualitative (rather than quantitative) information. And in each signal, this qualitative information may relate to different characteristics (not necessarily that each signal reports all the characteristics). For example, imagine a situation where the following signals come from a certain company:

a) We started to sink — we have to do something now!
b) We are still sinking heavily!
c) The decline seems to be slowing down.
d) We started to surface.
e) It's good to go, if we always go like this.
f) Oh no, we're sinking again!
g) We have passed the point of no return...
h) We're at the bottom.

What can be learned from this set of signals? It is possible to guess that events are formed within such an organization by observing some internal indicator that changes dynamically over time (possibly due to some controlling effects). The occurrence of events is related to the operation of comparing this indicator with some critical

values. If we use the proposed method, restore the dynamics of the non-observed internal indicator of the organization, construct a model of change of this dynamic, extrapolate and restart the process of formation of events, it is possible to predict future events (and perhaps even controlling influences).

What do we know from this type of signal that will help to restore the dynamics of the change of the non-observed indicator directly. It turns out that the following data can be extracted from the available signals: moments of time as points of receipt (possibly adjusted for delay in receipt); values of first derivative or position of extremity points; values of second derivative or position of inflection points; values of the most variable in some points. For greater completeness, also assume that integrals from the function of changing the same unobserved indicator can be observed (not included in the above example, but it can be assumed that if the indicator of interest indicates funds in the accounts, the accrued interest on these accounts over some time periods will be these integrals). It should be noted that all observations may be inaccurate, and both the observations themselves and the significance of those observations may be inaccurate. However, the error in the timing of the observation can be reduced in one way or another to an error in the values of the same observations (for example, if the true position of the extremum is slightly shifted from the observation point at which the value of the derivative is shifted from zero).

Denote the unknown search function of hidden dynamics as $f(t)$, which in the future will be restored according to available data in the form of spline. For simplicity let function $f(t)$ will be dimensionless and at the beginning of times t_0 the value of this function will be assumed as the reference value, and all other values will be expressed as a percentage of this reference value (i.e. assume that $f(t_0)=100$). Next, for example, let the original unknown function behave as shown in *Fig. 3*.

Initially it was said that we are dealing with primarily qualitative data, but this qualitative data will not be difficult to give a

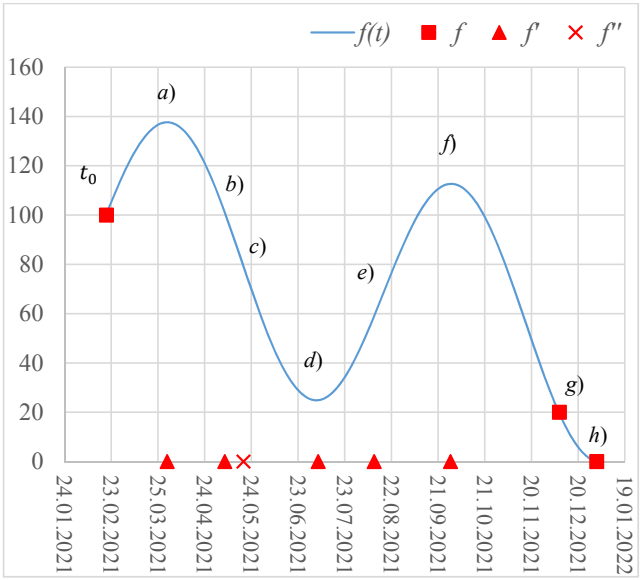


Fig. 3. **An example of the initial function $f(t)$ and the initial data**

Source: compiled by the author.

Note: Function $f(t)$ is dimensionless, shows the number (%) of the initial level (the value at the moment t_0 is taken as the initial). The function value data (events) are marked on the graph itself, the derivatives data are plotted on the horizontal axis and indicate their position.

quantitative estimate, and it is enough to give an approximate estimate (or the values could be initially approximated). So, in events *a*, *c*, *d*, *f* say that there are zero values of the first or second derivatives. For events *b* and *e* suppose we know the tangent of the angle of inclination, for event *g* suppose (or know) that the non-return rate starts at 20%. We gave an approximate estimate of values, and moments of time of events are observed and known. As a result, the available data may be as follows (*Table 2*).

Notice, the sample size can be very different, for the second derivative we have only one observation. Implementing the described function recovery method by functional (all coefficients μ, ν, ψ, α are equal 1), from the available data we will get the following result (*Fig. 4*). Code in R is presented in *Appendix 1*. The task proved poorly conditioned, there was insufficient information for a good function recovery (it was not specified that the function should have increased in the beginning).

If you add more information, for example, as integrals of the desired function, (*Fig. 5*),

Table 2

Available approximate data

t_f	y_f	t_{df}	y_{df}	t_{d^2f}	y_{d^2f}
20.02.2021	100	31.03.2021	0	19.05.2021	0
08.12.2021	20	07.05.2021	–1.75	–	–
01.01.2022	0	06.07.2021	0	–	–
–	–	11.08.2021	1.55	–	–
–	–	29.09.2021	0	–	–

Source: compiled by the author.

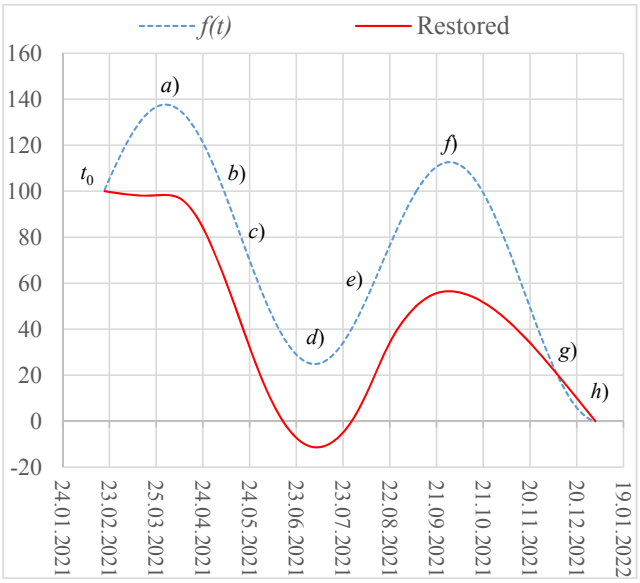


Fig. 4. **Poorly conditioned task. There was not enough information**

Source: compiled by the author.

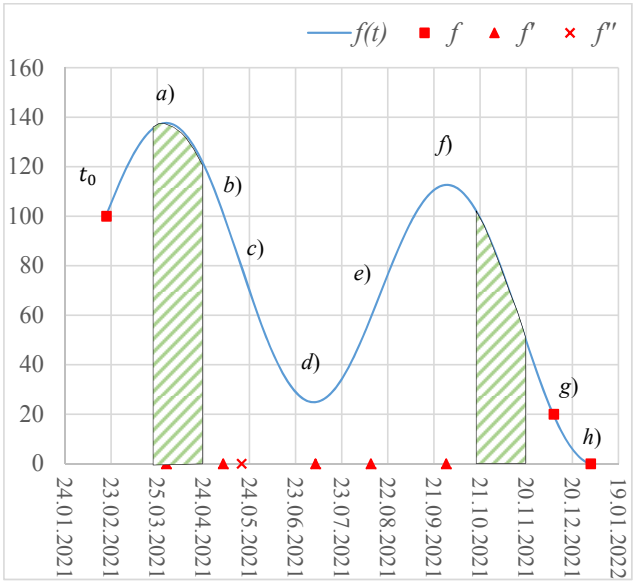


Fig. 5. **The original function $f(t)$ and the initial data with the additional information about the integrals**

Source: compiled by the author.

(*Table. 3*), then restore the original function gets much more accurate (*Fig. 6*).

Of course, recovery is still not perfect, but it is due to a lack of information, not a lack of mathematical method. If you add more observations, you can get a well recovery. Note that by adjusting the weights of both observations and observation groups, it is possible to adjust very flexibly which information should be paid more attention and which less.

Further forecasting of future events follows the same pattern as in the previous example (in this example we may not expect any future events).

5. AN APPROACH SCHEME FOR PREDICTING RARE EVENTS

Once again, separately, we give the idea of the approach of predicting rare events. If from rare events it is possible to restore the parameters of the process (dynamic), then further actions to predict rare events will be to identify the regularity of changes in these parameters of the process (their dynamics) over time. And to do so, all possible information should be used, patterns can be determined depending on changes in external observable factors, such as GDP, inflation, unemployment and other factors. Any existing mathematical methods can be used for this purpose. The purpose of

Table 3

Data with added information about integrals

t_f	y_f	t_{df}	y_{df}	t_{d^2f}	y_{d^2f}	t^a_{int}	t^b_{int}	Y_{int}
20.02.2021	100	31.03.2021	0	19.05.2021	0	25.03.2021	24.04.2021	4000
08.12.2021	20	07.05.2021	–1,75	–	–	21.10.2021	20.11.2021	2282
01.01.2022	0	06.07.2021	0	–	–	–	–	–
–	–	11.08.2021	1,55	–	–	–	–	–
–	–	29.09.2021	0	–	–	–	–	–

Source: compiled by the author.

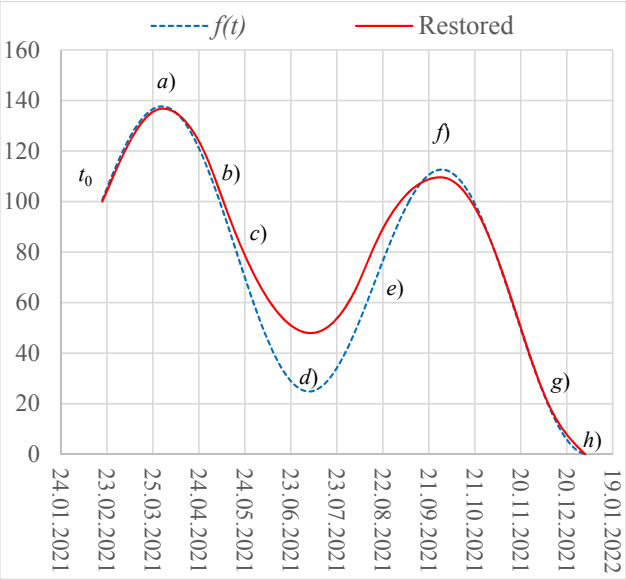


Fig. 6. **Restoration of a function from data with the additional information about integrals**

Source: compiled by the author.

this step will be to construct an appropriate model of changing the internal parameters of the process. The next step is to extrapolate process parameters for the future. If previously a model of dependence on external factors is built, it is necessary to extrapolate the values of external factors beforehand. Any known mathematical method can be used again for the extrapolation phase. Finally, if we have an idea of how the parameters of the process of making events will change in the future, and we can reproduce the mechanism of this process, it will be easy to get a forecast of future events by running the process itself with set parameters. The scheme of obtaining a forecast of future events is presented at *Fig. 7*.

The proposed approach cannot be directly compared with other methods. The fact is that none of the existing methods can predict when a future event will occur.

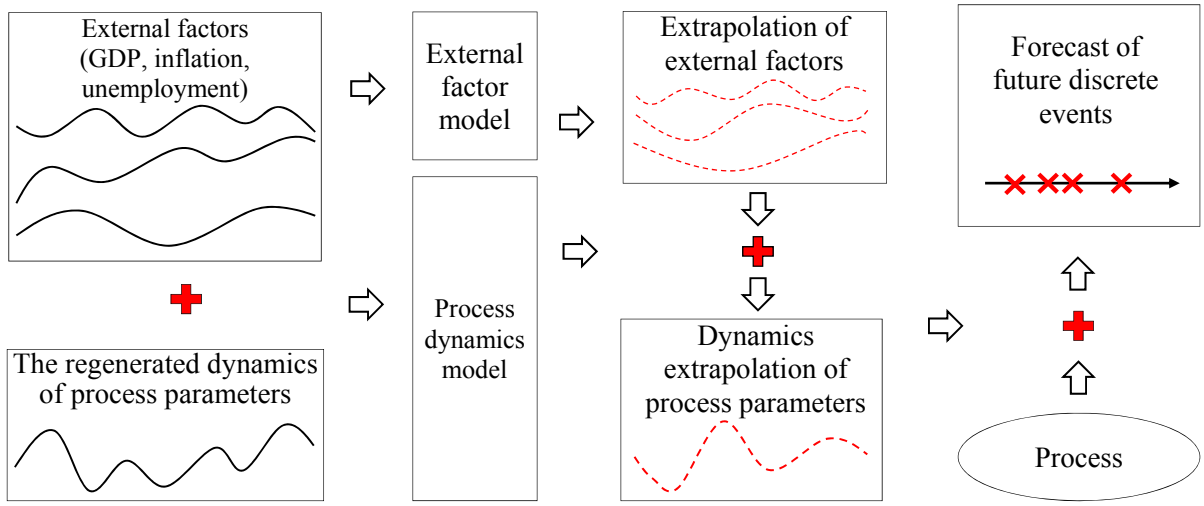


Fig. 7. **Scheme of the approach for predicting rare events in the economy**

Source: compiled by the author.

Some methods predict the occurrence of a given number of events in a certain amount of time or determine only the probability of one event in the next time interval (statistical methods based on Poisson and other flows of events, methods of analysis of irregular demand). Other methods measure the probability of an event if a sequence is observed among observed external factors or lag variables (all possible classification methods). Moreover, in the described approach it is proposed to build a model and extrapolate the restored dependencies of internal parameters over time, when in other methods all estimates are obtained static. In order to compare the proposed approach of prediction of events with other methods it will be necessary to adapt it to specific tasks of other methods. That is, with the help of the method it is necessary to give a forecast at once to a set of future events in an interval of time, and immediately from a multitude of sources (from different clients). This mixed prediction of a group of events can be compared with the prediction of other methods that work with events presented in the form of time series. However, it should be borne in mind that the proposed approach, in addition to the phase of recovery of internal parameters of the process, there are stages of identification of patterns and extrapolation of dynamics of the process, carried out by any known mathematical method. Depending on which methods the researcher chooses, the effectiveness of the proposed approach can be evaluated in different ways. In order to compare performance with other methods, a separate volumetric study is necessary, and maybe even not one.

CONCLUSION

As a result of this study, it was possible to develop a mathematical method of function recovery simultaneously for many different functionalities, taking into account the error in the observation of these functionalities. In *Appendix 1* the software implementation of the method in R language is presented. Due to the described method it is possible to analyze and predict rare events that are caused by some processes. In case it is processes of consumption, it is enough to consider data (purchases, credits, etc.) as integral and restore dependencies from this sequence of integrals. If these are more complex processes, some data can also be considered as first or second derivatives. If we can reconstruct the dynamics of the process of formation of rare events from the data available on these events, the next step will be to determine the pattern and extrapolate this dynamics to the future. And at this stage the researcher is not limited and can use any suitable mathematical method. After the extrapolation stage, it is possible to start the process of formation of events with set values of internal parameters and to get a forecast of future events in the economy. The described approach can be applied in various areas, for example, by analyzing the data on the haircut of some regular customer at the hairdresser, you can restore the function of the rate of accumulation of desire to cut [13], and analyzing the historical data of the Russo-Turkish wars, you can get the speed of the build-up of disagreements or the speed of preparation for another war [14] (the latter has a more demonstrative character). The analysis and prediction of rare events is very important. This will allow to prepare for such events, gain some benefit from it or reduce possible risks or losses.

ACKNOWLEDGEMENTS

The study was funded by the Russian Foundation for Basic Research (RFBR), project No. 19–010–00154. Financial University, Moscow, Russia.

REFERENCES

1. Shannon C. Works on information theory and cybernetics. Transl. from Eng. Moscow: Foreign Literature Publ.; 1963. 258 p. (In Russ.).
2. Croston J. D. Forecasting and stock control for intermittent demands. *Operational Research Quarterly*. 1972;23(3):289–303. DOI: 10.1057/jors.1972.50
3. Willemain T.R., Smart C.N., Schwarz H.F. A new approach to forecasting intermittent demand for service parts inventories. *International Journal of Forecasting*. 2004;20(3):375–387. DOI: 10.1016/S 0169–2070(03)00013-X
4. Kaya G.O., Sahin M., Demirel O.F. Intermittent demand forecasting: A guideline for method selection. *Sādhanā: Academy Proceedings in Engineering Sciences*. 2020;45(1):51. DOI: 10.1007/s12046–020–1285–8
5. Pinçe Ç., Turrini L., Meissner J. Intermittent demand forecasting for spare parts: A critical review. *Omega*. 2021;105:102513. DOI: 10.1016/j.omega.2021.102513
6. Korablev Yu.A. The function restoration method by integrals for analysis and forecasting of rare events in the economy. *Ekonomika i matematicheskie metody = Economics and Mathematical Methods*. 2020;56(3):113–124. (In Russ.). DOI: 10.31857/S 042473880010485–2
7. Kimeldorf G., Wahba G. Some results on Tchebychefian spline functions. *Journal of Mathematical Analysis and Applications*. 1971;33(1):82–95. DOI: 10.1016/0022–247X(71)90184–3
8. Lee C.H. A phase space spline smoother for fitting trajectories. *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics*. 2004;34(1):346–356. DOI: 10.1109/tsmcb.2003.817027
9. Xian J., Li Y., Lin W. Reconstruction of signal from samples of its integral in spline subspaces. In: Bubak M., van Albada G.D., Sloot P.M.A., Dongarra J., eds. Int. conf. on computational science: Computational science (ICCS 2004). Berlin, Heidelberg: Springer-Verlag; 2004:574–577. (Lecture Notes in Computer Science. Vol. 3037). DOI: 10.1007/978–3–540–24687–9_75
10. Green P.J., Silverman B.W. Nonparametric regression and generalized linear models. A roughness penalty approach. Boca Raton, London: Chapman & Hall/CRC; 1994. 194 p. (Monographs on Statistics and Applied Probability. No. 58).
11. Quinn B.G., Fernandes J.M. A fast efficient technique for the estimation of frequency. *Biometrika.* 1991;78(3):489–497. DOI: 10.1093/biomet/78.3.489
12. Quinn B.G., Hannan E.J. The estimation and tracking of frequency. Cambridge, New York: Cambridge University Press; 2001. 278 p. (Cambridge Series in Statistical and Probabilistic Mathematics. No. 9).
13. Korablev Yu.A., Golovanova P.S., Kostritsa T.A. Capacity method of rare events analysis in the area of services. *Ekonomicheskaya nauka sovremennoi Rossii = Economics of Contemporary Russia*. 2020;(3):132–142. (In Russ.). DOI: 10.33293/1609–1442–2020–3(90)-132–142
14. Korablev Yu.A., Golovanova P.S., Kostritsa T.A. Using the capacity method to analyze historical events. *KANT*. 2021;(1):27–32. (In Russ.). DOI: 10.24923/2222–243X.2021–38.6

Appendix 1

Implementation in R

Functionality recovery is implemented as a FunctionalSmoothingSpline function, at the input of which the available data is supplied.

```
FunctionalSmoothingSpline = function (
                        t_f = NULL,          # array of observation moments
                        values_f = NULL,     # array of observation values
                        weights_f = NULL,    # array of  weights
                        t_df = NULL,         # array of first derivative moments
                        values_df = NULL,    # array of first derivative values
                        weights_df = NULL,   # array of first derivative weights
                        coef_df = 1,         # coefficient of first derivative sum of squares
                        t_d2f = NULL,        # array of second derivative moments
                        values_d2f = NULL,   # array of second derivative values
                        weights_d2f = NULL,  # array of second derivative weights
                        coef_d2f = 1,        # coefficient of second derivative sum of squares
                        t_int_a = NULL,      # array of interals start moments
                        t_int_b = NULL,      # array of interals end moments
                        values_int = NULL,   # array of interal values
                        weights_int = NULL,  # array of interal weights
                        coef_int = 1,        # coefficient of integral sum of squares
                        knots = NULL,        # knots
                        knots_number = NULL, # number of knots
                        alpha = 1,           # smoothing parameter
                        x = NULL,            # output moments
                        info = FALSE)        # need info?
{

        if (is.null(knots_number) & is.null(knots))
        {
                knots_number = 0
                if (!is.null(t_f) & length(t_f)>0)
                        knots_number = length(t_f)
                if (!is.null(t_df) & length(t_df)>0 & knots_number<length(t_df))
                        knots_number = length(t_df)
                if (!is.null(t_d2f) & length(t_d2f)>0 & knots_number<length(t_d2f))
                        knots_number = length(t_d2f)
                if (!is.null(t_int_a) & length(t_int_a)>0 & knots_number<length(t_int_a))
                        knots_number = length(t_int_a)
        }

        if (knots_number<2)
                stop('knots_number or observations should not be less than 2')

        m = knots_number # for short

        # in case knots is not defined
        if (m != length(knots)) # when knots_number defined, but knots not defined
        {
                start_knot = + Inf
                end_knot = - Inf
                if (!is.null(t_f) & length(t_f)>0 )
                {
```

```
                if (start_knot>min(t_f))
                        start_knot = min(t_f)
                if (end_knot<max(t_f))
                        end_knot = max(t_f)
        }
        if (!is.null(t_df) & length(t_df)>0)
        {
                if (start_knot>min(t_df))
                        start_knot = min(t_df)
                if (end_knot<max(t_df))
                        end_knot = max(t_df)
        }
        if (!is.null(t_d2f) & length(t_d2f)>0)
        {
                if (start_knot>min(t_d2f))
                        start_knot = min(t_d2f)
                if (end_knot<max(t_d2f))
                        end_knot = max(t_d2f)
        }
        if (!is.null(t_int_a) & length(t_int_a)>0 & length(t_int_b)>0)
        {
                if (start_knot>min(t_int_a))
                        start_knot = min(t_int_a)
                if (end_knot<max(t_int_b))
                        end_knot = max(t_int_b)
        }
        knots=seq(start_knot,end_knot,length = m)
}

h = array(0,dim = m - 1) #array of distance between knots
h[1:(m - 1)] = knots[2:m] - knots[1:(m - 1)]

#Matrix Q
Q=matrix(0, nrow = m, ncol = m - 2)
for (i in 1:(m - 2))
{
        Q[i,i] = 1/h[i];
        Q[i + 1,i] = - 1/h[i] - 1/h[i + 1];
        Q[i + 2,i] = 1/h[i + 1]
}

#Matrix R
R = matrix(0, nrow = m - 2, ncol = m - 2)
for (i in 1:(m - 2))
{
        R[i,i] = 1/3*(h[i] + h[i + 1]);
        if (i<m - 2)
        {
                R[i + 1,i] = 1/6*h[i + 1];
                R[i,i + 1] = 1/6*h[i + 1];
        }
}
        #Matrix K calculation
```

```
inv_R = solve(R)
t_Q = t(Q)
K = Q %*% inv_R %*% t_Q

# ========= 1. Observation (t_f, values_f)  ===========

if (! is.null(t_f) & length(t_f)>0)
{
        nf = length(t_f) #number of observation coordinates
        if (length(values_f) != nf)
                stop('length of values_f and t_f must be same')
        if (is.null(weights_f))
                weights_f = rep(1,nf)
        if (length(weights_f) != nf)
                stop('length of weights_f and t_f must be same')
        Wf = diag(weights_f)

        #reorder observations (t_f, values_f) by appear time t_f
        ord = order(t_f,values_f)
        t_f = t_f[ord]
        values_f = values_f[ord]

        #Filling in Vf and Pf matrices
        Vf = matrix(0,nrow = nf, ncol = m)
        Pf = matrix(0,nrow = nf, ncol = m)
        k = 1 # start knot
        for (i in 1:nf)
        {
                while( (knots[k]<=t_f[i]) & (knots[k+1]<t_f[i]) & (k<knots_number)) #find first k, that
knots[k+1]>t_f[i]
                        k = k+1
                hk_m = t_f[i] - knots[k]
                hk_p = knots[k+1] - t_f[i]
                Vf[i,k] = hk_p/h[k]
                Vf[i,k + 1] = hk_m/h[k]
                Pf[i,k] = hk_m*hk_p*(h[k] + hk_p)/(6*h[k])
                Pf[i,k + 1] = hk_m*hk_p*(h[k] + hk_m)/(6*h[k])
        }
        Pf = Pf[1:nf,2:(m - 1)] #don't need first and last column

        #Matrix Cf calculation
        Cf = Vf - Pf %*% inv_R %*% t_Q
        t_Cf = t(Cf)
}

# ========= 2. Observation (t_df, values_df)  ===========

if (! is.null(t_df) & length(t_df)>0)
{

        ndf=length(t_df) #number of observation
        if (length(values_df) != ndf)
                stop('length of values_df and t_df must be same')
```

```
        if (is.null(weights_df))
                weights_df = rep(1,ndf)
        if (length(weights_df)!=ndf)
                stop('length of weights_df and t_df must be same')
        Wdf = diag(weights_df)

        ord = order(t_df,values_df)   #reorder observations (t_df, values_df)  by appear time t_df
        t_df = t_df[ord]
        values_df = values_df[ord]

        #Filling in Vdf and Pdf matrices
        Vdf = matrix(0, nrow = ndf, ncol = m)
        Pdf = matrix(0, nrow = ndf, ncol = m)
        k = 1 # start knot
        for (i in 1:ndf)
        {
                while( (knots[k]<=t_df[i]) & (knots[k+1]<t_df[i]) & (k<m)) #find first k, that knots[k +
1]>t_df[i]
                        k = k + 1
                hk_m = t_df[i] - knots[k]
                hk_p = knots[k + 1] - t_df[i]

                Vdf[i,k] = - 1/h[k]
                Vdf[i,k + 1] = 1/h[k]
                Pdf[i,k] = - h[k]/6+(hk_p)^2/(2*h[k])
                Pdf[i,k + 1] = h[k]/6-(hk_m)^2/(2*h[k])
        }
        Pdf = Pdf[1:ndf,2:(m-1)] #don't need first and last column

        #Matrix Cdf calculation
        Cdf = Vdf - Pdf %*% inv_R %*% t_Q
        t_Cdf = t(Cdf)

}

# ========= 3. Observation (t_d2f, values_d2f)  ===========

if (! is.null(t_d2f) & length(t_d2f)>0)
{

        nd2f = length(t_d2f) #number of observation
        if (length(values_d2f) != nd2f)
                stop('length of values_d2f and t_d2f must be same')
        if (is.null(weights_d2f))
                weights_d2f = rep(1,nd2f)
        if (length(weights_d2f) != nd2f)
                stop('length of weights_d2f and t_d2f must be same')
        Wd2f = diag(weights_d2f)

        #reorder observations (t_d2f, values_d2f)  by appear time t_d2f
        ord = order(t_d2f,values_d2f)
        t_d2f = t_d2f[ord]
        values_d2f = values_d2f[ord]
```

```
                #Filling in Pd2f matrices
                Pd2f=matrix(0, nrow = nd2f,ncol = m)
                k = 1 # start knot
                for (i in 1:nd2f)
                {
                        while( (knots[k]<=t_d2f[i]) & (knots[k+1]<t_d2f[i]) & (k<m)) #find first k, that knots[k+1]>t_d2f[i]
                                k = k + 1
                        hk_m = t_d2f[i] - knots[k]
                        hk_p = knots[k + 1] - t_d2f[i]

                        Pd2f[i,k] = - hk_p/h[k]
                        Pd2f[i,k+1] = - hk_m/h[k]
                }
                Pd2f = Pd2f[1:nd2f,2:(m - 1)] #don't need first and last column

                #Matrix Cd2f calculation
                Cd2f = - Pd2f %*% inv_R %*% t_Q
                t_Cd2f = t(Cd2f)

        }

        # ========= 4. Observation (t_int_a, t_int_b, values_int)  ===========

        if (! is.null(t_int_a) & length(t_int_a)>0)
        {
                nint=length(t_int_a) #number of observation
                if (length(t_int_b) != nint)
                        stop('length of t_int_b and t_int_a must be same')
                #if (length(values_int) != nint)
                        stop('length of values_int and t_int_a must be same')

                if (is.null(weights_int))
                        weights_int = rep(1,nint)
                if (length(weights_int) != nint)
                        stop('length of weights_int and t_int_a must be same')
                Wint=diag(weights_int)

                #reorder observations (t_int_a, t_int_b, values_int) by appear time t_int_a
                ord = order(t_int_a, t_int_b,values_int)
                t_int_a = t_int_a[ord]
                t_int_b = t_int_b[ord]
                values_int = values_int[ord]

                #Filling in Vint and Pint matrices
                Vint = matrix(0, nrow = nint, ncol = m)
                Pint = matrix(0, nrow = nint, ncol = m)
                k = 1 # start knot
                for (i in 1:nint)
                {
                        while( (knots[k]<=t_int_a[i]) & (knots[k + 1]<t_int_a[i]) & (k<m)) #find first k, that knots[k + 1]>t_int_a[i]
                                k = k + 1
```

```
                #finding L, it can be 0
                for (L in 0:(m - k - 1))
                        if (t_int_b[i] <= knots[k + L + 1])
                                break;
                l = 1;
                hk_m = t_int_a[i] - knots[k]
                hk_p=knots[k + 1] - t_int_a[i]
                hkL_m = t_int_b[i] - knots[k + L]
                hkL_p = knots[k + L + 1] - t_int_b[i]

                Vint[i,k] = (hk_p)^2/h[k]/2
                Pint[i,k] = h[k]^3/24 - (hk_m)^2*(hk_p + h[k])^2/h[k]/24
                while (l<=L)
                {
                        Vint[i, k + l] = (h[k + l - 1] + h[k + l])/2
                        Pint[i, k + l] = (h[k + l - 1]^3 + h[k + l]^3)/24
                        l = l + 1;
                }
                Vint[i, k + 1] = Vint[i, k + 1] - (hk_m)^2/h[k]/2
                Pint[i, k + 1] = Pint[i, k + 1] + (hk_m)^2*((hk_m)^2 - 2*h[k]^2)/h[k]/24
                Vint[i, k + L] = Vint[i, k + L] - (hkL_p)^2/h[k + L]/2
                Pint[i, k + L] = Pint[i, k + L] + (hkL_p)^2*((hkL_p)^2 - 2*h[k + L]^2)/h[k + L]/24
                Vint[i, k + L + 1] = (hkL_m)^2/h[k + L]/2
                Pint[i, k + L + 1] = h[k + L]^3/24 - (hkL_p)^2*(hkL_m + h[k + L])^2/h[k + L]/24

        }
        Pint=Pint[1:nint,2:(m - 1)] #don't need first and last column

        #Matrix Cint calculation
        Cint = Vint - Pint %*% inv_R %*% t_Q
        t_Cint = t(Cint)

}

# ============ Calculation =============

# matrix A
A = alpha * K
if (! is.null(t_f) & length(t_f)>0)
        A = A + t_Cf %*% Wf %*% Cf
if (! is.null(t_df) & length(t_df)>0)
        A = A + coef_df * t_Cdf %*% Wdf %*% Cdf
if (! is.null(t_d2f) & length(t_d2f)>0)
        A = A + coef_d2f * t_Cd2f %*% Wd2f %*% Cd2f
if (! is.null(t_int_a) & length(t_int_a)>0)
        A = A + coef_int * t_Cint %*% Wint %*% Cint

# matrix D
D = matrix(0, nrow = m, ncol = 1)
if (! is.null(t_f) & length(t_f)>0)
        D = D + t_Cf %*% Wf %*% values_f
if (! is.null(t_df) & length(t_df)>0)
```

```
            D = D + coef_df * t_Cdf %*% Wdf %*% values_df
        if (! is.null(t_d2f) & length(t_d2f)>0)
            D = D + coef_d2f * t_Cd2f %*% Wd2f %*% values_d2f
        if (! is.null(t_int_a) & length(t_int_a)>0)
            D = D + coef_int * t_Cint %*% Wint %*% values_int

        #Calculation of g and gamma
        g = solve(A , D)
        gamma = inv_R %*% t_Q %*% g   #After that spline is completely defined via g and gamma


        # ===== Calculating and returning spline values at x coordinates  =====

        g2 = c(0,gamma,0) #Second derivative on the edges was zero

        if (is.null(x))
            x = seq(knots[1],knots[m],by=1)

        y = rep(0,length(x))

        k = 1; #index of interval
        for (j in (1:length(x)))
        {
            while (x[j]>knots[k]+h[k] & k<m)
                k = k + 1;
            hk_m = x[j] - knots[k]
            hk_p = knots[k + 1] - x[j]
            y[j] = (hk_m*g[k + 1] + hk_p*g[k])/h[k] - 1/6*hk_m*hk_p*(g2[k + 1]*(1 + hk_m/h[k]) + g2[k]*(1 +
hk_p/h[k]) )
        }

        if (info)
        {
            error_total = 0
            error_f = 0
            error_df = 0
            error_d2f = 0
            error_int = 0
            error_penalty = 0
            fraction_error_f = 0
            fractio_error_df = 0
            fractio_error_d2f = 0
            fractio_error_int = 0
            fractio_penalty = 0
            relative_sqr_error_f = 0
            relative_sqr_error_df = 0
            relative_sqr_error_d2f = 0
            relative_sqr_error_int = 0
            relative_abs_error_f = 0
            relative_abs_error_df = 0
            relative_abs_error_d2f = 0
            relative_abs_error_int = 0
```

```
if (! is.null(t_f) & length(t_f)>0)
{
        V = values_f - Cf %*% g
        error_f = t(V) %*% Wf %*% V
        V = abs(V / values_f)
        relative_abs_error_f = (t(V) %*% Wf %*% V) / nf
        V = V^2
        relative_sqr_error_f = sqrt((t(V) %*% Wf %*% V) / nf)
}
if (! is.null(t_df) & length(t_df)>0)
{
        V = values_df - Cdf %*% g
        error_df = t(V) %*% Wdf %*% V
        V = abs(V / values_df)
        relative_abs_error_df = (t(V) %*% Wdf %*% V) / ndf
        V = V^2
        relative_sqr_error_df = sqrt((t(V) %*% Wdf %*% V) / ndf)
}
if (! is.null(t_d2f) & length(t_d2f)>0)
{
        V = values_d2f - Cd2f %*% g
        error_d2f = t(V) %*% Wd2f %*% V
        V = abs(V / values_d2f)
        relative_abs_error_d2f = (t(V) %*% Wd2f %*% V) / nd2f
        V = V^2
        relative_sqr_error_d2f = sqrt((t(V) %*% Wd2f %*% V) / nd2f)
}
if (! is.null(t_int_a) & length(t_int_a)>0)
{
        V = values_int - Cint %*% g
        error_int = t(V) %*% Wint %*% V
        V = abs(V / values_int)
        relative_abs_error_int = (t(V) %*% Wint %*% V) / nint
        V = V^2
        relative_sqr_error_int = sqrt((t(V) %*% Wint %*% V) / nint)
}
error_penalty = t(g) %*% K %*% g

error_total = error_f + coef_df*error_df + coef_d2f*error_d2f + coef_int*error_int + alpha*error_
penalty

fraction_error_f = error_f/error_total
fraction_error_df = coef_df*error_df/error_total
fraction_error_d2f = coef_d2f*error_d2f/error_total
fraction_error_int = coef_int*error_int/error_total
fraction_penalty = alpha*error_penalty/error_total

result = list(      x = x,
                    y = y,
                    error_total = error_total,
                    error_f = error_f,
                    error_df = error_df,
                    error_d2f = error_d2f,
                    error_int = error_int,
```

```
                                    error_penalty = error_penalty,
                                    fraction_error_f = fraction_error_f,
                                    fraction_error_df = fraction_error_df,
                                    fraction_error_d2f = fraction_error_d2f,
                                    fraction_error_int = fraction_error_int,
                                    fraction_penalty = fraction_penalty,
                                    relative_sqr_error_f = relative_sqr_error_f,
                                    relative_sqr_error_df = relative_sqr_error_df,
                                    relative_sqr_error_d2f = relative_sqr_error_d2f,
                                    relative_sqr_error_int = relative_sqr_error_int,
                                    relative_abs_error_f = relative_abs_error_f,
                                    relative_abs_error_df = relative_abs_error_df,
                                    relative_abs_error_d2f = relative_abs_error_d2f,
                                    relative_abs_error_int = relative_abs_error_int
                            )
        }
        else
                result = y
        return (result) }
```

Appendix 2

Computing in R

Example 1

Read data from a CSV file (which contains all corresponding columns):

```
#================ Data input ===================
library(lubridate)
filename = « F:/DIR/Sales.csv»;

#if CSV file was generated by Excel
MyData <- read.csv(file = filename, header = TRUE, sep = «;», stringsAsFactors = FALSE, dec =»,»)
t = as.numeric(dmy(MyData[[1]]))

#if CSV file was generated by R
#MyData <- read.csv(file = filename, header = TRUE, sep = «,», stringsAsFactors = FALSE, dec =».»)
#t = as.numeric(ymd(MyData[[1]]))
```

Deleting missing values:

```
# ================= Remove NA ====================
t = t[!is.na(t)]
n = length(t)
Y = MyData[[2]]
Y = Y[1:(n - 1)] # Last value not used
origin = dmy(MyData[[1]][1]) - t[1] #will need time origin for x-axis labes
```

Specify the number of nodes, you can take several times more than the number of all observations:

```
m = round(n*3)
```

Function set, result is saved to variable r. If Info = FALSE, the calculated values will be displayed immediately y.

```
#=========== Calculating spline =============
r = FunctionalSmoothingSpline(#t_f = t_f,
                #values_f = y_f,
```

```
                #weights_f = NULL,
                #t_df = t_df,
                #values_df = y_df,
                #weights_df = NULL,
                #coef_df = 1,
                #t_d2f = t_d2f,
                #values_d2f = y_d2f,
                #weights_d2f = NULL,
                #coef_d2f = 1,
                t_int_a = t[1:(n-1)],
                t_int_b = t[2:n],
                values_int = Y,
                weights_int = W,
                #coef_int = 1,
                #knots = NULL,
                knots_number = m,
                alpha = 10^(4),
                info = TRUE)
#r
y = r$y
```

Rendering a chart:

```
#================= Plotting spline with graph of average values ==============
x = r$x
#x = seq(t[1], t[n], by = 1) # In case Info = FALSE
x2 = t[1]:t[n] #for graph of average values
y2 = rep(0, length(x2))
for (i in 1:(n - 1))
  for (j in t[i]:t[i + 1])
    y2[j-t[1] + 1] = Y[i]/(t[i + 1] - t[i])


plot(x, y, col = “red”, type = “l”, lwd = “1”, lty = 1, xaxt = “n”, ylim = range(c(y,y2)), xlim = range(c(x,x2)))
axis.Date(1, at=seq(min(dmy(MyData[[1]])), max(dmy(MyData[[1]])), by = “months”), format = “%m-%Y”)
lines(x2, y2, col = “black”, type = “l”, lwd = “1”, lty = 1)
```

File output:

```
#================ Data output ===================
MyWriteData = data.frame(t = x + origin, Value = y, x2 = x2 + origin, y2 = y2)
s2 = “F:/DIR/f_spline_out.csv”;
write.csv(MyWriteData, file = s2,row.names=FALSE)
```

Example 2

Read data from a CSV file (which contains all corresponding columns):

```
#================ Data input ===================

Library (lubridate)
filename = “F:/Dir/DiscrSignals.csv”;

#if CSV file was generated by Excel
```

```
MyData < – read.csv(file = filename, header = TRUE, sep = ";", stringsAsFactors = FALSE, dec = ",")
t_f = as.numeric(dmy(MyData[[1]]))
y_f = MyData[[2]]
t_df = as.numeric(dmy(MyData[[3]]))
y_df = MyData[[4]]
t_d2f = as.numeric(dmy(MyData[[5]]))
y_d2f = MyData[[6]]
t_int_a = as.numeric(dmy(MyData[[7]]))
t_int_b = as.numeric(dmy(MyData[[8]]))
y_int = MyData[[9]]

#if CSV file was generated by R
#MyData < – read.csv(file = filename, header = TRUE, sep =",", stringsAsFactors = FALSE, dec =".")
#t_f = as.numeric(ymd(MyData[[1]]))
#t_df = as.numeric(ymd(MyData[[3]]))
#t_d2f = as.numeric(ymd(MyData[[5]]))
#t_int_a = as.numeric(ymd(MyData[[7]]))
#t_int_b = as.numeric(ymd(MyData[[8]]))
```

Delete missing values at the very end (CSV file had different column lengths, but missing values still read):

```
# ================= Remove NA ====================
t_f = t_f[!is.na(t_f)]
nf = length(t_f)
y_f = y_f[1:nf]
t_df = t_df[!is.na(t_df)]
ndf = length(t_df)
y_df = y_df[1:ndf]
t_d2f = t_d2f[!is.na(t_d2f)]
nd2f = length(t_d2f)
y_d2f = y_df[1:nd2f]
t_int_a = t_int_a[!is.na(t_int_a)]
nint = length(t_int_a)
t_int_b = t_int_b[!is.na(t_int_b)]
y_int = y_int[1:nint]

#will need time origin for x-axis labes
origin = dmy(MyData[[1]][1]) – t_f[1]
#origin = ymd(MyData[[1]][1]) – t_f[1]
```

Specify the number of nodes, you can take several times more than the number of all observations:

```
m = round(3*(nf + ndf + nd2f + nint))
```

Function set, result is saved to variable r. If Info = FALSE, the calculated values will be displayed immediately y:

```
#========== Calculating spline =============
r = FunctionalSmoothingSpline( t_f = t_f,
                               values_f = y_f,
                               t_df = t_df,
                               values_df = y_df,
```

```
                                    t_d2f = t_d2f,
                                    values_d2f = y_d2f,
                                    t_int_a = t_int_a,
                                    t_int_b = t_int_b,
                                    values_int = y_int,
                                    knots_number = m,
                                    alpha = 10^(0),
                                    info = TRUE)
y = r$y
```

Rendering a chart:

```
#============================ Plotting spline ============================
x = r$x
#x = seq (min(t_f, t_df, t_d2f, t_int_a), max (t_f, t_df, t_d2f, t_int_b), by = 1) # In case Info = FALSE
xrange = range (x)
yrange = range (y)
plot(x, y, col = "red", type = "l", lwd = "1", lty = 1,  xaxt = "n", ylim = yrange, xlim = xrange)
axis (1, at = c (t_f,t_df,t_d2f,t_int_a,t_int_b), labels = as. Date (c(t_f,t_df,t_d2f,t_int_a,t_int_b), origin=origin))
#lines(x,y2,col = "black", type = "l", lwd="2", lty = 1) # In case you want add some precalculated original function y2
```

File ouput:

```
#================ Data output ===================
#write to the csv file
MyWriteData = data.frame(x + origin, y)
s2 = «F:/DIR/f_spline_out.csv»
write.csv (MyWriteData, file = s2,row.names = FALSE
```

ABOUT THE AUTHOR

Yurii A. Korablev — Cand. Sci. (Econ.), Assoc. Prof., Department of System Analysis in Economics, Financial University, Moscow, Russia
http://orcid.org/0000-0001-5752-4866
yura-korablyov@yandex.ru

Conflicts of Interest Statement: The author has no conflicts of interest to declare.

The article was submitted on 21.10.2021; revised on 08.11.2021 and accepted for publication on 25.02.2022.
The author read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-226-240
JEL C32, C50, G15

Price Discovery of Agri Commodities: An Integrated Approach

A. Mishra, R.P. Kumar
National Institute of Food Technology Entrepreneurship and Management, Sonipat, India

ABSTRACT

The **purpose** of this paper is to examine how the introduction of the pan-India electronic trading portal (eNAM) impacted the information transmission and price discovery in informationally linked markets of India for agricultural commodities. We have applied the information share and component share **methods** using daily data of five agricultural commodities from 2005 to 2019. This paper offers two **findings**, first; evidence shows that commodity stock prices have more price discovery as compared to the market prices. Whereas market prices also generate a significant price discovery, but its volume is less. Second, market prices adjust more quickly than commodity stock prices to correct the disproportion between both markets. This scenario is more evident post eNAM era. **Conclusions** from this study can be used to understand the information flow and would be helpful to academicians, practitioners, policymakers, or business players of commodity markets.

Keywords: price discovery; agricultural commodity; information share; component share

For citation: **Mishra A., Kumar R.P. Price discovery of agri commodities: An integrated approach. *Finance: Theory and Practice*. 2022;26(3):226-240. DOI: 10.26794/2587-5671-2022-26-3-226-240**

INTRODUCTION

Agriculture is one of the major sectors of the Indian economy, which has a significant role in India's GDP and employment. Any strategic decision would impact people at the grassroots level who have associated with this sector anywhere. In the Indian context, there are three main aspects of strategic agri decision-making; (1) to empower the economic status of the farmers, (2) to strengthen the associated labours, and (3) to be consistent with the international commodity prices [1, 2]. Considering the importance of transparency in trade (through better price discovery), accessibility to farmers, and quick payments, the Government of India has introduced a pan-India electronic trading portal that networks the existing APMCs (physical markets established under the provisions of the APLM India Act proposed by the Agricultural Produce & Livestock Market Committee) to create a unified national market for agricultural commodities called National agriculture Market (eNAM) in 2016.[1]

Due to the technology adaptation and requirement of transparency in trade, price discovery occurs in dynamic commodity markets [3]. In this paper, we will be considering the price discovery process of five agri commodities traded in different markets. The price discovery process is often viewed as information transmission. Price discovery is the process of impounding new information into the commodity's market price and is one of the essential products of markets [4].

In this paper, we have investigated which market provides more useful information regarding fundamental value for these commodities in India. The commodities are (1) cotton — India is one of the largest producers accounting for about 26% of the world cotton production as well as the third-largest exporter of cotton.[2]. Cotton has posted significant positive growth of 68% in exports which is US$ 923 million to US$ 1,550 million between FY 20 and FY 21,[3] (2) maize — India ranks 4th in area and 7th in production if we only consider maize growing countries. During 1950–1951 India produced 1.73 million metric tons (MT) of maize, which has increased to 27.8 million MT by 2018–2019, recording close to 16 times increase in

[1] Department of Agriculture G of I. e-NAM Overview. URL: https://www.enam.gov.in/web/ (2021) (accessed on 02.11.2021).

[2] COCPC. National Cotton Scenario. URL: https://cotcorp.org.in/national_cotton.aspx (2020) (accessed on 02.11.2021).

[3] IBEF IBEF. Indian Agriculture and Allied Industries Report (July, 2021). URL: https://www.ibef.org/industry/agriculture-presentation (accessed on 02.11.2021).

production,[4] (3) wheat — India ranked second in wheat production after China, having a production share of 103.6 million MT in the year 2019,[5] (4) barley — is one of the four major feed grains (corn, barley, oats, and wheat) and is used commercially for animal feed, to manufacture malt, which is primarily used in beer production, for seed, and human food applications [5], (5) soybean — is the world's most crucial seed legume, contributing to approximately 25% of the world's edible oil, and about 65% of the global protein concentrate for livestock feeding. In the Indian context, the share of Soybean is approximately 40% of the total oilseeds and 25% of the edible oils [6].

Since Government in India is implementing the reforms to promote uniformity in the agriculture markets by integrating across the markets, measuring the price discovery is of particular relevance for policymaking also. The government is aiming to remove the information asymmetry between traders and farmers and promote real-time price discovery across the markets.[6] Understanding the price discovery in agriculture is important for market applicants and policymakers because it can contribute to better management decisions and more informed policy debates on market regulation [7]. It would be beneficial to study the flow of information between the markets and the commodities as it could be beneficial for traders and farmers both. The findings of this study can be used to understand the information flow and its impact on pricing to make relative trading strategies; if a commodity is being traded-in multiple markets. The farmers are trading directly at eNAM, and how far they get price discovered in commodity exchanges which a great concern for policy implications. It also regulates the public policy implication for the active participation of farmers in national-level commodity exchanges. As we proceed, the following section consists of a literature review, followed by the methodological framework, analysis and connectedness measurements, results and discussion, and lastly, conclusions.

[4] ICAR. India Maize Scenario. URL: https://iimr.icar.gov.in/india-maze-scenario/ (2020) (accessed on 02.11.2021).

[5] IBEF. Wheat production may cross 113 million tonnes: Skymet. URL: https://www.ibef.org/news/wheat-production-may-cross-113-million-tonnes-skymet (accessed on 02.11.2021).

[6] Department of Agriculture G of I. e-NAM Overview. URL: https://www.enam.gov.in/web/ (2021) (accessed on 02.11.2021).

REVIEW OF LITERATURE

Price discovery is the process by which the fundamental value of a commodity get reflects in the price [8]. Many researchers have discussed price discovery but the research work is restricted to the Future and Spot markets or prices. Man has investigated that the electronic trading system has more price discovery, and the share by the trading system depends on the volume, liquidity, and volatility [3]. Ahumada has developed a forecasting approach to test price discovery in a multivariate framework focusing on the soybeans market. They also found that future prices are the best predictors of future spot prices [9]. Dimpfl has examined eight commodities' spot and futures prices and found that efficient prices can be determined by the spot prices in the long run [10]. The results couldn't confirm the role of future markets in price discovery. They have explained that by understanding the market leadership in price discovery, we can look into the potential exposure of the actual prices that may be used for speculation in the future market. Ates has found that both floor and equity trading contribute significantly to price discovery [11]. However electronic trading is superior in terms of operational efficiency and relative liquidity. Dolatabadi has applied the "fractionally cointegrated vector autoregressive" (FCVAR) model to examine the relationship between spot prices and futures prices in five commodities [12]. They concluded that less evidence of long-run backward integration as compared to the non-fractional model, which Figuerola-Ferretti and Gonzalo, 2010 applied. Ahumada & Cornejo have used the time-series cross-sectional approach to examine the price formation in the US commodity market, which is determined by the demand and supply in the long run and demand-pull from China [13]. Baffes & Ajwad have explained the price linkages and the degree to which the prices (price discovery process) are related in different markets for cotton [14]. Liu & An have examined the price discovery in the US and Chinese commodity futures markets for copper and soybean spot contracts using M-GARCH and information share models [15]. Figuerola-Ferretti & Gonzalo have established an equilibrium model of spot and futures for non-ferrous metals prices traded on the London Metal Exchange [16]. They confirm that future prices are "information dominant" in

highly liquid futures markets. Baillie & Geoffrey have talked about the common factor models in the case when one commodity/asset is traded in more than one market [17]. They have compared the relationship between the two widely used common factor models; Hasbrouck (1995) [4], which considers the variance of the innovations to the common factor, and Gonzalo and Granger (1995) [18] which considers the components of the common factor and the error correction process. Karabiyik & Narayan have found that the spot market drives the price discovery as compared to future prices considering Islamic stocks from 19 markets [19]. Lien & Shrestha have proposed the Generalized Information Share (GIS) model to analyze the price discovery process in interrelated financial markets, which is found more efficient than Hasbrouck's (1995) [4] information share (IS) in their case [8]. Avino & Lazar [20] have examined multivariate GARCH models to generate a time-varying Hasbrouck (1995) [4] information share (IS) that can improve credit spread predictions. Kapar & Olmo have analyzed the price discovery process for Bitcoin spot and futures markets and found that future markets drive the price discovery [21]. Bohl & Siklos have examined price discovery for agricultural commodity markets and found that future markets contribute more to price discovery where speculation reduces the noise in future markets [22]. Grammig & Peter have examined the high-frequency data of stocks and found that contribution of the NYSE to price discovery has sharply declined from 2007 to 2012 [23]. Aggarwal & Thomas have examined the spot and futures markets for information flow and liquidity [24]. Putnins has examined three popular methods — the Hasbrouck information share, Harris-McInish-Wood component share, and information leadership share and found that ILS is correctly attributed to price discovery [25]. Narayan & Smyth have suggested the need for further research in econometric modeling integrating recent methods and empirical regularities in price discovery [26]. Recently, Hasbrouck has extended the application of his model by examining the price discovery in high-frequency data by estimating the multivariate time series models [27]. Here, *Table 1* shows the summary of some studies on price discovery for different types of commodities.

Since the risk transfer and price discovery are considered two primary functions of future markets [21], the price discovery process is well investigated in future markets. We have observed a gap where the price discovery process needs to examine for agri-commodities. These are commodities are traded in different markets. We want to understand the price discovery process for these agri-commodities in the commodity stock prices commodity market and National agriculture Market (eNAM) considering the following points:

1. To collect the evidence on the price discovery process in the agri-commodity markets.
2. To contribute to the knowledge of the relationships of the prices of the same commodity traded in different markets.
3. Understanding how the different markets are reacting to the new information share leads to price discovery specifically before and after the eNAM introduction.

METHODOLOGY

Based on the literature review, we can say that three popular techniques have been developed and are most widely used to estimate price discovery:

1. Information Share (IS), proposed by Hasbrouck (1995) [4].
2. Component Share (CS) proposed by Booth, So, and Tse (1999) [44], Chu, Hsieh, and Tse (1999) [45], Harris, McInish, and Wood (2002) [46], which is based on the permanent and transitory decomposition of Gonzalo and Granger (1995) [18].
3. The information leadership share (ILS) proposed by Putnins (2013) is a combination of both [25].

We would be using the first two methods to measure the price discovery.

Hasbrouck Information Share (IS) Measure:

To determine the contributions to price discovery is nothing but to identify each market's contribution to the variance of the innovations in the common random walk component. Since markets' innovations are potentially serially correlated, we require to extract the idiosyncratic effect of the markets independently. Referring to the methodology suggested by Shrestha and Lee [47], Hasbrouck Information Shares rely on a vector equilibrium correction of price changes in n markets. If Y_t be an $n \times 1$ vector of unit-root

Table 1

Summary of the Studies on Price Discovery Between Commodities

Study Reference	Methods	Period	Commodity / type	Summary
T. Vollmer [7]	Partially cointegrated error correction model, IS, CS & ILS	Jan 2002 to Apr 2016	agricultural (spot and futures)	In the context of efficiency or avoidance of noise, the Paris wheat futures market dominates price discovery, but this dominance decreases if price volatility increases
H. Karabiyik [19]	VECM	1982 to 2015	Islamic stocks	Spot markets lead the price discovery process for most the countries
B. Kapar [21]	IS, CFC	Dec 2017 to May 2018	Bitcoin	Futures market dominates the price discovery and a weighted combination of the futures and spot market called the "common factor" drives the prices
M.T. Bohl [22]	IS, CS & ILS	Jan1990 to Jun 2018	Agricultural, livestock (spot and futures)	Speculation (total and excessive) im- proves futures markets' price discovery
J. Grammig [23]	IS, GG	2004 to 2012	Equity	NYSE's contribution to price discovery has sharply declined, and later was overtaken by the NASDAQ
F. Balli [28]	VECM	2007–2016	Various commodities (spots and futures)	Commodity indexes are well connected and precisely during global financial crisis and oil price collapse (2014–16).
J. Hao [29]	VECM, CFW, PT & IS	2017–2018	Soybeans	Soybean options show stronger price discovery than soybean futures and call options trading volume has a stronger impact on Soybean options' price discovery than the put options.
T. Dimpfla [30]	IS, CS & ILS	Mar to Nov 2017	Cryptocurrency	Bitfinex (trading venue) is the leader in the price discovery process
J. Yang [31]	Recursive cointegration	Various (2004 to 2019)	Agricultural	Futures markets dominate the price discovery
K. Shrestha [32]	Generalized IS, CS	Various (1979 to 2017)	Agricultural (spot and futures)	Futures markets lead the price discovery process except for cocoa
A. Fassas [33]	CFW, IS, VECM	Jan 2013 to Dec 2014	Indices / futures Contracts	Futures markets lead the price discovery process
B. Frijns [34]	Structural VAR	Jan 2004 to Aug 2017	US & Canadian stocks	US market dominates the prices & algorithmic trading is negatively related to price discovery
P.K. Narayan [35]	VECM	Various (1979 to 2012)	Agricultural (spot and futures)	Spot markets lead the price discovery for nine commodities while future markets dominate only six
J. Wright [36]	VECM, Cluster analysis	May 2001 to Oct 2016	Livestock (Cattle)	Futures markets lead the price discovery for the US cattle market
B.S. Rout [37]	Cointegration, Granger causality, and VECM	Jan 2010 to Dec 2016	Agri and metal	Derivative market leads the price discovery for metals and spot for agri commodities
H. Karabiyik [38]	IS, PT	Jun 2017 to Dec 2017	Energy, foodstuffs, agri, livestock & metals (spot and futures)	IS and PT measures are consistent for Panel data measures of price discovery whereas time series only support PT

Table 1 (continued)

Study Reference	Methods	Period	Commodity / type	Summary
S.T.G. Nair [39]	Cointegration, VECM	2008 to 2019	Metals (spot and futures)	Futures markets lead the price discovery process
R. de Blasis [40]	Multivariate Markov chain	Jan 2016 to Dec 2017	Gold (spot and futures)	The author proposes a new measure of price discovery called price leadership share (PLS)
R. Manogna [41]	VECM, EGARCH	2010 to 2020	Agricultural (spot and futures)	Future markets lead the price discovery for six commodities while spot markets dominate only three
A.P. Fassas [42]	Recursive cointegration analysis, multivariate GARCH, IS	Jan 2018 to Dec 2018	Bitcoin	Futures markets lead the price discovery
M. Li [43]	VECM, cointegration, SADF	Sep 2004 to Sep 2017	Soybean futures markets	Price discovery well exists during price bubble periods as compared to non-bubble periods

Source: compiled by the authors.

Notes: VECM — vector error correction model; CFW — Common factor weight; IS — Information Share; CS — Component shares; ILS — Information leadership shares; PT — Permanent Transitory; CFC — common factor component; VAR — vector autoregression; GG — Gonzalo-Granger measure; GARCH — generalized autoregressive conditional heteroscedasticity; EGARCH — exponential generalized autoregressive conditional heteroscedasticity; SADF — Supremum Augmented Dickey-Fuller Test.

series where it is assumed that there are $(n-1)$ cointegrating vectors which implies that the system consists of a single common stochastic trend. So, we can represent the data generating process by a vector error-correction (VEC) model as below:

$$\Delta Y_t = \alpha\beta^T Y_{t-1} + \sum_{i=1}^{k} A_i \Delta Y_{t-i} + \varepsilon_t. \quad (1)$$

Where β and α are the $n \times (n-1)$ matrices of rank $(n-1)$. The columns of β consists of the $(n-1)$ cointegrating vectors, and each column of α represent the adjustment coefficients. The matrix $\alpha\beta^T$ is decomposed in such a way that $\beta^T Y_t$ consists of $(n-1)$ vectors of stationary series. Let Ω denote the covariance matrix of the innovation vector, i.e., $\mathbb{E}\left[\varepsilon_t \varepsilon_t^T\right] = \Omega$. The equation (1) can be transformed into the following vector moving average (VMA)

$$\Delta Y_t = \psi(L)\varepsilon_t. \quad (2)$$

Or

$$Y_t = Y_0 + \psi(1)\sum_{i=1}^{t}\varepsilon_t + \psi^*(L)\varepsilon_t, \quad (3)$$

where $\psi_0 = I_n$ is an identity matrix. Due to the assumed nature of the cointegrating relationship among these unit-root series, the Engle-Granger representation theorem [Engle and Granger (1987)] implies the following (De Jong (2002) and Lehmann (2002)):

$$\beta^T \psi(1) = 0 \text{ and } \psi(1)\alpha = 0. \quad (4)$$

Based on the above representations, Hasbrouck (1995) considers $\psi(1)\varepsilon_t$ to represent the long-run impact of the reduced form innovations on the unit-root series [4]. Since Hasbrouck (1995) assumes that the cointegrating relationship is one-to-one. Therefore, if n non-stationary series to have $(n-1)$ cointegrating vectors, the cointegrating vectors represented by columns of matrix β can be written as follows

$$\beta^T_{(n-1)\times n} = \left[\iota_{(n-1)} : -I_{(n-1)}\right] = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \vdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix}, \quad (5)$$

where $I_{(n-1)} = Diag\ (1, 1, \ldots, 1)$ and $\iota_{(n-1)}$ is an $(n-1)$ element column vector with all its elements

equal to 1. Given the unique nature of β, equation (4) implies that $\psi(1)$ has identical rows. Let $\psi = (\psi_1, \psi_2, \ldots, \psi_n)$ denote the identical row of $\psi(1)$. Then, the information share for market i (S_i) is given by Hasbrouck (1995) [4]:

$$S_i = \frac{\left([\psi F]_j\right)^2}{\psi \Omega \psi^t},$$

where F is the Cholesky factorization of Ω and $[\psi F]_j$ represents the j^{th} element of the row vector ψF. As the Cholesky factorization depends on the ordering, the S_i computed using the above equation will depend on the particular ordering. By considering all possible orderings, we can compute the upper and lower bounds for S_i. Then, the IS of market i (IS_i) is given by the average of the upper and lower bounds for S_i.

For a two series case, the covariance matrix Ω and F are given by:

$$\Omega = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \text{ \& } F = \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{(1-\rho^2)} \end{bmatrix}.$$

Therefore, the two IS bounds (IS_1 and IS_1^*) for the first market are given by

$$IS_1 = \frac{\left(\psi_1\sigma_1 + \psi_2\rho\psi_2\right)^2}{\psi_1^2\sigma_1^2 + \psi_2^2\sigma_2^2 + 2\psi_1\psi_2\sigma_{12}},$$

$$\text{\& } IS_1^* = \frac{\left(\psi_1\sigma_1\right)^2\left(1-\rho^2\right)}{\psi_1^2\sigma_1^2 + \psi_2^2\sigma_2^2 + 2\psi_1\psi_2\sigma_{12}}. \quad (6)$$

As the sum of the IS for the two markets is equal to 1, the IS for the second market can easily be computed. If the two elements of the reduced form innovation ε_t are independent, then a unique IS measure exists for each market. The IS for market i (IS_i) is given by

$$IS_i = \frac{\psi_i^2\sigma_i^2}{\psi_1^2\sigma_1^2 + \psi_2^2\sigma_2^2} \quad i = 1, 2. \quad (7)$$

Component Share (CS) Measure:

Gonzalo and Granger (1995) propose a way of decomposing the vector of non-stationary series Y_t into permanent component f_t (non-stationary series) and transitory (stationary) component $\hat{Y}_t$ where the identification of these components is achieved by assuming that (*i*) the permanent component is a linear function of the original series and that *(ii)* the transitory component does not Granger cause the permanent component in the long run. The permanent component f_t (under linearity condition) can be written as

$$f_t = \mu^T Y_t, \quad (8)$$

where μ is an $n \times 1$ permanent coefficient vector which can be shown to be orthogonal to the adjustment coefficient vector α, i.e., $\mu = \alpha_\perp$. From equation (4), it is clear that ψ and μ differ by a scalar multiple. The CS measure depends on the elements of μ. For a two series case, the CS for market i (CS_i) can be given as follows:

$$CS_i = \frac{\psi_i}{\psi_1 + \psi_2}, i = 1, 2. \quad (9)$$

There are some arguments regarding the interpretation of these measures. Grammig and Peter have used the volatility inclusions to identify a unique information share [23]. We have observed that Hasbrouck IS measure is sensitive to the variables' order and not unique when price innovations across markets are correlated. The contribution of a commodity market to price discovery is its information share. Information share (IS) is the proportion of efficient price innovation variance [4].

Research Design

1. Our analysis begins with performing the augmented Dickey-Fuller (ADF) to examine the price series' unit root.

2. It is essential to check the cointegration assumption for both the price series. We have used the Johansen Cointegration Test in our analysis.

3. After confirming that unit root exists in each price series and also that the two series are cointegrated, we set up a VAR model for both the price series and use AIC to determine the AR order in the model.

4. Based on the estimated VECM model, we compute the information share measures. For the Hasbrouck measure, upper and lower bounds are

calculated along with the component shares, which will measure the price discovery.

Data

We have collected the commodity price time series from NCDEX (commodity market index) — which is commodity market data from NCDEX (will refer as "Commodity stock Price") and another is from Agmarknet data — which is wholesale market data for eNAM or Agmarknet (will refer as "Market Price"). We have used daily data of prices for all five commodities — cotton, maize, wheat, barley, and soybean from 2005 to 2019. Our criteria for selecting the commodity are (1) commodity should be listed in more than one market, (2) Volume or quantity of trade in the last ten years for that commodity, and (3) food grain is being selected considering their importance in the food basket, (4) we have not considering the storable or non-storable categories of the commodities and last (5) also, not categorizing based on the "seasonal" and "non-seasonal" commodity. As an assumption for the study, we assume that the APMC mandi location, the operational cost, and the commissions do not impact commodity prices. However, these may influence the decision of farmers for choosing a marketplace for trade. India electronic trading portal (eNAM) was introduced on 14th April 2016, hence we are calling these periods pre-eNAM and post-eNAM. Later from the analysis perspective, we have converted these time series into log returns. We have not used the 2020 data to exclude the impact of the COVID-19 pandemic as it has a significant effect on various monetary aspects.

Data Analysis and Results

1. In our analysis, we have used the following tests
2. Unit Root Test and Lag Length Selection
3. Johansen Cointegration Test
4. Vector Error Correction Model (VECM)
5. Information Share (IS)
6. Component Share (CS)

Regarding stationary, a time series is called stationary if it doesn't have a trend or seasonal effects. Statistically, we need to check the mean and variance to identify if a series is stationary. As the first step for the commodity stock and market prices, the standard t-test fails to reject the null hypothesis that the differenced price series has a zero mean. Then we performed the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root tests listed in *Table 2*.

Later, we used the Johansen cointegration test to check the existence of the cointegrating relationship. We apply log-transformed prices in this test and the results are summarized in *Table 3*. Now we conclude that there exists a single cointegrating relationship between the commodity stock price and market prices as the hypothesis of having no cointegration is rejected at a 5% significance level and the hypothesis that there is one cointegration vector cannot be rejected.

Now, we applied the Hasbrouck method to get the information shared (IS) [4]. This method gives upper and lower bounds. The upper and lower bound of IS can be calculated by doing all the permutations for one or multiple markets. To get the upper bound of IS, we can place that's market price first and lower the last. Here, we automatically calculated the number of lags by referring to the Akaike information criterion (AIC). *Table 4* summarizes the results of Hasbrouck information share (IS) and component share (CS).

The results show that the commodity stock prices lead the price discovery for all the markets except cotton. This trend is visible in both pre and post eNAM time frames. Maize and soybean commodity stock prices dominates the price discovery during post eNAM, more precisely.

We also checked the importance of various trading components summarized in *Table 5* and respective regression results are shown in *Table A1* (*APPENDIX*).

We can see that max price on that day is the most significant component of trade across the commodities. It is a component of market prices but also has a significant impact on commodity stock prices. Overall there is no significant impact of commodity stock price on market prices but min price of wheat from market prices has some impact on commodity stock price. We also checked the Granger causality for the selected agricultural commodities and found that all the commodities have a bidirectional causality relationship.

CONCLUSION

Overall, the commodity index market dominates the price discovery for all agri commodities except cotton. For cotton, market prices lead to price discovery, which is the same for pre and post eNAM periods. Although all three measures are

Table 2

Unit-Root Test Results

	ADF (first difference)		PP (first difference)	
Commodity	**Intercept**	**Intercept & trend**	**Intercept**	**Intercept & trend**
Barley MP	–2.060	–3.700**	–2.790	–14.283***
Barley CSP	–1.902	–4.267***	–1.783	–5.843***
Cotton MP	–1.001	–46.920***	–1.809	–3.636**
Cotton CSP	–1.451	–59.069***	–1.990	–59.068***
Maize MP	–1.642	–44.215***	–2.124	–4.179***
Maize CSP	–1.466	–54.039***	–2.004	–54.962***
Soybean MP	–1.786	–44.262***	–2.094	–81.267***
Soybean CSP	–2.768	–55.693***	–1.991	–56.037***
Wheat MP	–0.997	–29.792***	–2.354	–39.818***
Wheat CSP	–0.668	–15.768***	–2.293	–19.674***
Critical values				
1% level	–3.960635			
5% level	–3.411076			
10% level	–3.127359			

Source: author's analysis.

Note: MP – market price; CSP – commodity stock price. The table contains the t-statistics of the ADF & PP tests results, where *** and ** indicate the significance of t-statistics at 1% and 5% level of significance, respectively.

Table 3

Johansen's Cointegration Test Results

	Number of Cointegrating Vectors			
	None		At most one	
	Max. Eigenvalue	**Trace**	**Max. Eigenvalue**	**Trace**
Barley	22.342*	23.338*	0.996	0.996
Cotton	24.884*	25.934*	1.050	1.050
Maize	32.800*	36.151*	3.351	3.351
Soybean	65.275*	66.436*	1.161	1.161
Wheat	51.829*	53.612*	1.783	1.783

Source: author's analysis.

Note: MP– market price; CSP – commodity stock price. Trace test indicates 1 cointegrating eqn(s) at the 0.05 level. Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level. ** MacKinnon-Haug-Michelis (1999) p-values.

Table 4

Results of information share measure

			Hasbrouck Information Share (IS)		Component Share (CS)
			Original ordering	Reversed ordering	
Barley	Overall	MP	0.347	0.092	0.181
		CSP	0.653	0.908	0.819
	Pre eNAM	MP	0.321	0.065	0.158
		CSP	0.679	0.935	0.842
	Post eNAM	MP	0.012	0.001	0.012
		CSP	0.988	0.999	0.988
Cotton	Overall	MP	0.656	0.656	0.787
		CSP	0.344	0.344	0.213
	Pre eNAM	MP	0.901	0.901	0.893
		CSP	0.099	0.099	0.107
	Post eNAM	MP	0.753	0.760	0.769
		CSP	0.247	0.240	0.231
Maize	Overall	MP	0.317	0.282	0.270
		CSP	0.683	0.718	0.730
	Pre eNAM	MP	0.046	0.036	0.101
		CSP	0.954	0.964	0.899
	Post eNAM	MP	0.847	0.787	0.561
		CSP	0.153	0.213	0.439
Soybean	Overall	MP	0.055	0.212	0.216
		CSP	0.945	0.789	0.784
	Pre eNAM	MP	0.056	0.215	0.221
		CSP	0.944	0.785	0.779
	Post eNAM	MP	0.020	0.010	0.050
		CSP	0.980	0.990	0.950
Wheat	Overall	MP	0.342	0.002	0.056
		CSP	0.658	0.998	0.944
	Pre eNAM	MP	0.725	0.105	0.398
		CSP	0.275	0.895	0.602
	Post eNAM	MP	0.015	0.043	0.142
		CSP	0.985	0.957	0.858

Source: author's analysis.

Note: MP — market price; CSP — commodity stock price.

Table 5

Importance of trading components

Commod-ity	Price series	Quantity Traded on that day	Traded Value (INR-In lacs)	Number of trades	Quantity Arrivals on that day	Min Price on that day	Max Price on that day	Closing Price on that day
Barley	MP	13.1%	2.3%	16.9%	5.6%	2.1%	25.8%	34.1%
	CSP	28.0%	24.0%	7.0%	1.4%	16.5%	5.4%	17.6%
Maize	MP	1.9%	3.5%	1.2%	0.1%	48.3%	43.2%	1.9%
	CSP	11.6%	4.7%	8.6%	1.1%	11.5%	54.7%	7.9%
Soybean	MP	0.8%	0.7%	0.3%	0.5%	26.6%	69.8%	1.4%
	CSP	2.1%	5.3%	3.0%	5.2%	4.4%	61.2%	18.7%
Wheat	MP	9.7%	13.3%	7.0%	2.3%	5.9%	36.4%	25.4%
	CSP	7.1%	12.8%	6.0%	0.9%	30.9%	22.5%	19.8%
Cotton	MP	1.7%	1.3%	0.6%	0.2%	35.1%	60.9%	0.2%
	CSP	31.1%	11.5%	15.4%	0.3%	7.3%	22.0%	12.4%

Source: author's analysis.

Note: MP — market price; CSP — commodity stock price.

showing similar outcomes, CS shows that the price discovery by commodity stock prices compared to IS is significantly higher for barley and wheat. For cotton, market prices are leading the price discovery significantly. Considering the pre-eNAM period, all the commodities follow similar trends except wheat & cotton. For this period only, with the original order (market price commodity stock price), the market price is leading the price discovery while commodity stock prices are taking the lead in reversing the order. CS shows the commodity markets taking the lead. For post eNAM, market prices lead the price discovery for maize and cotton. This domination of commodity stock prices is significant for IS. To summarize the results, the commodity stock prices are leading the price discovery for all the five commodities except cotton based on all measures of information flow used. But the strength of this dominance has been varying during pre eNAM and post eNAM. It is interesting to see that maize is leading post eNAM. As one of the prime objectives, eNAM generates some price discovery, and maize is one of the examples. In contrast to maize, soybean price discovery is dominated by commodity stock prices during post-eNAM, which is significantly higher as suggested by the CS measure. One important thing which needs to consider here is that commodity stock prices are represented at the overall national level. Still, market prices are more particular to the state or domestic markets. Conclusions from this study can be used to understand the information flow and would be helpful to academicians, practitioners, policymakers, or business players of commodity markets.

REFERENCES

1. Goyal A. Information, direct access to farmers, and rural market performance in Central India. *American Economic Journal: Applied Economics*. 2010;2(3):22–45. DOI: 10.1257/app.2.3.22
2. Ceballos F., Hernandez M.A., Minot N., Robles M. Grain price and volatility transmission from international to domestic markets in developing countries. *World Development*. 2017;94:305–320. DOI: 10.1016/j.worlddev.2017.01.015
3. Man K., Wang J., Wu C. Price discovery in the US treasury market: Automation vs. intermediation. *Management Science*. 2013;59(3):695–714. DOI: 10.1287/mnsc.1120.1559

4. Hasbrouck J. One security, many markets: Determining the contributions to price discovery. *The Journal of Finance*. 1995;50(4):1175–1199. DOI: 10.1111/j.1540–6261.1995.tb04054.x.
5. Tricase C., Amicarelli V., Lamonaca E., Rana R.L. Economic analysis of the barley market and related uses. In: Tadele Z., ed. Grasses as food and feed. London: IntechOpen; 2018:25–46. DOI: 10.5772/intechopen.78967
6. Agarwal D.K., Billore S.D., Sharma A.N., Dupare B.U., Srivastava S.K. Soybean: Introduction, improvement, and utilization in India — Problems and prospects. *Agricultural Research*. 2013;2(4):293–300. DOI: 10.1007/s40003–013–0088–0
7. Vollmer T., Herwartz H., von Cramon-Taubadel S. Measuring price discovery in the European wheat market using the partial cointegration approach. *European Review of Agricultural Economics*. 2020;47(3):1173–1200. DOI: 10.1093/erae/jbz040
8. Lien D., Shrestha K. Price discovery in interrelated markets. *The Journal of Futures Markets*. 2014;34(3):203–219. DOI: 10.1002/fut.21593
9. Ahumada H., Cornejo M. Out-of-sample testing price discovery in commodity markets: The case of soybeans. *Agricultural Economics*. 2016;47(6):709–718. DOI: 10.1111/agec.12267
10. Dimpfl T., Flad M., Jung R.C. Price discovery in agricultural commodity markets in the presence of futures speculation. *Journal of Commodity Markets*. 2017;5:50–62. DOI: 10.1016/j.jcomm.2017.01.002
11. Ates A., Wang G.H.K. Information transmission in electronic versus open-outcry trading systems: An analysis of U.S. equity index futures markets. *The Journal of Futures Markets*. 2005;25(7):679–715. DOI: 10.1002/fut.20160
12. Dolatabadi S., Nielsen M.O., Xu K. A fractionally cointegrated VAR model with deterministic trends and application to commodity futures markets. *Journal of Empirical Finance*. 2016;38(Pt. B):623–639. DOI: 10.1016/j.jempfin.2015.11.005
13. Ahumada H., Cornejo M. Explaining commodity prices by a cointegrated time series-cross section model. *Empirical Economics*. 2015;48(4):1667–1690. DOI: 10.1007/s00181–014–0827–5
14. Baffes J., Ajwad M.I. Identifying price linkages: A review of the literature and an application to the world market of cotton. *Applied Economics*. 2001;33(15):1927–1941. DOI: 10.1080/00036840010023788
15. Liu Q., An Y. Information transmission in informationally linked markets: Evidence from US and Chinese commodity futures markets. *Journal of International Money and Finance*. 2011;30(5):778–795. DOI: 10.1016/j.jimonfin.2011.05.006
16. Figuerola-Ferretti I., Gonzalo J. Modelling and measuring price discovery in commodity markets. *Journal of Econometrics*. 2010;158(1):95–107. DOI: 10.1016/j.jeconom.2010.03.013
17. Baillie R.T., Booth G.G., Tse Y., Zabotina T. Price discovery and common factor models. *Journal of Financial Markets*. 2002;5(3):309–321. DOI: 10.1016/s1386–4181(02)00027–7
18. Gonzalo J., Granger C.W.J. Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*. 1995;13(1):27–35. DOI: 10.2307/1392518
19. Karabiyik H., Narayan P.K., Phan D.H.B., Westerlund J. Islamic spot and index futures markets: Where is the price discovery? *Pacific-Basin Finance Journal*. 2018;52:123–133. DOI: 10.1016/j.pacfin.2017.04.003
20. Avino D., Lazar E,. Varotto S. Time varying price discovery. *Economics Letters*. 2015;126:18–21. DOI: 10.1016/j.econlet.2014.09.030
21. Kapar B., Olmo J. An analysis of price discovery between Bitcoin futures and spot markets. *Economics Letters*. 2019;174:62–64. DOI: 10.1016/j.econlet.2018.10.031
22. Bohl M.T., Siklos P.L., Stefan M., Wellenreuther C. Price discovery in agricultural commodity markets: Do speculators contribute? *Journal of Commodity Markets*. 2019;18:100092. DOI: 10.1016/j.jcomm.2019.05.001
23. Grammig J., Peter F.J. Tumbling titans? The changing patterns of price discovery in the U.S. equity market. 2018. URL: http://wp.lancs.ac.uk/finec2018/files/2018/09/FINEC-2018–017-Joachim-Grammig.pdf
24. Aggarwal N., Thomas S. When do stock futures dominate price discovery? *SSRN Electronic Journal*. 2011. DOI: 10.2139/ssrn.1916072
25. Putniņš T.J. What do price discovery metrics really measure? *Journal of Empirical Finance*. 2013;23:68–83. DOI: 10.1016/j.jempfin.2013.05.004

26. Narayan S., Smyth R. The financial econometrics of price discovery and predictability. *International Review of Financial Analysis*. 2015;42:380–393. DOI: 10.1016/j.irfa.2015.09.003
27. Hasbrouck J. Price discovery in high resolution. *Journal of Financial Econometrics*. 2021;19(3):395–430. DOI: 10.1093/jjfinec/nbz027
28. Balli F., Naeem M.A., Shahzad S.J.H., de Bruin A. Spillover network of commodity uncertainties. *Energy Economics*. 2019;81:914–927. DOI: 10.1016/j.eneco.2019.06.001
29. Hao J., He F., Liu-Chen B., Li Z. Price discovery and its determinants for the Chinese soybean options and futures markets. *Finance Research Letters*. 2021;40:101689. DOI: 10.1016/j.frl.2020.101689
30. Dimpfl T., Peter F.J. Nothing but noise? Price discovery across cryptocurrency exchanges. *Journal of Financial Markets*. 2021;54:100584. DOI: 10.1016/j.finmar.2020.100584
31. Yang J., Li Z., Wang T. Price discovery in Chinese agricultural futures markets: A comprehensive look. *The Journal of Futures Markets*. 2021;41(4):536–555. DOI: 10.1002/fut.22179
32. Shrestha K., Subramaniam R., Thiyagarajan T. Price discovery in agricultural markets. *American Business Review*. 2020;23(1):53–69. DOI: 10.37625/abr.23.1.53–69
33. Fassas A.P., Siriopoulos C. Intraday price discovery and volatility spillovers in an emerging market. *International Review of Economics & Finance*. 2019;59:333–346. DOI: 10.1016/j.iref.2018.09.008
34. Frijns B., Indriawan I., Tourani-Rad A. The interactions between price discovery, liquidity and algorithmic trading for U.S.-Canadian cross-listed shares. *International Review of Financial Analysis*. 2018;56:136–152. DOI: 10.1016/j.irfa.2018.01.005
35. Narayan P.K., Sharma S.S. An analysis of time-varying commodity market price discovery. *International Review of Financial Analysis*. 2018;57:122–133. DOI: 10.1016/j.irfa.2018.03.008
36. Wright J., Kim M.-K., Tejeda H.A., Kim H.-N. A tournament approach to price discovery in the U.S. cattle market. *Journal of Agricultural and Applied Economics*. 2021;53(1):21–36. DOI: 10.1017/aae.2020.26
37. Rout B.S., Das N.M., Rao K.C. Does commodity derivatives function effectively? A lengthy discussion. *IIM Kozhikode Society & Management Review*. 2021. DOI: 10.1177/2277975220985746
38. Karabiyik H., Westerlund J., Narayan P. Panel data measures of price discovery. *Econometric Reviews*. 2022;41(3):269–290. DOI: 10.1080/07474938.2021.1912973
39. Nair S.T.G. Price discovery and pairs trading potentials: The case of metals markets. *Journal of Financial Economic Policy*. 2021;13(5):565–586. DOI: 10.1108/jfep-06–2020–0139
40. de Blasis R. The price leadership share: A new measure of price discovery in financial markets. *Annals of Finance*. 2020;16(3):381–405. DOI: 10.1007/s10436–020–00371–3
41. Manogna R., Mishra A.K. Price discovery and volatility spillover: an empirical evidence from spot and futures agricultural commodity markets in India. *Journal of Agribusiness in Developing and Emerging Economies*. 2020;10(4):447–473. DOI: 10.1108/jadee-10–2019–0175
42. Fassas A.P., Papadamou S., Koulis A. Price discovery in bitcoin futures. *Research in International Business and Finance*. 2020;52:101116. DOI: 10.1016/j.ribaf.2019.101116
43. Li M., Xiong T. Do bubbles alter contributions to price discovery? Evidence from the Chinese soybean futures and spot markets. *Emerging Markets Finance and Trade*. 2019;55(15):3417–3432. DOI: 10.1080/1540496x.2019.1608178
44. Booth G.G., So R.W., Tse Y. Price discovery in the German equity index derivatives markets. *The Journal of Futures Markets*. 1999;19(6):619–643. DOI: 10.1002/(SICI)1096–9934(199909)19:6%3C 619::AID-FUT1%3E 3.0.CO;2-M
45. Chu Q.C., Hsieh W.G., Tse Y. Price discovery on the S&P 500 index markets: An analysis of spot index, index futures, and SPDRs. *International Review of Financial Analysis*. 1999;8(1):21–34. DOI: 10.1016/s1057–5219(99)00003–4
46. Harris F.H.deB., Jarrell S.L., McInish T.H., Wood R.A. Minority shareholder expropriation and asymmetric information flows in a global registered share: The saga of Daimler Chrysler. *SSRN Electronic Journal*. 2004. DOI: 10.2139/ssrn.628362
47. Shrestha K., Lee L.M.Q. Evaluation of information leadership share as price discovery measure. *SSRN Electronic Journal*. 2021. DOI: 10.2139/ssrn.3893392

APPENDIX

Table A1

Regerssion results for selected agricultural commodities

	Pre eNAM				Post eNAM			
Barley	**Coefficients**	**Std. Error**	***t***	**Sig.**	**Coefficients**	**Std. Error**	***t***	**Sig.**
(Constant)		7.968	47.491	0.000		11.420	0.792	0.428
Quantity Traded on that day	–0.899	0.002	–20.263	0.000	0.215	0.002	4.600	0.000
Traded Value (INR–In lacs)	0.772	0.012	20.734	0.000	–0.037	0.015	–0.945	0.345
Number of trades	0.225	0.020	6.784	0.000	–0.276	0.022	–8.411	0.000
Quantity Arrivals on that day	0.046	0.000	7.047	0.000	–0.091	0.000	–14.442	0.000
Min Price on that day	0.529	0.079	5.938	0.000	–0.034	0.088	–0.382	0.702
Max Price on that day	–0.175	0.075	–1.960	0.050	0.422	0.083	4.771	0.000
Closing Price on that day	0.564	0.013	38.839	0.000	0.557	0.016	38.839	0.000
Maize	**Coefficients**	**Std. Error**	***t***	**Sig.**	**Coefficients**	**Std. Error**	***t***	**Sig.**
(Constant)		6.117	14.559	0.000		1.964	3.250	0.001
Quantity Traded on that day	–0.307	0.001	–12.614	0.000	–0.021	0.000	–2.667	0.008
Traded Value (INR–In lacs)	0.124	0.004	7.022	0.000	0.038	0.001	6.907	0.000
Number of trades	0.228	0.009	8.930	0.000	–0.013	0.003	–1.611	0.107
Quantity Arrivals on that day	0.028	0.000	5.669	0.000	0.002	0.000	1.019	0.308
Min Price on that day	–0.304	0.075	–4.409	0.000	0.534	0.021	26.956	0.000
Max Price on that day	1.447	0.067	21.547	0.000	0.478	0.021	22.751	0.000
Closing Price on that day	–0.208	0.054	–3.902	0.000	–0.021	0.005	–3.902	0.000
Soybean	**Coefficients**	**Std. Error**	***t***	**Sig.**	**Coefficients**	**Std. Error**	***t***	**Sig.**
(Constant)		13.499	12.975	0.000		3.232	3.388	0.001
Quantity Traded on that day	–0.026	0.000	–1.428	0.153	0.008	0.000	1.720	0.085
Traded Value (INR–In lacs)	0.068	0.000	5.419	0.000	–0.007	0.000	–2.337	0.020
Number of trades	–0.038	0.003	–1.876	0.061	–0.003	0.001	–0.560	0.575
Quantity Arrivals on that day	0.066	0.000	21.713	0.000	–0.005	0.000	–6.201	0.000
Min Price on that day	–0.056	0.030	–2.122	0.034	0.273	0.005	60.418	0.000
Max Price on that day	0.778	0.053	14.489	0.000	0.717	0.006	116.03	0.000
Closing Price on that day	0.238	0.072	3.476	0.001	0.014	0.004	3.476	0.001

Wheat	Coefficients	Std. Error	*t*	Sig.	Coefficients	Std. Error	*t*	Sig.
(Constant)		8.467	10.843	0.000		6.104	19.378	0.000
Quantity Traded on that day	–0.242	0.000	–8.147	0.000	–0.170	0.000	–7.029	0.000
Traded Value (INR–In lacs)	0.435	0.004	10.773	0.000	0.232	0.003	6.987	0.000
Number of trades	–0.204	0.006	–7.196	0.000	–0.123	0.004	–5.294	0.000
Quantity Arrivals on that day	0.029	0.000	7.540	0.000	–0.041	0.000	–13.326	0.000
Min Price on that day	1.050	0.038	29.995	0.000	–0.102	0.032	–3.169	0.002
Max Price on that day	–0.763	0.036	–22.205	0.000	0.634	0.027	22.827	0.000
Closing Price on that day	0.673	0.020	36.192	0.000	0.443	0.011	36.192	0.000
Cotton	**Coefficients**	**Std. Error**	***t***	**Sig.**	**Coefficients**	**Std. Error**	***t***	**Sig.**
(Constant)		306.94	41.971	0.000		4.821	–10.084	0.000
Quantity Traded on that day	–1.826	0.007	–20.824	0.000	–0.018	0.000	–3.647	0.000
Traded Value (INR–In lacs)	0.676	0.026	12.944	0.000	0.014	0.000	4.962	0.000
Number of trades	0.906	0.132	10.759	0.000	0.007	0.002	1.481	0.139
Quantity Arrivals on that day	–0.019	0.003	–1.094	0.274	–0.002	0.000	–2.754	0.006
Min Price on that day	0.431	0.675	2.754	0.006	0.365	0.006	68.152	0.000
Max Price on that day	–1.294	0.885	–5.685	0.000	0.633	0.005	120.73	0.000
Closing Price on that day	0.729	1.291	2.254	0.024	0.002	0.000	2.254	0.024

Source: author's analysis.

Note: MP — market price; CSP — commodity stock price.

Table A2

Granger causality tests statistics for selected agricultural commodities

	Null Hypothesis	F-Statistic	Prob.	Direction	Relationship
Barley	CSP does not Granger Cause MP	17.2323	7.00E-17	Bi-directional	MP ↔ CSP
	MP does not Granger Cause CSP	16.5151	4.00E-16	Bi-directional	
Cotton	CSP does not Granger Cause MP	0.89476	0.04835	Bi-directional	MP ↔ CSP
	MP does not Granger Cause CSP	1.04827	0.03873	Bi-directional	
Maize	CSP does not Granger Cause MP	16.9235	1.00E-16	Bi-directional	MP ↔ CSP
	MP does not Granger Cause CSP	2.52123	0.0276	Bi-directional	
Soybean	CSP does not Granger Cause MP	102.032	9E-101	Bi-directional	MP ↔ CSP
	MP does not Granger Cause CSP	5.46742	5.00E-05	Bi-directional	
Wheat	CSP does not Granger Cause MP	9.83574	2.00E-09	Bi-directional	MP ↔ CSP
	MP does not Granger Cause CSP	10.0031	2.00E-09	Bi-directional	

Source: author's analysis.

Note: MP — market price; CSP — commodity stock price.

ABOUT THE AUTHORS

Arunendra Mishra — Research Scholar, Department of Food Business Management and Entrepreneurship, National Institute of Food Technology Entrepreneurship and Management, Sonipat (Delhi NCR), India
https://orcid.org/0000-0002-6070-1373
Corresponding author
arunendra.niftem@gmail.com

R Prasanth Kumar — PhD in Strategic Finance, Assistant Professor, Department of Food Business Management and Entrepreneurship,
National Institute of Food Technology Entrepreneurship and Management, Sonipat (Delhi NCR), India
https://orcid.org/0000-0001-5299-7701
prasanth@niftem.ac.in

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 02.11.2021; revised on 20.11.2021 and accepted for publication on 17.12.2021.
The authors read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-241-251
JEL F38, F13, K34

Implication of Goods and Services Tax (GST) Implementation in India on Foreign Trade

N. Singhal[a], **S. Goyal**[b], **S. Kumari**[c], **S. Nagar**[d], **A. Tyagi**[e]
[a] IIMT University, Meerut, India; [b, e] Amity University, Noida, India;
[c] Vidya Knowledge Park, Meerut, India; [d] G.L. Bajaj Institute of Technology and Management, Greater Noida, India

ABSTRACT

The **aim** of this research is to assess the impact of GST on India's foreign trade. The GST Bill is implemented to simplify India's complex tax system, allow commodities to move effortlessly across state borders, reduce tax evasion, enhance compliance, raise revenues, encourage growth, boost exports, and attract investors by making it easier to conduct business in India. The author uses the following methods of scientific research: Augmented Dickey-Fuller (ADF) for unit root tests, the Johansen-Juselius (JJ) for co-integration analysis and the Vector Error Correction (VECM) Model for short run and long run impact of GST on imports and exports from July 2017 to June 2021. The findings **concluded** that GST increases the exports of goods and services in both the long and short run. The VECM model's test statistics reveal that imports rose after the imposition of GST. Because of the emergence of a uniform national market and tax system, GST has simplified commercial operations in India. Future studies on the effects of the GST introduction can examine the impact of GST on foreign trade by state or commodities.

Keywords: goods and services tax; imports; exports; balance of trade; foreign trade; India

For citation: **Singhal N., Goyal S., Kumari S., Nagar S., Tyagi A. Implication of goods and services tax (GST) implementation in India on foreign trade.** ***Finance: Theory and Practice.*** **2022;26(3):241-251. DOI: 10.26794/2587-5671-2022-26-3-241-251**

1. INTRODUCTION

Goods & Services Tax is a comprehensive, multistage, destination-based tax (GST) that is levied on each stage where manufacturing and sales of goods/services take place. It is levied upon the consumption of goods and services. It has three subcategories:

- Central Goods and Services Tax (CGST): It is a tax levied on Intra State supplies of both goods and services by the Central Government.
- State Goods and Services Tax (SGST): It is a tax levied on Intra State supplies of both goods and services by the State Government.
- Integrated Goods and Services Tax (IGST): It is a tax levied on all Inter-State supplies of goods and/or services. IGST will be applicable on any supply of goods and/or services in both cases of import into India and export from India.

The Goods and Services Tax (GST) has revolutionized the way Indian businesses operate. The Bill is implemented to simplify India's complex tax system, allow commodities to do across state borders, reduce tax dodging, enhance compliance, raise revenues, encourage growth, promote exports, and encourage investment by making it easier to conduct business in India [1]. But GST has also bought some challenges because of its dual structure. India has adopted dual GST where centre and state collect taxes at their own level. It has made the entire structure of GST fairly complicated in India. The centre will have to coordinate with all states and union territories to collect and share tax. Moreover, states have also lost their autonomy in determining rates under GST.

The implementation of GST has transformed the way businesses operate in India and also have a substantial influence on foreign trade of commodities since it overhauled the framework of imports and exports duty and eliminated several indirect taxes and deductions. It also causes changes in the small-scale and medium-scale industries. As a result, the worldwide market's competitiveness has shifted. Trade is one of the

most important variables that contribute to a country's economic prosperity (imports-exports). As a result, it is critical to monitor international trade performance following the adoption of the GST [2].

Prior to the adoption of GST, the status of imports of goods and services was subject to a variety of state and centralized taxes, including customs duty, countervailing duty (similar to excise duty), and special extra charge (equal to value-added tax). All of these taxes have been replaced by the GST, which is a single integrated goods and services tax (IGST) [3]. However, some imports continue to be subject to basic customs duty, education cess, and other protective taxes, such as anti-dumping duty and safeguard duty, in addition to the IGST. GST is not to be levied on products or services exported from India. If the supply of commodities qualifies as an export out of India under the Place of Supply Rules, the transaction is classified as a «zero-rated supply.» The supplier must be permitted to export the goods/services tax-free and to claim the CGST/SGST and IGST credits paid on inputs and input services. If he or she is unable to use the credit, he or she may apply for a credit refund under Section 38 of the proposed Central GST Act, 2016.

In a nutshell, IGST will apply to both imports and exports. Exports will be zero-rated, and ITC will be refunded. Imports will be subject to IGST as well as Basic Custom Duty.

We have discussed the impact of GST on a few sectors relating to overseas commerce in the sections below.

Easy Business Doing

One element on which the sector has agreed is that GST, as a destination-based tax (where the tax is levied at the point of supply or consumption rather than the site of production), makes life simpler for businesses in India. Companies are no longer required to file tax returns with different departments, but instead, use a single web-based application to do so [4]. The country is eventually become a single market, with uniform pricing across states and efficient management, rendering our goods more attractive.

Increase Exports

Can exports be far behind if GST makes doing business easier? One apparent result of uniform taxes and cost reductions due to a reduction in time and money in shipping would be that 'Made in India' items would now be more cost-competitive in global markets [5]. Our exports were sinking under the prior tax regime because we also exported a significant percentage of taxes. Due to double taxes, indigenous producers were unable to capitalize. All of this is altered after the implementation of GST. And, ultimately, the country's export is increased. GST is not to be levied on products or services exported from India. If the supply of commodities qualifies as export out of India under the Place of Supply Rules, the transaction is classified as a "zero-rated supply".

The Revolution in Logistics

GST has completely transformed the logistics and transportation business. Existing studies have already highlighted that the simplified tax procedure will lower the cost of maintaining inventory by half because merchandise need not to be stacked up in multiple warehouses and elimination of toll booth delays.

The Goods and Services Tax (GST) has transformed India's diverse tax framework into a single uniform tax rate, transforming the country into a single marketplace. This allows for the seamless delivery of commodities across borders with drastically reduced transit times and increased demand for logistics services.

Table 1 compares the various fiscal and foreign trade parameters namely indirect taxes, indirect taxes as a percentage of GDP, import and export in the pre and post-GST era.

The data obtained from the official sources (see *Table 1*) over the years on indirect taxes reveals that tax collection in India is rising over the years, still India needs to increase its share of taxes in GDP because it has not been improved enough after GST implementation. The higher proportion of tax to GDP shows faster economic development as it enhances the government's ability to finance the expenditure. In 2020–2021, indirect taxes' share in GDP was about 12.36% of GDP while in 2016–2017 immediately before GST implementation it was 10.95%. Similarly, the

Table 1

Government's fiscal parameters pre and post GST implementation

Years	Indirect tax (₹ Crore)	Indirect tax as percentage of GDP	Exports (₹ Crore)	Imports (₹ Crore)
2010–2011	741348	9.41	1142922	1683467
2011–2012	871505	9.84	1465959	2345463
2012–2013	1036732	10.35	1634318	2669162
2013–2014	1119772	9.86	1905011	2715434
2014–2015	1217289	9.71	1896445	2737087
2015–2016	1466981	10.81	1716384	2490306
2016–2017	1662518	10.95	1849434	2577675
2017–2018	1856945	11.07	1956515	3001033
2018–2019	2032864	10.69	2307726	3594675
2019–2020	2161306	10.63	2219854	3360954
2020–2021	2441371	12.36	2154339	2909937

Source: Handbook of Statistics on Indian Economy for various years.

Notes: Data for 2019–2020 are Revised Estimates and data for 2020–2021 are Budget Estimates; Amount of Indirect tax are Gross tax collection.

export and import in 2020–2021 was Rs. 2154339 crore and Rs. 2909937 crores as compared to Rs. 1849434 crore and Rs. 2577675 crores in 2016–2017. However, due to coronavirus restrictions, the import and export growth is affected and the growth of foreign trade has slowed down which can be seen in Table 1. Export and import have reached Rs. 2307726 crores and Rs. 3594675 crores in 2018–19 and declined to Rs. 2154339 crore and 2909937 crores in 2021.

This study's contribution and novelty may be seen in numerous areas. From a theoretical and methodological standpoint, this study empirically tested the effect of GST on the import and export of goods and services in India that contribute to the study's novelty. The existing studies primarily belong to the impact of GST on the Indian economy and used only qualitative information. The studies on the impact of GST on foreign trade are not well explored which might be due to the nascent stage of GST implementation in India, which makes it difficult to deploy comprehensive quantitative data. Hence in this study, we have adopted monthly time series data from July 2017 to June 2021 to assess the impact of GST on import and export in India. The impact of GST on import and export is a contentious issue, and the available literature presents GST implications on the economy in general necessitating more research in this field. From a geographical standpoint, this is the first attempt to assess the impact of GST on foreign trade in India. It is a well-known fact that the GST is a remarkable indirect tax reform in India that transforms the way of doing business in India, which has seen a cumbersome tax structure over the years. Earlier studies are focused on the GST in developed and developing countries that have implemented GST decades back.

Section 2 includes a study of relevant literature. Section 3 describes the data and methods used, as well as the variables. Section 4 reports on the empirical results and discussion. Finally, Section 5 presents the paper's conclusions and underlines its implications.

2. LITERATURE REVIEW

This section will focus on a review of the extant literature related to GST before and after its implementation in India. India has

implemented GST in July 2017 which took India a significant step forward. It is the most significant tax change in decades. Many scholars disagree with the implementation of GST as a replacement for VAT and considered it an indirect tax reform because it subsumes various indirect taxes both at the central level and state level. Authors have used several approaches to determine the impact of GST on the Indian economy.

Leemput and Wiencek [6] explored the country's multi-layered taxes, such as the national and state indirect tax structures. According to the report, the new tax regime removed ten different forms of indirect taxes including state level VAT (however VAT is still applicable in some items such as petrol and alcohol), central excise duty, services tax, additional customs duty, surcharges, Octroi, etc., and combined them into a single tax system that applied throughout the country. The report described the tax rates imposed on various commodities and services traded in India. They examined the impact of GST using the foundational model of commerce and geography, which included all of the states in the nation. They envisioned India as a single entity comprised of 30 diverse states that trade agricultural and manufactured commodities both locally and globally. They calculated internal and international obstacles to assess the impact of GST. The report also discusses the link and differences between India's classical and contemporary tax systems. They studied the country's imports and exports to determine the impact of the tax reform via a trade model. According to the study, GST is intended to improve general Indian well-being and to be a comprehensive strategy in the sense that it would benefit all Indian states. According to the trade model, GST would result in real GDP gains of 4.2%.

The structure of GST was developed by the Empowered Committee, which includes state finance ministers and cabinet ministers because the structure of GST is dual. The tax is levied by the central and state governments in the form of CGST (levied by central government) and SGST (levied by state governments) and IGST (levied by central government but distributed between central and state governments) [7]. The article of Raj [7] shed some insight on the economist's and taxpayer's predictions. The author revealed that from the perspective of economists GST will boost exports by 3.2 to 6.3 percent because it moderates the cost and the taxpayer feels that it promotes neutrality, efficiency, clarity, simplicity, effectiveness, and justice. He has also discussed some of the disadvantages of GST, claiming that a goods and services tax would not increase growth but will increase consumer price inflation and that there appear to be some loopholes in the planned GST. Coordination between 29 states and 7 union territories to create such a tax scheme will be tough.

GST makes life simpler for Indian enterprises as a result of the creation of a common national market [8]. One apparent result of even taxation and cost-effectiveness due to reduced time and expenses in shipping would be that 'Made in India' items would now be more cost-competitive in global markets. It also emphasises that its implementation represents a cohesive tax structure that absorbs the majority of existing indirect taxes, resulting in increased output, more job possibilities, and a thriving GDP in the long run.

Some authors [8, 9] in early phases of GST implementation have revealed that if GST is implemented properly at the ground level, with tax exemptions for certain goods such as agricultural commodities, this will result in higher revenue at the Centre as the tax collection system becomes understandable, eliminating the problem of tax evasion and leading to economic growth, allowing Indians to reclaim the wealth they have lost within the country. Nayyar and Singh mentioned that GST would have an influence on all industries in India, including manufacturing, services, telecommunications, automobiles, and small and medium-sized enterprises (SMEs). One of the most significant tax reforms, GST, would link the whole country under a single taxing system rate. Experts expect that GST would enhance tax collections, boost India's economic development, and eliminate all tax barriers between the state and central governments [9].

Sahoo and Swain [10] emphasised the impact of GST after one year of its introduction. It emphasises that government officials and experts are now considering the need to make several changes in the existing GST architecture, such as removing the 28 percent tax rate and transitioning towards fewer tax slabs by merging 12 percent and 18 percent rates, and systematically bringing electricity, real estate, and petroleum products under its purview. It also streamlines submission by eliminating the obligation to submit so many returns at such short intervals, while simultaneously strengthening anti-evasion procedures.

Two recent studies by Deshmukh, Mohan & Mohan [11] and Neog & Gaur [12] have studied the effects of GST implementation in India. Deshmukh, Mohan & Mohan [11] presented a general macroeconomic analysis of the extent to which the adoption of GST has improved existing tax administration and the resultant general economic well-being of a democratic political economy like India in light of the situation-actor-process; learning-action-performance analysis framework for the case analysis. The analysis revealed that India has observed a tremendous increase in tax base vis-à-vis revenue collection. Yet, some efforts are desired to improve the low tax to GDP ratio, skewed GST payers base, negative stakeholders' perception of GST, and the evil of tax evasion. The other merits realized by the economy are presented as benefits to the consumers, MSMEs, improved ease of doing business ranking and fostering make-in-India and *AatmanirbharBharat* move by the government.

Neog & Gaur [12] highlighted GST is not the only policy that transformed the tax structure of India; other major policy changes have also taken place. Reforms are the ongoing process that takes place when needed. The Tax Reform Committee (TRC) report of 1991 is regarded as one of the most productive and structured policy recommendations of that time. At the state level, sales tax reform in the form of Value Added Tax (VAT) in 2005 becomes a fruitful policy initiative. Again, when the need arises one more tax reform has taken place in form of GST. However, the tax collections at both national and sub-national levels are still low as compared to the international standards.

The smooth functioning of GST is based on the well-developed infrastructure in a form of information system, a mechanism of sharing of GST between state and centre and well-defined rules and regulations. Hence, the government should be extremely clear that in order for GST to function successfully, information systems should be adequately built throughout India [13–15]. Moreover, the central government should assure the state governments that the GST is functioning well. Furthermore, every effort should be taken to cover all things under GST which have still been left out so that no item is left out of the GST preview; otherwise, the goal of adopting GST would be defeated. In India, certain items are not yet subject to GST such as petrol and alcohol.

Having cited the above literature, the researchers observed a lack of primary and secondary literature linking indirect tax reform in an economy and its impact on foreign trade. The studies cited primarily belong to the impact of GST on the economy and used only qualitative information, which was used to establish the viewpoint. However, due to the nascent stage of GST implementation in India, it becomes difficult to deploy comprehensive data. Hence in this study, we have adopted monthly time series data from July 2017 to June 2021 to assess the impact of GST on import and export in India.

3. RESEARCH METHODOLOGY

3.1. Description of variables and data sources

Imports of goods and services (IMPORT): Imports of goods and services is the total import of the country. The essential information on import and export is taken from the Department of Commerce, Ministry of Commerce and Industry, Government of India[1] from July 2017 to June 2021.

Exports of Goods and Services (EXPORT): Exports of goods and services are the total Export of the country. The data on export is taken from the Department of Commerce, Ministry of Commerce and Industry, Government of India[1] from July 2017 to June 2021.

[1] URL: https://tradestat.commerce.gov.in/meidb/default.asp (accessed on 10.02.2022).

Table 2

ADF Unit root tests results

Variables	Intercept		Trend and Intercept		None	
	Level	First Difference	Level	First Difference	Level	First Difference
Export	−5.097 (0.101)	−8.071 (0.000***)	−5.439 (0.210)	−8.000 (0.000***)	0.576 (0.837)	−8.112 (0.000***)
Imports	−2.624 (0.095*)	−8.240 (0.000***)	−2.607 (0.279)	−8.144 (0.000***)	0.335 (0.778)	−8.311 (0.000***)
GST	−29.99 (0.301)	−8.012 (0.000***)	−28.94 (0.057*)	−7.917 (0.000***)	−0.010 (0.673)	−8.113 (0.000***)
GDP	−2.594 (0.103)	−6.658 (0.000***)	−2.673 (0.251)	−6.583 (0.000***)	0.407 (0.797)	−6.708 (0.000***)
TO	−3.940 (0.090*)	−9.168 (0.000***)	−3.898 (0.190)	−9.063 (0.000***)	−0.278 (0.580)	−9.263 (0.000***)
ER	−1.517 (0.515)	−5.685 (0.000***)	−1.696 (0.737)	−5.728 (0.000***)	1.315 (0.950)	−5.535 (0.000)

Source: authors' estimations using STATA.

*Note:** , ** and *** denote significance at the 10%, 5% and 1% levels, respectively, P value in paratheses.

Goods and Services Tax (GST): Goods and Services tax consists of revenue collection from July 2017 to June 2021. The data on GST has been collected from the GST portal of the Government of India.[2]

Gross Domestic Products (GDP): GDP is the total output of the country which is considered as the current price of Gross Domestic Product in India. The data on GDP has been taken from FRED Economic data Series.[3]

Exchange Rate (ER): It is measured as the National currency to US dollar exchange rate for which data has been collected from FRED Economic data Series.[3]

Trade Openness (TO): It is calculated as exports plus imports as percent of GDP.

3.2. Statistical Techniques

The research used a Vector Error Correction (VECM) model to examine the impact of GST on India's foreign trade from July 2017 to June 2021. Before beginning the estimate method, the monthly data series for the study were submitted to a unit root pre-test. However, the argument for using the VECM model is also cointegrated. The VECM considers all variables to be a priori endogenous and so adjusts for the interplay of endogenous and exogenous factors. Exogenous variables might be integrated into the VECM model in some circumstances.

4. EMPIRICAL RESULTS AND DISCUSSION

Our empirical research seeks to provide light on the impact of GST on India's foreign trade. Since the implementation of GST, several articles have stated that GST would change the way businesses operate in India and would have a substantial influence on foreign trade of commodities since it altered the structure of imports and export taxation and eliminated many indirect taxes and exemptions. However, there has been no empirical study that reveals the influence of GST on foreign trade, leaving the question of whether GST supports or hinders imports and exports of goods and services unresolved. Hence. We use the VECM model to put the research topic to the test.

4.1. Unit root test:

To begin, Augmented Dickey-Fuller (ADF) root tests were used to determine if the

[2] URL: https://www.gst.gov.in/download/gststatistics (accessed on 10.02.2022).

[3] URL: https://fred.stlouisfed.org/tags/series?t=gdp%3Bindia%3Bquarterly (accessed on 10.02.2022).

Table 3

Johansen cointegration test

Estimated Models	Maximum rank	LL	Eigen Value	Trace Statistic	5% critical value
Imports	0	516.31	–	111.733	68.52
	1	543.51	0.693	57.33	47.21
	2	556.43	0.429	31.482	29.68
	3	567.77	0.389	8.803*	15.41
	4	570.58	0.114	3.190	3.76
	5	572.17	0.067		
Exports	0	502.85	–	108.26	68.52
	1	529.66	0.688	54.65	47.21
	2	542.53	0.428	28.91*	29.68
	3	552.57	0.353	8.82	15.41
	4	555.41	0.111	3.14	3.76
	5	556.98	0.066	–	–

Source: authors' estimations using STATA.

Note: *Acceptance of null hypothesis of no cointegration because Trace value is more than critical value.

series was stationary or not. The Schwarz Information Criterion (SC), Akaike Information Criterion (AIC), and Hannan-Quinn Information Criterion (HQ) were used to estimate the lag duration during the unit root tests, the results of which are accessible on demand. *Table 2* shows the ADF unit root test findings for the variables utilised in this investigation.

The results of the ADF tests performed on the variable levels demonstrate that the variables were not stationary, however their first-degree differences were, showing that the variable difference was stationary.

4.2. Co-integration test

Because all of the variables in the model were found to be stationary at the first degree, a co-integration analysis was carried out. The Johansen–Juselius (JJ) test was utilised in the study for co-integration analysis, and the findings are shown in *Table 3*. The Trace statistic and Eigenvalue are both employed in the cointegration test result to assess the number of cointegrating equations to the 5% (.05) critical value. In this study, we only give trace statistics; eigenvalues will be accessible upon request. If the trace value is smaller than the crucial value, the null hypothesis of no cointegration is accepted.

The results for imports reveal that there are three cointegrating equations at a 5% level of significance using the Trace statistic. As a result, at most three cointegrating equations exist, indicating the presence of long-run interactions between the variables in the system. In a nutshell, the null hypothesis of no cointegration is rejected in the model in favour of the alternative of cointegrating linkages.

The export statistics, on the other hand, demonstrate that there are two cointegrating equations at a 5% level of significance under the Trace statistic. As a result, at most two cointegrating equations exist, indicating the presence of long-run interactions between the variables in the system. However, because the prevailing result demonstrated the existence of long-run interactions among the variables in the system, the VECM rather than the VAR model estimate is required.

Table 4

The vector error correction model (VECM) test results for Imports

Variables in Logarithms	Coefficients	Standard Error	t statistics	Probability
Long–Run Scenario				
GST	0.927	0.130	–7.12	0.000***
GDP	0.249	0.300	–0.83	0.040**
TO	–0.759	0.359	2.11	0.035**
ER	0.655	0.189	–3.46	0.001***
Short-run Scenario (error correction)				
Cointegrating equation	–0.048	0.122	0.40	0.069*
GST	0.007	0.018	0.41	0.680
GDP	–0.648	0.488	–1.33	0.184
TO	–0.451	0.341	–1.32	0.187
ER	–2.427	1.387	–1.75	0.080*
Constant	0.006	0.008	0.70	0.487

Source: authors' estimations using STATA.

Note: *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.

Table 5

The vector error correction model (VECM) test results for Exports

Variables in Logarithms	Coefficients	Standard Error	t statistics	Probability
Long-Run Scenario				
GST	1.927	0.262	7.35	0.000***
GDP	2.777	0.606	–4.58	0.000***
TO	1.852	0.380	–4.86	0.000***
ER	1.209	0.723	–1.67	0.095*
Short-run Scenario (error correction)				
Cointegrating equation	–0.145	0.084	–1.73	0.043*
GST	0.35	0.025	1.38	0.016**
GDP	0.713	0.805	0.89	0.037**
TO	0.584	0.547	1.07	0.286
ER	–4.063	1.890	–2.15	0.032**
Constant	0.006	0.012	0.53	0.599

Source: authors' estimations using STATA.

Note: *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.

4.3. Vector Error Correction (VECM) Model

Based on the normal pre-testing of the variables of interest, the Johansen cointegration test result revealed the presence of a cointegrating equation in the model, which goes a long way toward confirming the existence of long-run linkages among the variables in the model. Furthermore, the lag order selection criterion revealed that Lag 2 is the proper or optimal lag selection based on the SC, AIC, and HQ. As a consequence, the VECM estimate is the best model for examining the short-run speed of adjustment from equilibrium; accordingly, the conclusion is presented and analysed as follows:

Table 4 and *5* show the VECM estimation results for imports and export respectively. The Johanson approach's expression of the cointegrating link confirmed the presence of long-run correlations among the variables of interest (GST and Foreign trade) in the research. However, the cointegrating finding from the Johansen result served as the foundation for calculating the VECM in order to determine the rate of adjustment from long-run equilibrium to short-run equilibrium [16].

Table 4 shows that in the long term, all factors, particularly GST, GDP, and ER, have a considerable positive influence on imports of goods and services, but TO has a significant but negative influence on imports. In the short run, the cointegrating equation's VECM estimate output shows that the previous year's divergence from long-run equilibrium is rectified in the present period at a rate of roughly 4% (see table 4). In the short run, none of the factors are important.

Table 5 shows that in the long term, all factors, notably GST, GDP, TO, and ER, have a favorable impact on the exports of products and services. While GST, GDP, and TO are statistically significant at 5%, ER is statistically significant at 10% in determining exports in India.

Table 5 also displays the result of VECM model for export in short run. In the short run, the cointegrating equation's VECM estimate output shows that the previous year's divergence from long-run equilibrium is rectified in the present period at an adjustment speed of roughly 14 percent. In the near run, a percentage shift in FD is associated with a 35 percent rise in exports on average ceteris paribus. In the near run, a percentage rise in GDP is associated with a 71 percent increase in exports, ceteris paribus, for the GDP coefficient. In the near run, a percentage rise in TO is connected with a 58 percent increase in exports on average ceteris paribus. Furthermore, in the near run, a percentage rise in ER is associated with a 4% decline in exports on average ceteris paribus. To summarise, all factors except TO are relevant in the short run.

5. CONCLUSION AND IMPLICATION OF STUDY

The level of development of the economic system depends on foreign trade [11, 17]. The strategic role of foreign trade in providing sustainability to the economic system is that foreign trade is based on the realization of competitive advantages by removing the boundaries and deepening the movement of labour. The introduction of GST in the Indian economy has resulted in a significant improvement in the level of trade and thereby strong economic system. Exports have been more comfortable since taxes on all levels of exports have been reduced, and imports have likewise gotten more convenient and inexpensive. The VECM model was used in the study to assess the impact of GST on imports and exports from July 2017 to June 2021. The findings indicated that GST increases the exports of goods and services in both the long and short run. The VECM model's test statistics revealed that imports rose after the imposition of GST. Because of the emergence of a uniform national market and tax system, GST has simplified commercial operations in India. One apparent result of even taxation and cost-effectiveness due to reduced time and expenses in shipping would be that 'Made in India' items would now be more cost-competitive in global markets.

The theoretical implications of the study are evident through its propensity to address the knowledge gap encompassing an empirical inquiry linking GST and foreign trade scenario in the Indian context. It also established that tax structure requires constant supervision and

amendments. The findings drawn from our study are also consistent with the earlier studies [3, 6] which conceptualized that GST will increase foreign trade.

The limitation of this study is that since the GST is implemented four years ago, a short time period of data has been included. Also, GST implementation in India is still in the introduction phase; therefore, not enough panel data could be generated to establish the relationship between GST tax reform and foreign trade state-wise and commodity-wise using econometric modeling. Future studies can show the relationship between the two indicators using quantitative or econometric modeling in a panel to ensure the broader generalizability of the findings. More so, the effectiveness of the GST implementation can also be measured by quantitatively comparing the set objectives with several fiscal indicators.

REFERENCES

1. Maheshwari T., Mani M. Benefits of goods and services tax implementation in India: An analytical hierarchy process approach. *Journal of Public Affairs*. 2020: e2578. DOI: 10.1002/pa.2578
2. Ojha R., Vrat P. Implications of goods and services tax reform on the make in India initiative: A system dynamics perspective. *System Research and Behavioural Science*. 2019;36(4):551–563. DOI: 10.1002/sres.2570
3. Bindal M., Gupta D.C. Impact of GST on Indian economy. *International Journal of Engineering and Management Research (IJEMR)*. 2018;8(2):143–148.
4. Pearl D. The policy and politics of reform of the Australian goods and services tax. *Asia & the Pacific Policy Studies*. 2016;3(3):405–411. DOI: 10.1002/app5.149
5. Bhattacharya G. Evaluation and implementation of GST in Indian growth: A study. *International Journal of Commerce and Management Research*. 2017;3(11):65–68.
6. Van Leemput E., Wiencek E.A. The effect of the GST on Indian growth. International Finance Discussion Paper Note. 2017. URL: https://www.federalreserve.gov/econres/notes/ifdp-notes/the-effect-of-the-gst-on-indian-growth-20170324.pdf
7. Raj R. Goods and service tax in India. *SSRN Electronic Journal*. 2017. DOI: 10.2139/ssrn.2946241
8. Nisa S. The impact of GST on India's foreign trade. *SSRN Electronic Journal*. 2017. DOI: 10.2139/ssrn.2954353
9. Nayyar A., Singh I.A. A comprehensive analysis of goods and services tax (GST) in India. *Indian Journal of Finance*. 2018;12(2):57. DOI: 10.17010/ijf/2018/v12i2/121377
10. Sahoo A., Swain A. Implementation and its impact of GST on industrial sectors: An analytical study. *International Journal of Innovative Research in Science, Engineering and Technology*. 2017;6(7):15173–15180. DOI: 10.15680/IJIRSET.2017.0607379
11. Deshmukh A.K., Mohan A., Mohan I. Goods and services tax (GST) implementation in India: A SAP-LAP-Twitter analytic perspective. *Global Journal of Flexible System Management*. 2022;23(2):165–183. DOI: 10.1007/s40171–021–00297–3
12. Neog Y., Gaur A.K. Tax structure and economic growth: A study of selected Indian states. *Journal of Economic Structures*. 2020;9:38. DOI: 10.1186/s40008–020–00215–3
13. Adhana D. Goods and services tax (GST): A panacea for Indian economy. *International Journal of Engineering and Management Research*. 2015;5(4):332–338.
14. Thowseaf M.A.S. A study on GST implementation and its impact on Indian industrial sectors and export. *International Journal of Management Research and Social Science (IJMRSS)*. 2016;3(2):27–30.
15. Singhal N., Goyal S., Singhal T. The relationship between insurance and economic growth in Asian countries: A regional perspective. *Macroeconomics and Finance in Emerging Market Economies*. 2021. DOI: 10.1080/17520843.2021.1957599
16. Singhal N., Goyal S., Singhal T. Insurance-growth nexus: Empirical evidence from emerging Asian markets. *Transnational Corporations Review*. 2020;12(3):237–249. DOI: 10.1080/19186444.2020.1756170
17. Yadigarov T.A. Econometric assessment of the level of development of balanced foreign economic relations in conditions of uncertainty. *Finance: Theory and Practice*. 2022;26(1):79–90. DOI: 10.26794/2587–5671–2022–26–1–79–90

ABOUT THE AUTHORS

Nikita Singhal — Assistant Professor, School of Commerce and Management, IIMT University, Meerut, India
https://orcid.org/0000-0002-0700-8086
Corresponding author
nikitagoyal.nikki@gmail.com

Shikha Goyal — Research Scholar, Amity University, Noida, India
https://orcid.org/0000-0002-7634-4403
shikhagoyal.sg94@gmail.com

Dr Sapna Kumari — Assistant Professor, Vidya Knowledge Park, Meerut, India
https://orcid.org/0000-0001-7504-4428
adharshila.sapna@gmail.com

Shweta Nagar — Assistant Professor, G.L. Bajaj Institute of Technology and Management, Greater Noida, India
https://orcid.org/0000-0002-4661-6949
ashweta5434@gmail.com

Ayushi Tyagi — Research Scholar, Amity University, Noida, India
https://orcid.org/0000-0002-7984-8893
tyagi.ayushi17@gmail.com

Author's declared contribution:
N. Singhal — defined the research problem, objectives of the study and developed the conceptual framework of the study.
S. Goyal — analyzed the literature, collected the data and conducted the analysis.
S. Kumari — compiled the tables, interpreted the result, discussed the results.
S. Nagar — wrote the conclusions of the research and implication of the study.
A. Tyagi — drafted the final paper and improve English grammar.

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 03.03.2022; revised on 30.03.2022 and accepted for publication on 17.04.2022.
The authors read and approved the final version of the manuscript.

ORIGINAL PAPER

DOI: 10.26794/2587-5671-2022-26-3-252-270
UDC 336.2(045)
JEL H22, H25, H26

Why did the Consolidated Taxpayer Group Regime Fail?

A.V. Korytin, **N.S. Kostrykina**
Russian Presidential Academy of National Economy and Public Administration (RANEPA), Moscow, Russia

ABSTRACT

The object of the study is the consolidated taxpayer group regime (CTG) introduced in Russia in 2012 for the largest groups of companies in which a group of companies is treated as a single entity for tax purposes. The aim of this paper is to investigate the reasons for the decision to cancel the CTG regime starting from 2023. The method of the study is an analysis of the available official reports, scientific publications and officials' statements on the topic. We have also compared the initial objectives of the CTG regime introduction with the criteria by which the regime was further evaluated. In the paper we show that the main reasons for criticizing and abolishing the CTG regime do not correspond to the initial goals of its introduction. In the result, we found that the reasons for criticism stem from the fact that the CTG mechanism does not always allow the Russian regions to achieve predictability of their budgets and prevent them from diminishing. In prospect, this experience should be taken into account when making the decision on the future of the CTG regime to avoid a disappointment of the 'losing' regions.

Keywords: consolidated taxpayer group regime (CTG); corporate profits tax; transfer pricing; regional budgets; budget federalism

For citation: **Korytin A.V., Kostrykina N.S. Why did the consolidated taxpayer group regime fail?** ***Finance: Theory and Practice.*** **2022;26(3):252-270. DOI: 10.26794/2587-5671-2022-26-3-252-270**

INTRODUCTION

Consolidated taxpayer groups (CTGs) in Russia are a progressive tax administration tool that allows fiscal authorities to interact with a group of interdependent companies as a single entity for tax purposes.

A consolidated taxation regime is applied in many advanced economies. Austria, Germany, Denmark, Italy, France, and the United States allow the consolidation of profits and losses within a taxpayer group. Australia, New Zealand, the Netherlands, and Japan go further, using the principle of tax unity of the consolidated group. A number of other developed countries provide interdependent companies with certain opportunities to move either only profits (Norway, Sweden, and Finland), or only losses (Great Britain, Ireland) [1, p. 81].

Foreign practices also show that the introduction of a consolidated taxation regime is not a simple political process. For example, Australia has come a long way from the first proposals for tax consolidation in 1975 to a decisive reform in 2002 that allowed the holding of many subsidiaries to be managed as a single taxpayer. The main efforts throughout this period were aimed at developing rules that exclude the possibility of aggressive tax planning within the framework of the consolidation regime [2].

Canada in 1985 considered the possibility of introducing federal rules for the carry-forward of losses within a group of companies, similarly to the experience of the UK. However, the law was not adopted due to the resistance of the provinces, especially those losing tax revenues [3, p. 131]. This idea was discussed again in 2010–2012, but meeting objections from the regional authorities, in 2013 the Canadian government refused to introduce such mechanisms, expressing its intention to continue coordinating these plans with the provinces.[1]

In 2016, the discussion (which began in 2001) about the consolidated taxation for transnational companies within the European Union developed with renewed vigor [4]. Consolidated profits and losses of interdependent companies and distribution of the consolidated tax base between the EU member states according to the formula (in proportion to three equilibrium indicators: headcount, assets, and revenue) are assumed. This approach is expected to prevent base erosion and profit transfer (BEPS project) within the EU [5]. However, the expected redistribution of tax revenues was the main obstacle to reforms, since the countries that are profit centers (Ireland, Luxembourg, Malta, and the Netherlands) will inevitably lose part of the budget revenues [6].

By 2008, there were grounds for applying the consolidated tax institution in the Russian Federation: ten sectoral inspectorates for the largest taxpayers were created in the structure of the Federal Tax Service, and the quality of work of tax authorities with this category of payers has seriously improved, tax collection for the largest taxpayers is at a relatively high level [7]. But the practical implementation of this institution has faced problems similar to those faced by Australia, Canada, and the EU.

The CTG regime in Russia began in 2012 as an experiment in relation to a number of very large holdings, and if the mechanism was successful, it was supposed to be extended to a wider range of companies.[2]

Although, following the results of the first year of the CTG regime, the Ministry of Finance of the Russian Federation announced the achievement of its objectives,[3] however, in

[1] Government of Canada. Jobs, Growth and Long-Term Prosperity — Economic Plan 2013. URL: https://www.budget.gc.ca/2013/doc/plan/budget2013-eng.pdf (accessed on 12.05.2022).

[2] Statement by the Deputy Minister of Finance of the Russian Federation S.D. Shatalov on December 7, 2010, at a conference on the development of the tax system in the Russian Federation, organized by the Association of European Businesses. Information is available on the official website of the Ministry of Finance of the Russian Federation. URL: https://minfin.gov.ru/ru/press-center/?id_4=31942-vyskazyvaniya_s.d._shatalova_informatsionnym_agentstvam_na_konferentsii_razvitie_rossiiskoi_nalogovoi_sistemy (accessed on 12.05.2022).

[3] "The main directions of the tax policy of the Russian Federation for 2014 and for the planning period of 2015 and 2016" (approved by the Government of the Russian Federation on May 30, 2013). URL: https://minfin.gov.ru/common/upload/library/2015/09/main/ONNP_2013–06–05.docx (accessed on 12.05.2022), and similar documents for subsequent periods. URL: https://minfin.gov.ru/ru/statistics/docs/budpol_taxpol/ (accessed on 12.05.2022).

subsequent years, the tone of the ministry's statements about this regime became more cautious. The main direction of criticism was the decline in revenues of the consolidated and some regional budgets as a result of the CTG regime. In the course of analyzing the reasons for this decline, a moratorium was announced on the creation of CTGs, as well as on the accession of new participants to the existing CTGs. During the period of the moratorium, a number of measures were introduced to prevent manipulation by enterprises in the distribution of the tax base, as well as limiting the possibility of reducing the tax base by offsetting losses. However, in the end, it was decided to eliminate the mechanism of consolidated tax: the existing agreements on the creation of CTGs will cease to be valid no later than January 1, 2023.[4]

AVAILABLE DATA LIMITATIONS

In order for the reader to understand what determines the choice of the research method underlying this article, we consider it important to make the following explanation.

The possibilities of analyzing the functioning of the CTG regime are limited by the available data, a list of which is given in *Table 1*.

Assessing the extent to which the tax base has been redistributed between the regions of Russia as a result of the introduction of CTG is not an easy task, since the size of the tax base attributable to a particular region depends on a large number of factors, in addition to the distribution formula. This problem cannot be solved with the help of data available to an independent researcher.

Firstly, the very indicator of the CTG tax base for income tax, even before its distribution among the constituent entities of the Russian Federation, is influenced by many factors, including changes in sales volumes, prices, the introduction of new tax benefits, the offset of tax losses, etc. Because of this, the comparison, for example, of the absolute values of income tax revenues from the CTG regime to the budgets of the subjects for the period 2012–2020 without factor analysis is unpromising.

Secondly, CTGs were formed gradually — some of the existing CTGs began to operate in 2012, some in 2013, and some in 2014 (further on, a moratorium on the creation of new CTGs was announced). At the same time, there is no information on the tax base of individual CTGs, there are only aggregated data on the tax base for all CTGs. In addition, the holdings, mostly at their own discretion, included and excluded organizations from CTG that met the criteria established by law,[5] and information about which organizations were included in the CTG in each of the tax periods is not public. Because of this, a comparison of the shares of subjects in revenues from CTG is also inconclusive. In addition, it should be taken into account that, at the initiative of the Moscow authorities, in 2012 amendments were adopted to the legislation on taxes and fees for CTGs formed by the Gazprom holding, which established a 5-year transition period and did not allow the organizations' simultaneous redistribution of income tax revenues paid by OAO Gazprom from the budget of Moscow to the budgets of other regions where the assets and employees of the respective consolidated group are concentrated.

Ideally, it would be necessary to evaluate the contribution of each of the CTGs to the tax base of each specific region, since they can be multidirectional. At the same time, it is necessary to compare the change in the tax bases of the CTG regime and their distribution across the regions of Russia with and without consolidation. However, such calculations can only be made by the taxpayers themselves or by the Federal Tax Service on the basis of data provided by the taxpayers.

In view of these limitations, we choose a research method in the form of analysis

[4] P. 3 Art. 3 of the Federal Law "On Amendments to Parts One and Two of the Tax Code of the Russian Federation" dated August 3, 2018, No. 302-FZ.

[5] At the same time, the Federal Law of November 24, 2014, No. 366-FZ, and the Federal Law of November 28, 2015, No. 325-FZ introduced restrictions that prevent the arbitrary formation of CTG.

Table 1

Main sources of information on the activity of consolidated tax groups

Source	Data description
The main directions of the tax policy of the Russian Federation	Contain a very brief assessment of the effectiveness of the CTG regime, and only for the period 2012–2014. There is not enough data to draw independent conclusions even for the available period
Reports of the Accounts Chamber of the Russian Federation	The publicly available report contains information for 2012–2013 only. Press release and transcript of the presentation of the 2012–2016 result report contain episodic information
Statistical tax reporting forms No. 5-KGN (CTG)	Available for 2012–2019, as well as 9 months of 2020 Contain aggregated data on the accrual of income tax on CTG

Source: compiled by the authors.

and criticism of sources, which will allow us to answer the question posed in the title of the article using the available data. We use aggregated CTG income tax assessments to calculate the volatility of the regional tax base.

AIM, HYPOTHESIS AND METHOD OF THE RESEARCH

The aim of this study is to test the validity of the decision to cancel the CTG regime, which is as follows:

1) determine the criteria for evaluating the effectiveness of the functioning of the CTG institution in official documents and scientific literature;

2) show how these criteria have changed from the moment the CTG regime was introduced until the decision to abolish this instrument;

3) check the validity of the estimates in the publications under consideration in terms of the analysis method used in them and the sufficiency of the data used.

The hypothesis of the study is that the main points of criticism of the consolidated tax mechanism, on which the decision to abolish the CTG regime is based, do not correspond to the CTG objectives when it was introduced. In particular, one of the goals of the introduction of the CTG regime, i.e. instant set-off of losses between its participants, began to be considered as the main reason for criticism.

The research method consists of the analysis and comparison of various sources of information on this issue close to the introduction of the CTG regime and during its functioning. Such sources of information include scientific publications, bills with accompanying documentation, statements by authorized officials, official reports (Main Directions of Tax Policy (ONNP), Reports of the Accounts Chamber of the Russian Federation), and minutes of public discussions in government bodies, in the public domain.

THE CTG GOALS

To determine the initial objectives that were set for the CTG regime, we analyzed the following documents and other sources of information:

1) an explanatory note to the draft law No. 392729–5,[6] which was supposed to introduce the CTG institution;

2) conclusion of the State Duma Committee on Budget and Taxes on the draft Federal Law No. 392729–5;

3) statements of authorized persons;

[6] Draft Law No. 392729–5 (archived) "On Amendments to Parts One and Two of the Tax Code of the Russian Federation in Connection with the Creation of a Consolidated Taxpayer Groups.". URL: https://sozd.duma.gov.ru/bill/392729–5 (accessed on 12.05.2022).

4) ONNP 2012–2016;[7]

5) a report of the Accounts Chamber on the results of the functioning of the CTG regime for 2012–2013;

6) press release of the Accounts Chamber on the results of the CTG activities for 2012–2016;[8]

7) a transcript of the meeting in the Federation Council on the topic "Results of the application of the institution of CTG for the period 2012–2017." December 12, 2017.[9]

Based on the results of a comprehensive analysis of these sources, it can be concluded that the following objectives were set for the CTG Institute.

Objective 1. Fair distribution of income tax among the constituent entities of the Russian Federation, in particular, the redistribution of profits from Moscow and St. Petersburg to regions where there are personnel and fixed assets of CTGs.

Objective 2. Summation of financial results between CTG participants, which provided instant compensation for losses.

Objective 3. Simplification of tax administration.

Objective 4. Exclusion from transfer pricing (TP) control of transactions between members of CTG.

Deputy Minister of Finance of the Russian Federation in 2000–2015 S. D. Shatalov explains in an interview that the objective of a fair distribution of income tax among the constituent entities of the Russian Federation is a priority from the point of view of the state: "In Russia, the income tax is arranged in such a way that almost nothing goes to the federal budget; this tax is actually regional, although it is federal. Accordingly, the budget into which this tax is paid is extremely important. Due to transfer pricing, the center of profit easily moves both within Russia and abroad. If it moves within a country, problems arise such as loss trading, where organizations agree on which of them will have a larger loss and which will have a smaller one. There are other possibilities as well. For example, it is very convenient for me to pay taxes to my beloved governor, because this beloved governor will give me other benefits, up to budget support, on occasion. We have come across this many times. Tax laws should not encourage such examples".[10]

The same idea is largely repeated in ONNP-2016: "... if before the CTG regime, the individual constituent parts of the holding could more arbitrarily pay income tax to the budgets of the constituent entities of the Russian Federation (using transfer pricing mechanisms), then after the formation of the group [CTG], the distribution of tax payments for income tax is determined by the legislation on taxes and fees, ... in proportion to the share of the number [*of personnel*] located in the region (payroll budget) and the value of fixed assets in general group indicators".

Auditor of the Accounts Chamber of the Russian Federation S. I. Shtogrin also explained in the framework of the round table in the Federation Council that the main goal of the CTG is a more equitable distribution of profits, so that it goes from the main profit centers — Moscow and St. Petersburg — to the regions in which it is generated.[11] The share of Moscow was 29.1% in 2010 and 29.4% in 2011. The share of St. Petersburg was 6.6% in 2010 and 5.7% in 2011. In other words, the two subjects received 35, 1% of all profits received by all organizations in Russia.

The explanatory note to the draft law also states that, since the CTG regime is a new

[7] ONNP-2012 refers to the document "The main directions of the tax policy of the Russian Federation for 2012 and the planned periods of 2013–2014." etc.

[8] The creation of a consolidated taxpayer group regime has led to a significant reduction in corporate profit tax receipts. URL: https://ach.gov.ru/news/sozdanie-konsolidirovannyh-grupp-nalogoplatelshikov-privelo-k-znachitelnomu-snizheniyu-postupleniya-32238 (accessed on 12.05.2022).

[9] Transcript of the round table meeting on the topic "Results of the application of the institution of CTG for the period 2012–2017". URL: http://council.gov.ru/media/files/UZUY 2ZKXh50i jwPeQ5x9Q3PS 5hq9H3qz.pdf (accessed on 12.05.2022).

[10] Statements by S.D. Shatalov to news agencies in the State Duma of Russia. The Ministry of Finance opposes the exclusion of domestic transactions from the control system when determining prices for taxation. RBC. 16.02.2010, Moscow. URL: https://minfin.gov.ru/ru/press-center/?id_4=31726 (accessed on 12.05.2022).

[11] Transcript of the meeting in the Federation Council on the topic "Results of the application of the CTG institution for the period 2012–2017" dated December 12, 2017.

object of administration for the tax authorities, at the first stage it is proposed to limit the circle of organizations that may be part of it.

PUBLICATIONS CONFIRMING THE ACHIEVEMENT OF THE OBJECTIVES

Evaluation of the CTG functioning in official reports

Based on the results of the analysis of the CTG activities for 2012, ONNP-2014 concluded about "the achievement of the goal set during their [*consolidated groups*] introduction — the economically justified distribution of corporate income tax among the constituent entities of the Russian Federation in proportion to the production base (the value of fixed assets) and the number of employees. Since the main losses fell on such profit centers as Moscow and St. Petersburg, and most regions won (in 2012, income in 65 out of 83 subjects),[12] the new distribution of corporate profit tax between regional budgets turned out to be more even. The ONNP-2015 indicates that 74.0% of the decrease in revenues in 2013 (47.2 billion rubles) falls on 4 constituent entities of the Russian Federation. In turn, the budgets of 63 regions received an increase in revenues in the same amount (47.4 billion rubles).

An analysis of the first results of the CTG functioning by the Accounts Chamber of the Russian Federation is presented in the report for the period 2012–2013. Although the report notes that two years is not enough time to conduct a comprehensive analysis of the impact of CTG on the budget revenues of the Russian Federation and the national economy as a whole, nevertheless, it draws conclusions about the achievement of these two out of three objectives (the issue of simplifying administration is not considered in the report).

Firstly, the existence of a consolidation mechanism allowed members of CTG to exercise the right to net profits and losses within the group. Secondly, more equitable redistribution of income tax has been carried out from the entities that are "profit centers" to the regions where the main production is concentrated. At the same time, the report notes that a number of constituent entities of the Russian Federation, in which the budgetary security does not exceed the average Russian level (Belgorod, Volgograd, Nizhny Novgorod, Samara and Perm regions, the Republic of Karelia), need targeted financial assistance due to the impact of the CTG activities.

Evaluation of the CTG functioning in the scientific literature

Evaluation of the CTG functioning in the scientific literature

In the scientific literature, after the CTG introduction in 2012, a number of works dwell on the advantages of the consolidated tax mechanism for business, tax authorities, regional budgets, and the economy as a whole referring to the disadvantages only the insufficiently widespread use of the mechanism[13] [8–10].

N.A. Kondrashova notes that the CTG regime allows large companies to avoid problems with controlled transactions, which is important for modern business, which is characterized by a holding structure [11].

Given the situation in the Irkutsk region, E.O. Gorbatenko concludes that the need to create a CTG mechanism is associated with the consolidation of national holdings absorbing regional enterprises [12]. M.A. Evnevich concludes that the CTG institution stimulates the registration of groups of companies, contributing to a more perfect tax control of such associations by the state [13].

According to the article by D.A. Babenko, the CTG introduction contributes to increasing

[12] We are not talking about an absolute increase or decrease in income from profit ax for the group as a whole — the calculations show how much the income of regional budgets would change if each specific participant was not included in CTG.

[13] High entry thresholds, which cut off a significant part of the business from CTG, were sometimes seen as an unfair advantage for large companies. Some saw the disadvantage that the consolidation only applies to income tax, not including VAT and property taxes. It was impossible not to notice that the consolidation of profits and losses led to the loss of part of budget revenues, but it is clear that it is impossible to reduce the tax burden without losses to the budget.

Table 2

The assessment of tax consolidation goals in scientific literature and official reports

Goals of the CTG regime	Positive appraisal	Critical appraisal
Goals of the CTG regime before its introduction		
Objective 1 — fair distribution of income tax among the constituent entities of the Russian Federation	Bill, 2011; ONNP-2014; ONNP-2016; Babenko, 2014b; Vitvitskaya, 2015; Bannova et al., 2016; Nikitin et al., 2016; Accounts Chamber, 2014; Kondrashova, 2013; Konovalova, 2016	Grinkevich, 2015; Nikitin et al., 2016; SF, 2017; Accounts Chamber, 2018; Ilyin, Povarova, 2019
Objective 2 — instant set-off of losses between the CTG participants	ONNP-2008; Kireeva, 2008; Kizimov and Shegurova, 2009; Bill, 2011; ONNP-2014; Babenko, 2014b; Vitvitskaya, 2015; Grinkevich, 2015; Bannova et al., 2016; Kondratieva and Shalneva, 2016; Kondrashova, 2013; ONNP-2015; Leontieva and Zaugarova, 2014; Malis, Grundel, 2015; Shatalov, 2015; Nikitin et al., 2016; ONNP-2016; Gorbatenko, 2015; Konovalova, 2016	Not found
Objective 3 — simplified tax administration		
Objective 4 — exclusion from the TP control perimeter of transactions between CTG participants		
Other goals		
	Achieved	**Failed**
The principle of tax neutrality	Kizimov and Shegurova, 2009; Kondratieva, Shalneva, 2016	–
Competitiveness of business at the global level	Babenko, 2014a; Grinkevich, 2015	–
Consolidated budget revenue stability	–	ONNP-2015; Malis, Grundel, 2015; SF, 2017; Accounts Chamber, 2018; Ilyin, Povarova, 2019
Predictability of corporate profit tax revenues to regional budgets	–	ONNP-2015; ONNP-2016; Leontieva and Zaugarova, 2014; Nikitin et al., 2016; Accounts Chamber, 2014; Malis, Grundel, 2015; Konovalova, 2016; Ilyin, Povarova, 2019

Source: compiled by the authors.

the competitiveness of large holdings at the international level and increases the investment attractiveness of the country as a whole [14].

Analyzing the first results of the CTG functioning (for 2012–2013) both from the point of view of business and from the point of view of the budget, D.A. Babenko comes to the conclusion that the applied mechanism for the distribution of tax revenues between the constituent entities of the Russian Federation better reflects the real contribution of each region to the growth of national income, stimulates regional authorities to increase

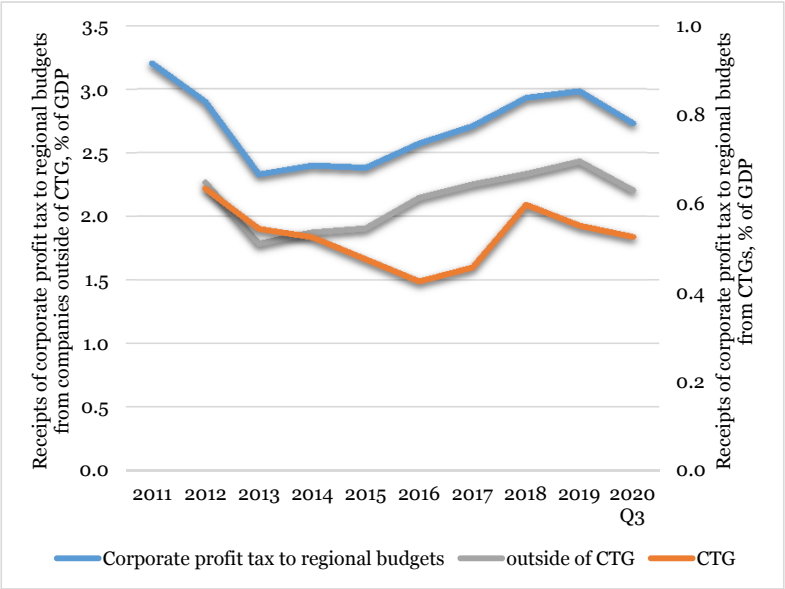


Fig. 1. **Receipts of profits tax by regional budgets from consolidated groups and other taxpayers in 2009–2019**
Source: authors' calculations based on data of Federal Tax Service.

their investment attractiveness, and is also more resistant to possible manipulations aimed at shifting the tax base [15].

Unfortunately, these works are not supported by sufficient empirical analysis. The authors base their conclusions on theoretical considerations based on the international experience of consolidated taxation and those expected results of the CTG functioning, which were formulated in official documents.

CHANGING THE EVALUATION CRITERIA OF THE CTG REGIME AS A BASIS FOR ITS CRITICISM

None of the publications we analyzed (see Table 2) concluded that any of the objectives 2–4 is not achieved or is not sufficiently achieved. The objectives to control transfer pricing, simplify administration or create preferences for large businesses do not cause any complaints.

At the same time, the literature gives a different assessment of the first objective — a fair distribution of income tax among the constituent entities of the Russian Federation. If initially, scientific articles supported the official position on the fairness of such a measure, then, starting from 2015, when ONNP-2016 appeared, this distribution mechanism has now been criticized in a number of works.

L. S. Grinkevich speaks of a conflict of interest "between individual subjects of the Russian Federation due to the lack of a scientifically based and fair mechanism for the redistribution of income tax" within the framework of CTG [16, p. 41]. K.M. Nikitin et al. explain that the formula for the distribution of the tax base, similar to that used for companies with separate divisions, reflects ideas about the factors of value creation that are inadequate to the realities of the 21st century and is a choice "in favor of the simplicity of the solution to the detriment of its validity" [17, p. 39]. However, the authors consider the distribution by production factors to be more equitable than by the financial result.

At the same time, statements that "compensation from the federal budget for the revenues of the regional budgets that fall as a result of the CTG regime" is needed [18, p. 80], should be regarded as an implicit statement that the previous distribution of the holding's tax base according to the financial results of its subsidiaries in the regions was more preferable.

In addition, in later sources, two new objectives are set for the CTG regime.

The first is the stability of budget revenues. The distribution of income tax within the framework of the CTG regime was initially positioned as more equitable since it was supposed to provide the regions with a large share of tax revenues from well-known profit centers — Moscow and St. Petersburg. Apparently, it was assumed that the rest of the regions would benefit not only relatively, but absolutely, i.e. will receive an increase in budget revenues. Although it is impossible to maintain the income of the consolidated budget at the same level, subject to the achievement of objective 2 (reducing the tax burden on business by adding profits and losses among CTG members), however, a compromise between these requirements is possible if the shortfall (due to the consolidation of profits and losses) incomes turn out to be small, and the effect of the redistribution of the CTG tax base compensates for the majority of regions for losses due to profit centers.

The second is to ensure the predictability of income tax revenues to the budgets of the subjects. If individual regions lose tax revenues due to the CTG activities, they still have the opportunity to save the regional budget through subsidies from the federal one. To do this, they need to correctly predict the magnitude of the decrease in tax revenues.

Both of these objectives are affected by factors such as intra-group *offsetting of profits and losses*, as well as the possibility of taxpayers manipulating *the distribution of the tax base* of the CTG. In addition, the predictability of tax revenues to the regional budget is difficult due to *the lack of data available* to the regional financial authorities.

Let's look at each of these problems in more detail.

LOSS OF INCOME DUE TO CONSOLIDATION OF PROFIT AND LOSSES

Before the introduction of the CTG regime A.S. Kizimov and T.A. Shegurova noted that "the introduction of the consolidation regime leads to a significant decrease in tax revenues," the compensation of which may require an increase in tax rates [19].

Many officials express concern about the decrease in tax revenues from CTGs. *Figure 1* shows that after 2015, the dynamics of income tax receipts from CTGs noticeably lag behind corporate profit tax receipts.

S.I. Shtogrin noted that the decrease or increase in the tax revenues of the constituent entities of the Russian Federation on income tax from CTG is affected not only by the institution itself but also by the economic performance of companies: a decrease in taxpayers' revenues, a decrease in prices in the foreign market, etc.[14]

The Ministry of Finance of the Russian Federation calculated the shortfall in the income of the constituent entities of the Russian Federation directly from the functioning of the CTG institution. ONNP-2016 explains the shortfall in income, firstly, by adding up profits and losses within the framework of the CTG regime, and also by the fact that "the tax base is often redistributed to the constituent entities of the Russian Federation, where lower income tax rates are established".[15] However, the report does not provide a quantitative assessment of the impact of these factors.

The second report of the Accounts Chamber on the results of the activities of the CTG for 2012–2016,[16] noted that the amount of shortfalls in revenues of the regional budgets of the Russian Federation due to the consolidation of profits and losses within the framework of the CTG for the period 2012–2016 amounted to 8, 16, 65, 126 and 78 billion rubles in each year, respectively.[17]

Based on the analysis of the shortfall in income from CTG at the level of the subjects of the Russian Federation, regional legislation, and official reporting of the subjects of the

[14] Transcript of the round table meeting on the topic "Results of the application of the CTG institution for the period 2012–2017" December 12, 2017.

[15] Draft Guidelines for the Budget Policy for 2016 and for the planning period of 2017 and 2018. URL: https://minfin.gov.ru/common/upload/library/2015/07/main/Proekt_ONBP_2016–18.pdf (accessed on 12.05.2022).

[16] Published only as a press release. The full text of the report is not in the public domain.

[17] Transcript of the round table meeting on the topic "Results of the application of the CTG institution for the period 2012–2017" December 12, 2017, p. 5.

Russian Federation, N. S. Kostrykina and A. V. Korytin [20] give an upper estimate of the amount of annual budget losses from the competition of the constituent entities of the Russian Federation for the CTG tax base. According to their calculations, they do not exceed 10 billion rubles, which is not critical for the consolidated budget. The paper also shows that the amount of shortfall in income tax (from all companies) due to the reduction in the rate by region fluctuates around 0.09% of GDP both before and after the introduction of the CTG regime. From this, a logical conclusion is made that the remaining losses are associated with loss compensation [20].

It should be noted that the offset of losses was initially positioned as one of the objectives of the CTG (Objective 2), and, as noted above, there is a contradiction between the objective not to reduce the income of the consolidated budget and Objective 2.

It can be assumed that the Ministry of Finance of the Russian Federation still had an idea of the permissible amount of shortfall in the income of the consolidated budget due to intra-group offsetting of losses, despite the fact that it was not announced when the CTG regime was introduced. Guided by this idea, the Ministry of Finance of the Russian Federation considered it appropriate to introduce temporary restrictions, in connection with which the CTGs can offset the losses of their unprofitable organizations in an amount not exceeding 50% of the profits of profitable participants.[18]

Thus, the problem of reducing tax revenues through the consolidation of profits and losses is partially solved in such a way that it returns the need to take into account the financial results of individual enterprises, while there are other approaches that are more appropriate within the framework of the concept of a consolidated taxpayer, for example, the use of increased tax rates on income for CTG participants or the abolition of the use of other tax benefits.

[18] Paragraph 1 of Art. 278.1 of the Tax Code of the Russian Federation as amended by Federal Law No. 401-FZ dated November 30, 2016.

LACK OF INFORMATION ABOUT THE CTG ACTIVITIES

A number of publications point to the lack of information on the CTG activities, which is necessary for the regional tax authorities to make an adequate forecast of tax revenues and, in case of their lack, to request advising subsidies from the federal budget.

The ONNP-2015 states that "the current CTG regime has led to difficulties in forecasting budget revenues at the level of an individual subject of the Russian Federation".

Based on these forms of tax reporting 5-KGNM for 2012–2013 N. I. Malis and L. P. Grundel conclude that the complexity of forecasting income from CTG at the regional level is a significant drawback, to eliminate which they propose "to establish the procedure for disclosing information on the profit and loss of each CTG participant" [21].

T. V. Konovalova brings clarity to the question of why the regional authorities are reporting difficulties in forecasting tax revenues from CTG. The only available reporting form 5-KGNM does not disclose what constitutes the taxable income of CTG and its share attributable to the region. As a result, regional authorities cannot even understand the reason for the decline in tax revenues [22].

The report of the Accounts Chamber for 2012–2013 stated that the regional tax authorities do not have sufficient data for an independent reliable calculation of the impact of the formation of CTG on the receipt of payments to the budget, since the law does not establish special requirements for the CTG administration during the period of tax audits. Information about the tax base that would have been formed in the absence of CTG is provided by the responsible CTG participants at the request of the central office of the Federal Tax Service of Russia on a voluntary basis and, in fact, is an expert assessment.

The lack of information necessary for the constituent entities of the Russian Federation to predict income tax was also noted in the

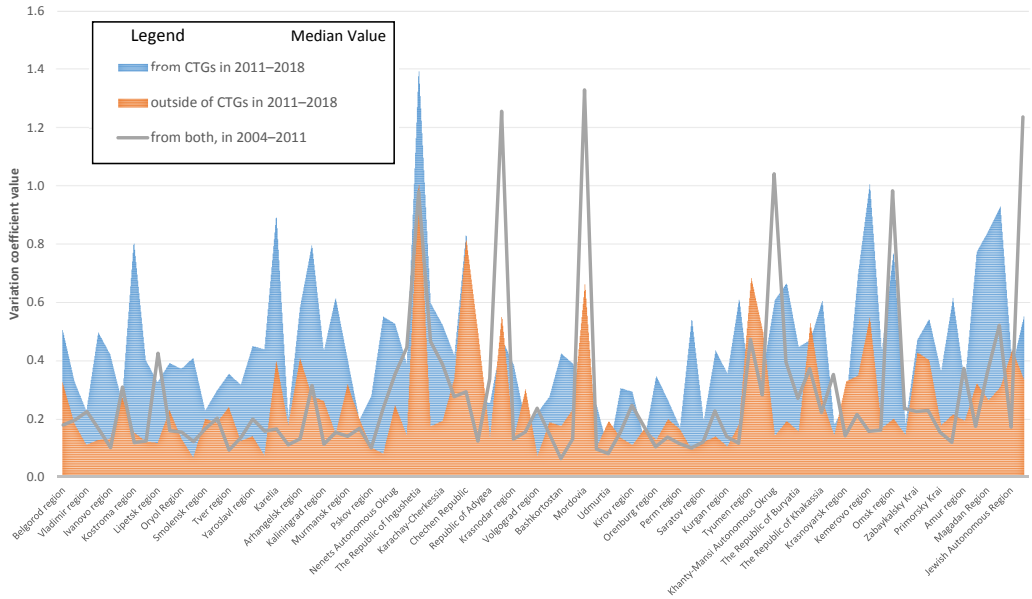


Fig. 2. **Variation coefficients of profit tax base for CTGs and organizations beyond CTG in 2011–2018 compared with the profit tax revenue variation from all organization in 2004–2011, by federal subjects**

Source: compiled by the authors with data from Federal Tax Service and Federal State Statistics Service.

Federation Council.[19] Considering that one of the main responsibilities of the subject may be to ensure the growth of tax revenues, limiting the availability of information resources can lead to problems in achieving this indicator.

As part of the ONNP-2016, the Ministry of Finance formulated provisions that allow the authorities of the constituent entities of the Russian Federation to obtain additional information from the tax reporting of CTG participants.

At the moment, the responsible CTG member is obliged to submit to the tax authority at the place of its registration information on the projected revenues from CTG for corporate profit tax to the budgets of the constituent entities of the Russian Federation in the current financial year, for the next financial year and planning period and on the factors affecting the planned receipts of corporate income tax.[20]

Let us analyze the volatility of the regional base for corporate income tax after the CTG introduction. As indicators of variability from CTG and enterprises outside CTG, we use the coefficient of variation calculated as $v(b)=\frac{\sigma(b)}{\mu(b)}$, where $\sigma(b)$ — the standard deviation of the series b, $\mu(b)$ — the average value of the series b.

To be able to compare the value of the tax base for different periods of time, regardless of economic fluctuations and inflation, we take b_i as the ratio of TB_i — the tax base of the i-th region to GDP — Russia's GDP for each year from 2004 to 2018. Then we calculate the variation of this indicator for each subject of the Russian Federation for the period 2004–2011,[21] according to forms 1-NM.

$$b_i = \frac{TB_i}{GDP}.$$

For 2012–2018, we will also calculate, using the data of the 5-PM and 5-KGN forms, the spread of the corresponding indicators for the tax base of the CTG and

[19] In particular, this opinion was expressed by representatives of the Belgorod and Kursk regions, the Krasnoyarsk Territory, the Komi Republic, and the Chairman of the Federation Council Committee on Budget and Financial Markets. Transcript of the round table meeting on the topic "Results of the application of the CTG institution for the period 2012–2017" December 12, 2017.

[20] Federal Law No. 302-FZ dated August 3, 2018, subparagraph 9, paragraph 3, art. 25.5.

[21] From a series of 2004–20112009 is excluded due to particularly strong crisis phenomena in the Russian economy, especially affecting the profits of companies.

enterprises that are not included in the CTG, and compare them.

Fig. 2 shows the result obtained. It shows that the variation in the CTG tax base (median value — 0.40) significantly exceeds the variation in the tax base of enterprises that are not included in CTG (median value — 0.19) in almost all regions of Russia. In our opinion, this is due to the predominance of CTGs in the commodity sector, whose financial indicators are tied to world commodity markets, which are characterized by high volatility. Consequently, it is more difficult for the budgetary authorities of the constituent entities of the Russian Federation to plan receipts from CTGs due to the volatility of their tax base.

At the same time, there is no noticeable difference in the variability of corporate profit tax receipts from all companies before 2012 (median value — 0.17) and after 2012 (median value — 0.16). This means that the introduction of the CTG regime in itself does not cause greater uncertainty regarding the tax revenues of the region, although it requires a completely different approach to forecasting regional budget revenues.

MANIPULATION OF THE TAX BASE DISTRIBUTION

The CTG regime has been criticized for the possibility of artificially distributing the holding's tax base among the constituent entities of the Russian Federation, including the one stimulated by the respective regions by establishing a reduced income tax rate. The manipulation became possible largely due to the arbitrary determination by the holdings of the composition of the CTG participants, as well as the arbitrary choice of the parameters of the distribution formula.[22]

In particular, due to the problem of manipulation, in 2014, changes were made to the legislation on taxes and fees aimed at suspending the process of creating new and expanding existing CTGs in 2015. When making this decision, it was assumed that during 2015 there would be an analysis of their activities and impact on budget revenues, as well as a further strategy in relation to such groups.

J.G. Leontieva and E.V. Zaugarova proposed to limit the possibility of determining the composition of CTGs on the principle of "all or none", identifying potential participants according to the control criterion of 50% + 1 share [23].

K.M. Nikitin and co-authors considered this problem in detail and developed a set of measures to solve it. They advocate the "all or none" principle, allowing them to be excluded if the volume of trade between dependent legal entities is extremely small. In addition, the authors believe to exclude all possibilities of manipulating the distribution of the tax base by setting strict rules for its calculation according to the existing formula [17].

V.A. Ilyin and A.I. Povarova support the introduction of strict restrictions on the determination of the CTG perimeter [18].

The subjects of the Russian Federation also made a proposal to limit the possibility of determining the composition of the CTG regime on the principle of "all or none" for organizations with a stake of 90% or more. In addition, they proposed to introduce a moratorium on the provision of corporate profit tax benefits directly by the subjects of the Russian Federation.[23]

Many of the proposals considered have been implemented. As part of the ONNP-2016, the Ministry of Finance proposed a number of changes to the CTG legislation, including limiting the possibility of changing the composition of participants in consolidated

[22] We are talking about the possibility of choosing between the indicator of the payroll budget and the average headcount. For example, the Atomenergoprom Group of Companies began to use the payroll instead of the average headcount, as a result of which the share of the Moscow budget immediately increased from 5.7% to 10.2%, Rostov — from 10% to 15%, and a number of constituent entities of the Russian Federation lost. LUKOIL switched from average headcount to payroll, as a result of which Moscow's share increased from 8.4% to 17%.

[23] As part of the proposals for the development of the CTG institute, submitted to the Ministry of Finance of the Russian Federation. Transcript of the round table meeting on the topic "Results of the application of the CTG institution for the period 2012–2017", p. 6–7.

groups. The distribution formula indicators are now not subject to change during the entire term of the agreement on the creation of CTG.[24] In addition, the subjects of the Russian Federation actually lost the opportunity to reduce the regional income tax rate.[25]

Although at the moment not all sources of manipulation in the CTG legislation have been eliminated, measures have already been taken. The experience of using this institution in other countries and the limited scale of this problem allow us to conclude that the problem of artificially shifting the base for the corporate income tax of CTG can be solved.

CRITICISM OF REDISTRIBUTION OF INCOME TAX BETWEEN REGIONS

The objective of a fair redistribution of the tax base within the CTG framework was criticized by a number of participants in the round table in the Federation Council on December 12, 2017, on the topic "Results of the application of the institution of a CTG regime in 2012–2017".

According to the results of the second report of the Accounts Chamber presented there, in 2016 compared to 2011, the share of Moscow in profit tax (from all organizations) decreased insufficiently — from 29.4% to 25.9%, and the share of St. Petersburg did not change at all — 5.7%. This allowed the speaker (S.I. Shtogrin) to conclude that the main objective of CTG — the redistribution of profits from the two main profit centers — Moscow and St. Petersburg — has not been achieved.[26]

As a result of the round table, proposals were put forward to change the distribution formula. One of the proposals was to distribute the tax base among the subjects not only in proportion to the value of fixed assets and the number of employees but also depending on the financial result of enterprises in the region.

Based on this discussion, we can conclude that the practical implementation of the CTG regime, which allows the creation of consolidated groups at the initiative of the taxpayer, in some cases leads to a violation of the interests of the regions.

If there is a very profitable enterprise in a disadvantaged region, then the proceeds from it can account for a significant share of budget revenues (for example, Karelsky Okatysh mine, which is part of the Severstal holding, formed 50% of corporate profit tax revenues to the budget of the Republic of Karelia). The inclusion of such an enterprise in CTG will lead to the distribution of its tax base among all constituent entities of the Russian Federation where CTG is present and will exacerbate the budgetary problems of this entity (in 2016, the share of Karelsky Okatysh mine in Severstal's profit was 12.7%, at the same time, its share in the number of employees was 5.7%, in the residual value of fixed assets — 2.5%, as a result, the share of tax revenues from the enterprise to the budget of the Republic of Karelia decreased to 5–10%).[27] Regardless of the validity of such a redistribution, it creates objective problems by changing the system of budgetary relations between the federation and the regions.

Another problem may arise when an enterprise that brings a stable income to a certain region is included in CTG and is not one of the largest in its CTG (for example, Orenburgneft in

[24] Changes were made by Federal Law No. 325-FZ dated November 28, 2015.

[25] From January 1, 2023, the constituent entities of the Russian Federation cannot introduce reduced rates for corporate profit tax other than those provided for by Chapter 25 of the Tax Code of the Russian Federation. The reduced rates for corporate profit tax established by the laws of the constituent entities of the Russian Federation, adopted before the date of entry into force of Federal Law No. 302-FZ of August 3, 2018, are subject to application by taxpayers until the expiration date of their validity, but no later than January 1, 2023.

[26] At the same time, S.I. Shtogrin explained that the high share of Moscow in 2016 was due to the fact that in 2016 Gazprom made refunds of income tax overpayments on advance payments from the budgets of other subjects. For comparison, in 2015 the share of Gazprom Group of Companies was only 28.1%. Despite the fact that in 2015 the Gazprom Group of Companies still used a slightly different procedure for redistributing profits between the constituent entities of the Russian Federation than the rest of the Group of Companies, according to which the Moscow budget received more than with the standard distribution formula.

[27] Transcript of the round table meeting on the topic "Results of the application of the CTG institution for the period 2012–2017", p. 16, 25.

Table 3

Evaluation of the results of the goal for CTG tax base reappointment from Moscow and St. Petersburg to the rest of federal subjects

No.	CTG holding company	Consolidated revenue in 2016, bln. rubles	Consolidated profit in 2016, bln. rubles	Evaluation of the result
1	Gazprom	5 966	997	Objective achieved. The share of Moscow in 2011 is 62.2%, and in 2015[1] – 28,1%
2	Lukoil	4 744	208	Objective achieved. The share of Moscow decreased from 19.9% in 2011 to 8.4% (17% after the group changed the average headcount indicator to the payroll)
3	Rosneft	4 134	201	Unable to assess*
4	Gazprom Neft	1 696	210	Unable to assess*
5	Surgutneftegaz	1 006	(62)	Objective not applicable**
6	X5 Retail Group	1 034	22	Information not available
7	Transneft	818	233	Information not available
8		700	93	Objective achieved. The share of Moscow decreased from 52.1% in 2011 to 5.7% (10.2% after the group changed the average headcount indicator to the payroll)
9	Atomenergoprom	580	106	Objective not applicable**
10	Tatneft	537	265	Objective achieved. The share of Moscow decreased, in 2011 it was 51.1%
11	NOVATEK	508	61	Objective not applicable**
12	NLMK	517	(13)	Information not available
13	EVRAZ	549	167	Information not available
14	Nornickel	392	102	Objective not applicable**
15	Severstal	316	26	Information not available
16	Megaphone	276	9	Unable to assess *

Note: * zero tax base, ** no tax base concentration in Moscow and St. Petersburg, *** we give the figure for 2015 due to the fact that, according to S.I. Shtogrin, the high share of Moscow in 2016 is due to refund of overpayment on income tax on advance payments from the budgets of other organizations, made by Gazprom, however, the value of this factor is not given, which makes it impossible to determine the share of Moscow without the influence of the overpayment return rate.

Sources: Accounts Chamber, 2014; RBC-500 rating in 2016, consolidated accounts of Rosatom and Gazprom Neft; Transcript of the round table "Results of the application of the institution of a consolidated taxpayer group in 2012–2017"; authors' analysis.

the Orenburg Region, which is part of the Rosneft structure). Then the income of this region begins to depend entirely on the financial result, which is formed not in it, but somewhere outside it (in this example, in Moscow and the regions of Northern Siberia). This radically changes the existing relations between regional authorities and business.

VALIDITY OF A NEGATIVE ASSESSMENT

According to the report of the Accounts Chamber (Accounts Chamber, 2014), a total of 17 CTGs operated in Russia, one of which worked with the responsible participant JSC RTKomm.RU for only 2 years (2012 and 2013) and was liquidated by the decision of its members.

Table 3 shows the names of the holdings that formed the remaining 16 CTGs, as well as the data contained in the report by S.I. Shtogrin, which allow assessing whether the objectives of redistributing the tax base of CTG to other entities have been achieved. In addition, we also provide the consolidated financial performance of holdings as a reference.

It is important to understand that the financial performance of the holdings and CTGs created by them are not equal due to legislative restrictions[28] and the decisions of the holdings themselves.[29] However, due to the lack of information on the financial results of individual CTGs and the perimeter of their entities, we use the consolidated financial performance of holdings formed by these CTGs. Such information provides insight into the scale of the holding and the maximum possible scale of the respective CTG.

Table 3 shows that four of the sixteen CTGs did not concentrate their tax base in Moscow and St. Petersburg, so the objective of redistributing the tax base from profit center regions to other subjects of the Russian Federation is not applicable to them.

The results of the redistribution of tax bases for three more CTGs cannot be assessed due to losses incurred in 2016 or in previous tax periods.

At the same time, the tax base of the largest CTGs formed by Gazprom, Lukoil, and Rosatom holdings has been significantly redistributed in favor of regions other than Moscow and St. Petersburg. It is also indicative that when choosing the average headcount indicator instead of the payroll in the formula for the distribution of the tax base,[30] the share of other regions was even higher.

There is no information about the five remaining CTGs in the transcript. It can be assumed that these regions were not mentioned due to the lower value of their tax base in the budgets of the regions. However, when considering the consolidated financial performance of the holdings, it becomes clear that this hypothesis is valid only for the CTGs formed by the holdings Megafon, X5 Retail Group, and, possibly, EVRAZ. As for the CTGs formed by the holdings of Norilsk Nickel and especially Transneft, we would like to see the results of the redistribution of their tax base.

In any case, the information contained in the report by S.I. Shtogrin, does not allow us to conclude that the objective of the CTG regime, which is to redistribute the tax base from Moscow and St. Petersburg to other regions, has not been achieved. On the contrary, the indicators of individual CTGs show the opposite.

In addition, it was noted at the round table that thanks to CTG, oil and steel companies, which were mainly part of the CTG regime left the tolling scheme — the creation of structures that can be located in any region of the Russian Federation and make a profit on a give-and-take basis. As a result, businesses began to build real, not artificial, production chains.[31]

In addition, if during the testing period of the new institution, CTG could form groups that declared a significant part of their profits in Moscow and St. Petersburg (i.e. those for which there was an understanding that their tax base is distributed unfairly across the constituent entities of the Russian Federation), then precedents similar to Karelsky Okatysh mine and Stoilinsky GOK would not have

[28] For example, for an organization to become a CTG member, the share of direct or indirect participation of the parent organization in it must be at least 90%.

[29] Holdings arbitrarily form the CTG perimeter and may not be included in the consolidated group of organizations, although they meet the conditions for entering the CTG perimeter.

[30] CTGs were formed by Lukoil and Rosatom holdings.

[31] Transcript of the round table meeting on the topic "Results of the application of the CTG institution for the period 2012–2017", p. 52–53.

happened. This is due to the fact that, as noted above, the NLMK and Severstal groups did not declare a significant share of their profits in Moscow and St. Petersburg. And the CTG formation by them led to the redistribution of profits among other, including subsidize, regions.

According to early documents and scientific publications, the CTG regime was intended to simplify the tax administration of large holdings, solving the problem of transfer pricing and the transfer of profits between the subjects. It seemed desirable to link the amount of tax revenues to the regional budget not with the amount of profit declared in the region, but with real indicators of economic activity: the cost of fixed assets and the size of the payroll budget. Preference in the form of adding up profits and losses was recognized as the goal of CTG. At the same time, a number of researchers suggested an increase in the profit tax rate as possible compensation.

According to the Deputy Minister of Finance in 2000–2015 S.D. Shatalov, "consolidation is not only economically justified but also contributes to a more equitable distribution of profit tax between regions". However, "the new institution did not appear during a period of economic growth, ... which exacerbated issues of interbudgetary relations also due to the fact that the losses of individual members of the group reduce the total profit of the entire group and, accordingly, the amount of tax to be distributed" [24].

Thus, based on the goals of the consolidated tax regime, declared before its introduction in 2012, we arrive at the following results:

1. The consolidated taxpayer group regime is a modern instrument for taxing large holdings when branches and subsidiaries within the same business have equal tax status [25]. Simplification of administrative costs is the main source of increased economic efficiency as a result of consolidated tax [26].

2. Lowering the tax base through profit and loss consolidation is a key feature of the CTG regime. Such an effect from the introduction of consolidated taxation was expected, and certain ways were proposed to prevent a decrease in tax revenues, for example, by raising the income tax rate.

3. The income tax base distribution formula for CTGs encourages regions to create jobs and invest in fixed assets. The new rules for participation in CTG prevent significant manipulation by businesses in the distribution of the tax base.

4. Issues of fiscal federalism, whether it be the difficulty of predicting the regional tax base and budget revenues or the reduction in income tax revenues in some subsidized regions, should be addressed separately. The provision of sufficient tax information to regional authorities to forecast revenues should be required by law.

5. The problem of subsidized regions must be solved with the help of measures of fiscal federalism. The abolition of the CTG institution will cause a new redistribution of the income tax base, which will also lead to problems for regional budgets. The losers will be those regions to which income tax was redistributed from Moscow and St. Petersburg, as well as a number of other regions that benefited from the redistribution under the current formula. In connection with the fears of the regions about the upcoming abolition of the CTG regime, the Ministry of Finance of the Russian Federation is forced to look for "how to ensure the stability of the tax base in the context of a changing income tax".[32]

CONCLUSIONS

The experience of consolidated taxpayer groups shows that changing the rules for the distribution of tax revenues in a federal state is not an easy task, requiring coordination of the interests of the federal center, businesses, and regional authorities. Initially, the objectives of the CTG institute included:

1) distribution of the income tax base between regions in proportion to production assets, and not to profit;

[32] The Federation Council hosted an "Open Dialogue" with the participation of the Minister of Finance of the Russian Federation. URL: http://council.gov.ru/events/chairman/130580/ (accessed on 12.05.2022).

2) unification of tax administration procedures of group members;

3) reduction of the tax burden due to the consolidation of profits and losses;

4) elimination of the need to comply with and control the transfer pricing rules.

At first glance, these tasks reflect the interests of all three parties — the federal authorities, businesses, and regions, however, the further implementation of the CTG regime has demonstrated the need for more careful consideration of the interests of the regions.

In the process of introducing the institution of consolidated tax on a trial basis, CTGs were formed at the initiative of taxpayers. If, on the contrary, the state, within the framework of the test period, ensured the formation of only those CTGs that declare profits in the regions — profit centers, the task of a fair distribution of profits between regional budgets could be completed to a greater extent. It would also avoid the consequences for which the CTG regime was most criticized — the situations with the Karelsky Okatysh mine (KGP, formed by the Severstal holding) and Stoilensky GOK (CTG, formed by the NLMK holding). Neither Severstal nor NLMK concentrated a significant part of their profits in Moscow or St. Petersburg, so the creation of CTG by these groups led to a redistribution of profits among other regions, including subsidized regions.

If the distribution of the income tax base between regions in accordance with production assets, and not the financial performance of enterprises as a whole, seems to be more equitable, then the redistribution of the tax base of individual holdings may cause problems, which are indicated in the examples above.

Although, based on the results of the CTG functioning, the tasks set were completed, in fact, the requirements for the power consolidation of the constituent entities of the Russian Federation were not taken into account:

1) transparency and predictability of changes in tax revenues to regional budgets;

2) maintaining the tax revenues of subjects at a sufficient level.

As the main scientific result of the study, we state that the decision to abolish CTG was not justified by the criteria that were formulated in the form of consolidated tax objectives (1–4), but was already based on these specified criteria. Consequently, the unsuccessful implementation of the CTG regime led to the fact that the provisions of the bill initially did not ensure the observance of the indicated interests of the regions, and then the decision-making bodies were unable to fully adapt the work of the CTG regime for the objectives.

At the same time, the study shows that the decision to abolish the CTG regime in 2023 cannot be considered fully justified. In our opinion, it is necessary to return to the revision of this issue when the Accounts Chamber analyzes the results of the functioning of the CTG regime for 2017–2020 and publishes them. This will make it possible to assess the redistribution of the tax base of those CTGs that in 2014–2016 had a zero tax base. Also, when assessing the change in the shares of subjects in income tax revenues from CTGs, it is necessary to adjust the coefficient of return of overpayments on advance payments from the budgets of the subjects of the Russian Federation.

The lack of a sufficient number of scientific papers that analyze in detail the consequences of the introduction of CTGs is due to the lack of available information about the results of their activities. Therefore, in connection with the decision taken, one can only assume that the abolition of consolidated taxation will lead to a new redistribution of tax revenues between regions, as well as to additional costs for taxpayers. Abolition of the CTG regime and narrowing the scope of transfer pricing control rules actually returns the distribution of income tax between regions, as it was before 2012, i.e. when two regions accounted for more than 35% of the total profit of all Russian organizations.

ACKNOWLEDGEMENTS

This article was prepared as part of the research work of the state assignment, RANEPA, Moscow, Russia.

The authors express their gratitude to Michael Alexeev for his professional and friendly support in the preparation of the article.

REFERENCES

1. Migukina N.E. Foreign experience of state regulation of institute of the consolidated taxation within providing economic security. *Voprosy ekonomiki i prava = Economic and Law Issues*. 2015;(90):81–87. (In Russ.).
2. Cooper G.S. Policy forum: A few observations on managing the taxation of corporate groups — The Australian experience. *Canadian Tax Journal / Revue Fiscale Canadienne*. 2011;59(2):265–294.
3. Mintz J. Europe slowly lurches to a common consolidated corporate tax base: Issues at stake. In: Schön W., Schreiber U., Spengel C., eds. A common consolidated corporate tax base for Europe — Eine einheitliche Körperschaftsteuerbemessungsgrundlage für Europa. Berlin, Heidelberg: Springer-Verlag; 2008:128–138. DOI: 10.1007/978–3–540–79484–4_7
4. Khan Niazi S.U. Re-launch of the proposal for a Common Consolidated Corporate Tax Base (CCCTB) in the EU: A shift in paradigm. *Legal Issues of Economic Integration*. 2017;44(3):293–314.
5. Cavalier G. Harmonization of the corporate tax in the EU — The CCCTB Proposals (First part). *Revue Internationale des Services Financiers/International Journal for Financial Services*. 2017;(1):114–117. URL: https://hal-univ-lyon3.archives-ouvertes.fr/hal-01561193/document
6. Hentze T. The challenge of moving to a Common Consolidated Corporate Tax Base in the EU. Institut der deutschen Wirtschaft. IW-Report. 2019;(2). URL: https://www.econstor.eu/bitstream/10419/191535/1/1047134152.pdf
7. Kireeva A.V. Prospects for the introduction of taxpayers' consolidation institution in the Russian Federation. *Ekonomiko-politicheskaya situatsiya v Rossii*. 2008;(2):56–58. (In Russ.).
8. Vitvitskaya T.V. The advantages and disadvantages of introducing the institution of consolidated groups of taxpayers. *Sotsial'no-ekonomicheskii i gumanitarnyi zhurnal Krasnoyarskogo GAU = Socio-Economic and Humanitarian Magazine Krasgau*. 2015;(2):19–27. (In Russ.).
9. Bannova K.A., Dolgikh I.N., Kuz'mina N.A. Improving the methodology for distribution of regional tax revenues from consolidated taxpayers group based on added value. *Finansy i kredit = Finance and Credit*. 2016;(3):29–39. (In Russ.).
10. Kondratieva E.A., Shalneva M.S. Problems of development of the institution of the consolidated group of taxpayers in Russia: The risk of abuse for the purpose of tax evasion. *Rassledovanie prestuplenii: problemy i puti ikh resheniya — Investigation of Crimes: Problems and Solution*. 2016;(1):44–52. (In Russ.).
11. Kondrashova N.A. Prerequisites for creating consolidated groups of taxpayers. *Mezhdunarodnyi bukhgalterskii uchet = International Accounting*. 2013;(2):28–34. (In Russ.).
12. Gorbatenko E.O. Consequences of establishing consolidated group of taxpayers to regional budget revenues in example of Irkutsk area. *EKO: vserossiiskii ekonomicheskii zhurnal = ECO Journal*. 2015;(4):177–183. (In Russ.).
13. Evnevich M.A. Consolidated group of taxpayers as a holding-like institution. *Vestnik Sankt-Peterburgskogo universiteta. Ekonomika = St Petersburg University Journal of Economic Studies (SUJES)*. 2013;(3):118–128. (In Russ.).
14. Babenko D.A. About development of foreign practice of application of the consolidated taxation concept. *Transportnoe delo Rossii — Transport Business of Russia*. 2014;(1):235–237. (In Russ.).
15. Babenko D.A. Introduction of the concept of the consolidated group of taxpayers in Russia: Preliminary results. *Finansovyi zhurnal = Financial Journal*. 2014;(2):135–142. (In Russ.).
16. Grinkevich L.S. Consolidated group of taxpayers as a tool for equitable distribution of tax revenues. *Sibirskaya finansovaya shkola = Siberian Financial School*. 2015;(2):40–44. (In Russ.).

17. Nikitin K.M., Avdienkova M.A., Zakharova O.O., Gilmanova E.A. Consolidated groups of taxpayers: How we should use the moratorium. *Finansovyi zhurnal = Financial Journal*. 2016;(5):37–51. (In Russ.).
18. Ilyin V.A., Povarova A.I. Consolidated taxation and its consequences for regional budgets. *Ekonomika regiona = Economy of Region*. 2019;15(1):70–83. (In Russ.). DOI: 10.17059/2019–1–6
19. Kizimov A.S., Shegurova T.A. Consolidated taxpayer institution: History of development and principles of functioning. *Finansy i kredit = Finance and Credit*. 2009;(30):11–16. (In Russ.).
20. Kostrykina N.S., Korytin A.V. Why did the consolidated tax regime cause massive losses in tax revenue in Russia? *Journal of Tax Reform*. 2020;6(1):6–21. DOI: 10.15826/jtr.2020.6.1.072
21. Malis N.I., Grundel L.P. The analysis of the consolidated taxpayer institution in Russia. *Finansy i kredit = Finance and Credit*. 2015;(3):38–46. (In Russ.).
22. Konovalova T.V. The results of functioning of the consolidated taxpayer groups in 2012–2015 on the example of Republic of Mari El. *Innovatsionnoe razvitie ekonomiki = Innovative Development of Economy*. 2016;(2):113–115. (In Russ.).
23. Leontieva J., Zaugarova E. Improvement of Russian consolidated groups of taxpayers taxation system. *Zhurnal pravovykh i ekonomicheskikh issledovanii = Journal of Legal and Economic Studies*. 2014;(2):109–115. (In Russ.).
24. Shatalov S.D. On the main directions of tax policy at the present stage. *Finansy = Finance*. 2015;(5):3–7. (In Russ.).
25. Shu-Chien J.C. Neutrality as tax justice: The case of Common Consolidated Corporate Tax Base under the EU law. *European Studies: The Review of European Law, Economics and Politics*. 2018;5:33–66.
26. Mintz J. Corporate tax harmonization in Europe: It's all about compliance. *International Tax and Public Finance*. 2004;11(2):221–234. DOI: 10.1023/B: ITAX.0000011401.67566.14

ABOUT THE AUTHORS

Andrey V. Korytin — Research Fellow, Laboratory for Tax Policy Studies, Russian Presidential Academy of National Economy and Public Administration (RANEPA), Moscow, Russia
https://orcid.org/0000-0001-6635-7190
Corresponding author
Korytin-av@ranepa.ru

Natalya S. Kostrykina — Research Fellow, Laboratory for Tax Policy Studies, Russian Presidential Academy of National Economy and Public Administration (RANEPA), Moscow, Russia
https://orcid.org/0000-0002-9810-0573
kostrykina-ns@ranepa.ru

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 19.11.2021; revised on 06.12.2021 and accepted for publication on 17.12.2021.
The authors read and approved the final version of the manuscript.