Vol. 24, No. 1, February 2023 ISSN 1234-7655

STATISTICS IN TRANSITION

new series

STATISTICS OF UKRAINE

An International Journal of the Polish Statistical Association and Statistics Poland

Scientific and information journal of the National Academy of Statistics, Accounting and Audit

A NEW ROLE FOR STATISTICS: Joint Special Issue

EDITOR

Włodzimierz Okrasa — *University of Cardinal Stefan Wyszyński, Warsaw and Statistics Poland, Warsaw, Poland*
e-mail: w.okrasa@stat.gov.pl; phone number +48 22 – 608 30 66

EDITORIAL BOARD

Dominik Rozkrut (Co-Chairman)	*Statistics Poland, Warsaw, Poland*
Waldemar Tarczyński (Co-Chairman)	*University of Szczecin, Szczecin, Poland*
Czesław Domański	*University of Łódź, Łódź, Poland*
Malay Ghosh	*University of Florida, Gainesville, USA*
Graham Kalton	*University of Maryland, College Park, USA*
Mirosław Krzyśko	*Adam Mickiewicz University in Poznań, Poznań, Poland*
Partha Lahiri	*University of Maryland, College Park, USA*
Danny Pfeffermann	*Professor Emeritus, Hebrew University of Jerusalem, Jerusalem, Israel*
Carl-Erik Särndal	*Statistics Sweden, Stockholm, Sweden*
Jacek Wesołowski	*Statistics Poland, and Warsaw University of Technology, Warsaw, Poland*
Janusz L. Wywiał	*University of Economics in Katowice, Katowice, Poland*

ASSOCIATE EDITORS

Arup Banerji	*The World Bank, Washington, USA*	Andrzej Młodak	*Statistical Office Poznań, Poznań, Poland*
Misha V. Belkindas	*ODW Consulting, USA*	Colm A. O'Muircheartaigh	*University of Chicago, Chicago, USA*
Sanjay Chaudhuri	*National University of Singapore, Singapore*	Ralf Münnich	*University of Trier, Trier, Germany*
Eugeniusz Gatnar	*National Bank of Poland, Warsaw, Poland*	Oleksandr H. Osaulenko	*National Academy of Statistics, Accounting and Audit, Kiev, Ukraine*
Krzysztof Jajuga	*Wrocław University of Economics, Wrocław, Poland*	Viera Pacáková	*University of Pardubice, Pardubice, Czech Republic*
Alina Jędrzejczak	*University of Łódź, Poland*	Tomasz Panek	*Warsaw School of Economics, Warsaw, Poland*
Marianna Kotzeva	*EC, Eurostat, Luxembourg*	Mirosław Pawlak	*University of Manitoba, Winnipeg, Canada*
Marcin Kozak	*University of Information Technology and Management in Rzeszów, Rzeszów, Poland*	Mirosław Szreder	*University of Gdańsk, Gdańsk, Poland*
Danute Krapavickaite	*Institute of Mathematics and Informatics, Vilnius, Lithuania*	Imbi Traat	*University of Tartu, Tartu, Estonia*
Martins Liberts	*Bank of Latvia, Riga, Latvia*	Vijay Verma	*Siena University, Siena, Italy*
Risto Lehtonen	*University of Helsinki, Helsinki, Finland*	Gabriella Vukovich	*Hungarian Central Statistical Office, Budapest, Hungary*
Achille Lemmi	*Siena University, Siena, Italy*	Zhanjun Xing	*Shandong University, Shandong, China*

EDITORIAL OFFICE

ISSN 1234-7655

Scientific Secretary
Marek Cierpiał-Wolan, *Statistical Office in Rzeszów, Rzeszów, Poland, e-mail: m.cierpial-wolan@stat.gov.pl*
Managing Editor
Adriana Nowakowska, *Statistics Poland, Warsaw, e-mail: a.nowakowska3@stat.gov.pl*
Secretary
Patryk Barszcz, *Statistics Poland, Warsaw, Poland, e-mail: p.barszcz@stat.gov.pl, phone number +48 22 – 608 33 66*
Technical Assistant
Rajmund Litkowiec, *Statistical Office in Rzeszów, Rzeszów, Poland, e-mail: r.litkowiec@stat.gov.pl*

© Copyright by Polish Statistical Association, Statistics Poland and the authors, some rights reserved. CC BY-SA 4.0 licence

Address for correspondence
Statistics Poland, al. Niepodległości 208, 00-925 Warsaw, Poland, tel./fax: +48 22 – 825 03 95

Preface

This volume is the result of the need of the moment – similarly felt by both of us, the undersigned – to meet the demand of the international statistical community for first-hand knowledge of the multiple consequences of the war in Ukraine for the functioning of the national statistical system as well as for statistics as a discipline, and as “statistics without borders”. On behalf of the editorial offices and scientific boards and committees of the *Statistics in Transition new series* and *Statistics of Ukraine*, about half a year ago we invited researchers and practitioners to submit manuscripts to a joint Special Issue devoted to statistical data production in wartime conditions.

From the descriptions of situations in which the national statistical system functions, including evidence on organizational and methodological problems and challenges, presented in this volume, a vision for a new role of statistics and statisticians emerges as important participants in ongoing processes.

We present this volume to the Readers in the hope that it will shed light on these issues and draw attention to those that require immediate attention and reflection by members of the international statistical community.

Włodzimierz Okrasa
Editor of Statistics in Transition new series

Oleksandr H. Osaulenko
Guest Editor of Statistics in Transition new series & Editor of Statistics of Ukraine

CONTENTS

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. V

Submission information for Authors

Statistics in Transition new series (SiTns) is an international journal published jointly by the Polish Statistical Association (PTS) and Statistics Poland, on a quarterly basis (during 1993–2006 it was issued twice and since 2006 three times a year). Also, it has extended its scope of interest beyond its originally primary focus on statistical issues pertinent to transition from centrally planned to a market-oriented economy through embracing questions related to systemic transformations of and within the national statistical systems, world-wide.

The SiT*ns* seeks contributors that address the full range of problems involved in data production, data dissemination and utilization, providing international community of statisticians and users – including researchers, teachers, policy makers and the general public – with a platform for exchange of ideas and for sharing best practices in all areas of the development of statistics.

Accordingly, articles dealing with any topics of statistics and its advancement – as either a scientific domain (new research and data analysis methods) or as a domain of informational infrastructure of the economy, society and the state – are appropriate for *Statistics in Transition new series.*

Demonstration of the role played by statistical research and data in economic growth and social progress (both locally and globally), including better-informed decisions and greater participation of citizens, are of particular interest.

Each paper submitted by prospective authors are peer reviewed by internationally recognized experts, who are guided in their decisions about the publication by criteria of originality and overall quality, including its content and form, and of potential interest to readers (esp. professionals).

Manuscript should be submitted electronically to the Editor:
sit@stat.gov.pl,
GUS/Statistics Poland,
Al. Niepodległości 208, R. 296, 00-925 Warsaw, Poland

It is assumed, that the submitted manuscript has not been published previously and that it is not under review elsewhere. It should include an abstract (of not more than 1600 characters, including spaces). Inquiries concerning the submitted manuscript, its current status etc., should be directed to the Editor by email, address above, or w.okrasa@stat.gov.pl.

For other aspects of editorial policies and procedures see the SiT Guidelines on its Web site: https://sit.stat.gov.pl/ForAuthors.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. VII–VIII

Policy Statement

The broad objective of *Statistics in Transition new series* is to advance the statistical and associated methods used primarily by statistical agencies and other research institutions. To meet that objective, the journal encompasses a wide range of topics in statistical design and analysis, including survey methodology and survey sampling, census methodology, statistical uses of administrative data sources, estimation methods, economic and demographic studies, and novel methods of analysis of socio-economic and population data. With its focus on innovative methods that address practical problems, the journal favours papers that report new methods accompanied by real-life applications. Authoritative review papers on important problems faced by statisticians in agencies and academia also fall within the journal's scope.

***

Abstracting and Indexing Databases

Statistics in Transition new series is currently covered in:

BASE – Bielefeld Academic Search Engine

CEEOL – Central and Eastern European Online Library

CEJSH (The Central European Journal of Social Sciences and Humanities)

CNKI Scholar (China National Knowledge Infrastructure)

CNPIEC – cnpLINKer

CORE

Current Index to Statistics

Dimensions

DOAJ (Directory of Open Access Journals)

EconPapers

EconStore

Electronic Journals Library

Elsevier – Scopus

ERIH PLUS (European Reference Index for the Humanities and Social Sciences)

Genamics JournalSeek

Google Scholar

Index Copernicus

J-Gate

JournalGuide

JournalTOCs

Keepers Registry

MIAR

Microsoft Academic

OpenAIRE

ProQuest – Summon

Publons

QOAM (Quality Open Access Market)

ReadCube

RePec

SCImago Journal & Country Rank

TDNet

Technische Informationsbibliothek (TIB) – German National Library of Science and Technology

Ulrichsweb & Ulrich's Periodicals Directory

WanFang Data

WorldCat (OCLC)

Zenodo

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. IX–XIV

From the Editors

The tragic events currently taking place in Ukraine have affected all aspects of life and activity, in private and public spheres, including an unspeakably difficult situation of the state statistics services.

The presented to the readers issue, entitled *A New Role for Statistics,* is the product of the jointly undertaken task by *Statistics in Transition new series* and *Statistics of Ukraine* to showing some of the enormity of problems experienced by the statisticians of Ukraine and the ways they are dealing with them. As first-hand accounts, articles by Ukrainian statisticians also provide information about the disruptions and types of assistance expected as well. A large part of possible reactions from the international community of statisticians has already been preliminarily identified and addressed in the opening of this volume basing on summary of the presentations that panelists representing various types of institutions and organizations gave at the session devoted to these issues within the last FCSM2022 conference (see *The Post-Conflict Reconstruction of the Statistical System in Ukraine...*).

This issue contains15 papers, which focus on functioning of statistical system in war conditions demonstrating the role of statistics in documenting the effects ofRussian aggression on the economy and society of the invaded country. Particular emphasis is put on the humanitarian crisis and the degradation of people's well-being, and on challenges faced by statisticians along with new tasks and approaches to overcome tchem.

This Joint Special Issue gives us also the opportunity to express our appreciation and thanks for all our contributors: authors, reviewers and all the participants of the editorial process.

This Joint Special Issue starts with the paper ***Problems realating to the statistical research of the national market of logistics services in war conditions*** by **Nataliia Hrynchak, Olha Yatsenko, Olena Bulatova,** and **Olena Ptashchenko**. The article discusses the theoretical principles of statistical research with regard to the national market of logistics services during wartime, and identifies the main structural changes that occurred due to the hostilities faced by the country. The authors determine the main factors influencing the functioning of the analysed market during war, as the statistical study of these factors is considered necessary for the transformation and development of logistics services. According to the results of the evaluation, analysis and structuring of relevant indicators and factors affecting the development of the logistics services market, their priority is determined according to the type of logistics services, which makes it possible to identify new opportunities for development both at the micro- and at the macro level.

The article entitled ***Using Big Data by Ukrainian official statistics when martial law applies: problems and solutions*** by **Oleksandr H. Osaulenko** and **Olena Horobets** focuses on issues of the secure operation of official statistics in Ukraine during the application of martial law. The level of digitalisation in Ukraine as the basis for using Big Data was analysed by the proposed indices of internetisation, social progress and digital transformation, and several problems (methodological, legal, financial, and managerial) were identified as vital for statistical offices on their way to the implementation of Big Data in statistical processes. proposals concern tools for Big Data processing, The authors discuss the proposals such as Data Hypercube as a way for presenting Big Data for their visualisation, applications of Web scraping in estimating the consumer prices index, analyses of labour and real estate markets, and the applications of specialised software for the collection, processing and analysis of Big Data sets.

Nataliia Reznikova, Iryna Zvarych, Roman Zvarych, and **Ivashchenko Oksana** in their paper ***The impact of the Russian-Ukrainian war on the green transition and the energy crisis: Ukrainian scenario of circular economy development*** analyse how to minimise the impact of the energy crisis on the environment as one of the ways of getting rid of carbon footprints resulting from the growth of the russian energy and building a circular sustainable ecosystem in Ukraine. The paper determines the impact that the war has on the practice of applying resource nationalism associated with a wide variety of modern global problems. It also identifies the dominant diversification tendencies in the EU in terms of the circularity of the economy. The proposed concept of a global inclusive circular economy can be considered as a complex multidimensional system, whose main components are based on the economic, sociological, environmental and circular aspects of life.

The next article ***A statistical study of climate change in Ukraine under martial law*** by **Tetiana Kobylynska, Iryna Legan,** and **Olena Motuzka** presents the development of theoretical and methodological foundations of statistical research in the field of national environmental and economic accounting, which forms the basis for the development of indicators of climate change under martial law and shapes the adaptation to these changes. The paper studies issues of producing ecological information relating to Ukraine according to statistical data, and describes the main problems which arise during the construction of national environmental accounts were characterised. The article identified the key factors which influence to the largest extent the quality of statistical data and calculations, and which are necessary for the transformation and development of the statistical estimation of climate change under Russian military aggression.

Olha Lubenchenko, Svitlana Shulga, and **Halyna Pavlova** discuss ***Method of auditing in conditions of martial law***. The authors consider methodical recommendations on the actions of auditors during martial law that relate to such stages of the audit as the preparatory phase, the planning phase, the task implementation and the final phase. Under martial law, new risks are emerging, systematized by the authors and related to the identification of persons involved in terrorist activities and the proliferation of weapons of mass destruction. The paper has been developed also to assess ethical threats in the light of martial law. The war in Ukraine has forced auditors to tackle new challenges in complying with the latest legal requirements for identifying those involved in military aggression against Ukraine, on the one hand, and requiring careful compliance with International Standards on Auditing.

In the next manuscript entitled ***Current challenges related to the consumer price index (CPI) in Ukraine*** **Olga Vasyechko** analyses how to contribute to the maintenance and compilation of the consumer price index (CPI) in the current extreme situation caused by the Russian military aggression against Ukraine. The interaction between the ideal and conditional concepts of the index and their practical implementation is considered as a potential source of compilation improvement. The author argues that the main factor of the modern criticism of the CPI is the systematic deviation of the practical form of the index from its theoretical foundations. The revision of the paradigm of primary data sources allows for a significant reduction in the methodological and organizational limitations imposed by the extreme conditions of Russia's military aggression against Ukraine. In the conditions caused by the war, this kind of information allows regular estimates of the consumer price index for a large number of goods without the loss of quality, and control the structure of consumption both in general and by region, and opens prospects for reducing discrepancies between conventional concept of the CPI, its ideal concepts and their practical application.

Volodymyr Sarioglo and **Maryna Ogay's** article presents ***Approach to population estimation in Ukraine using mobile operators' data*** discussing the task of developing effective approaches to estimating the population size using data from existing sources, in particular the data of mobile operators regarding the number, location and mobility of subscribers. The article highlights the results of a study on the use of data from mobile operators, data from administrative registers, and the results of a special population sample survey on the use of mobile communication for the purpose of estimating the population. It also provides the results of experimental calculations of the population size in Ukraine as a whole and in particular regions. The developed approaches can be used to assess and monitor the number and location of the population of Ukraine, provided the availability and proper preparation of data of mobile operators, the availability of administrative records containing information about the population, the availability of sample surveys, in particular on the peculiar use of mobile communications by the population.

Taisiia Bondaruk, Liudmyla Momotiuk, and **Iryna Zaichko** focus on ***Budgetary policy of Ukraine in time of challenges and its impact on financial security.*** The aim of the study is to deepen the theoretical and methodological foundations of the creation and implementation of budgetary policy in Ukraine, evaluation of its impact on the financial security in time of challenges. The study uses methods of comparative analysis, grouping in the process of evaluating the current state of budgetary policy indicators, methods of normalization and standardization of data, modelling, and graphical analysis of data for normalizing the financial security indicators and determining the dynamics of financial security components. The materials and reports containing statistical data from the Ministry of Finance of Ukraine and the State Statistics Service of Ukraine served as the basis of the study. It was determined that the components of the state's financial security in the face of martial law and pandemic do not take into account the impact of budgetary policy. Therefore, in the course of comprehensive integrated assessment of the financial security of the state, additional indicators were proposed.

The paper by **Tetyana Chala, Oleksiy Korepanov, Juliia Lazebnyk, Daryna Chernenko,** and **Georgii Korepanov** deals with ***Statistical modelling and forecasting of wheat and meslin export from Ukraine using the singular spectra analysis.*** The article presents the problems related to the functioning of the worldwide market of wheat and meslin. The structure of wheat export by Ukrainian regions is analysed in comparison with the total export. The localisation coefficient is applied to measure the regional unevenness of the distribution of wheat export volumes and the total export by regions of the country. The modelling and forecasting of the volumes and prices of export of wheat and meslin from Ukraine are based on Singular Spectrum Analysis. The study particularly focuses on the individual components of time series, such as trend, annual, semi-annual, four-month, three-month seasonal components. The reliability of the forecast is confirmed by the calculation of the MAPE forecast error and Henry Theil’s inequality coefficient. The article proposes an algorithm for calculating the relative indicators of the structure for the individual components of the reconstructed time series, identified through the singular spectral analysis.

The next article prepared by **Halyna Holubova** ***A comparative analysis of the principal component method and parallel analysis in working with official statistical data*** *describes* the basic conceptual approaches to the definition of principle components. Moreover, the methodological principles of selecting the main components are presented. A comparative analysis of the eigenvalues was performed by means of two methods: the Kaiser criterion and the parallel Horn analysis on the example of several data sets. The study shows that the method of parallel analysis produces more valid results with actual data sets. The author believes that the main advantage of Parallel analysis is its ability to model the process of selecting the required number of main components by determining the point at which they cannot be distinguished from those generated by simulated noise. The Parallel analysis method uses multiple data simulations to overcome the problem of random errors. This method assumes that the components of real data must have greater eigenvalues than the parallel components derived from simulated data which have the same sample size and design, variance and number of variables.

Oleg Krekhivskyi and **Olena Salikhova** in their manuscript consider ***A new industrial strategy for Europe – new indicators of the results of its implementation***, The paper discusses the experiences resulting from EU's adoption and implementation of a wide variety of policy measures in response to the COVID-19 crisis. These measures included stimulating the relocation and expansion of manufacturing to reduce vulnerability, depending on imports, ensuring the stability and development of industrial production. The study proposes and tests a new approach to assessing the consequences of relocation policies aimed at developing the local production potential, increasing the value added by activity, and expanding the share of local value added in industry exports. The manuscript focuses on the formation of statistical analysis tools for assessing the changes of the specialisation and identifying the country's comparative advantages. The authors propose new indicators: RSP – coefficient of Revealed Specialisation of Production, CAVA – coefficient of Comparative Advantage in Value Added by Activity and EVA – coefficient of Comparative Advantages in the Domestic Value Added Exports.

The paper entitled ***Assessing the maturity of the current global system for combating financial and cyber fraud*** by **Olha Kuzmenko, Hanna Yarovenko,** and **Larysa Perkhun** assesses the maturity of systems for counteracting financial and cyber fraud with the view of their future integration at global-level. The calculations made by the authors were based on indicators for 76 countries, which characterized each country's level of cybersecurity and its ability to combat financial fraud in 2018. The authors conducted a bifurcation analysis of the maturity of current global system for combating financial and cyber fraud and produced its phase portraits. It was found to be mature („Government Efficiency Index – Ease of Doing Business" and „Ease of Doing Business – Crime Index") and insufficient mature („Government Efficiency Index – Crime Index"), with the components' imbalance indicating high system's sensitivity to react on changes. The constructed 'Equilibrium States' phase portraits showed non-equilibrium phase portraits of the 'saddle' type. The obtained results made it possible to identify determinants of a global integrated system's instability to combat financial and cyber fraud.

Ella Libanova and **Oleksii Pozniak** in their paper ***War-driven wave of Ukrainian emigration to Europe: an attempt to evaluate the scale and consequences (the view of Ukrainian researchers)*** evaluate the scale and consequences of the emigration of

Ukrainians triggered by the military aggression of the Russian Federation. The paper also attempts to determine the composition of the refugees. According to the estimation of the Ptukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine based on the data from the State Border Guard Service, the number of 'refugees from the war in Ukraine' reached 3 million as of the end of June 2022. The potential amount of irreversible migration losses, depending on the military and economic factors, ranges from 600–700 thousand to 5–5.5 million people. Considering the fact that approximately 3 million Ukrainians had already been staying (working) abroad before 2022, the war is likely to result in a demographic catastrophe for Ukraine, whose demographic potential has been utterly exhausted.

The article prepared by **Maryna Puhachova** and **Oleksandr Gladun** entitled ***Using electronic registries to study the COVID-19 pandemic and its consequences*** analyses systems of electronic information resources (registers and databases) in the field of the healthcare in different countries. These systems provide information to support the treatment of patients, and also also accumulate large amounts of statistics, thus enabling their qualitative operational analysis. The authors summarise information on the use of electronic registers and databases to create an information base for the study of the COVID-19 pandemic and its consequences in different countries, and formulate proposals for the improvement of electronic health systems in Ukraine. On the basis they propose a list of electronic registers that can significantly improve the analysis of both, the course and the consequences of the coronavirus disease.

Deepika Rajoriya and **Diwakar Shukla's** manuscript ***Under military war weapon support the economic bond level estimation using generalized Petersen graph with imputation*** presents a sample based estimation methodology for estimating the meaneconomic bond value among countries involved in the military support or business. The problem is derived from current Russia-Ukraine war situation. A node sampling procedure is proposed whosebias, mean-squared error and other properties are derived. Results are supported with empirical studies. Findings are compared with particular cases and confidence intervals are usedas a basic tool of comparison. Pattern imputation is used together with a new proposal of CI-Imputation method who has been proved useful for filling the missing value, specially whensecret economic support data from involved countries found missing.

Włodzimierz Okrasa
Editor Statistics in Transition new series

Oleksandr H. Osaulenko
Guest Editor of Statistics in Transition new series & Editor of Statistics of Ukraine

© Włodzimierz Okrasa, Oleksandr H. Osaulenko. Article available under the CC BY-SA 4.0 licence

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 1–12, DOI 10.59170/stattrans-2023-001

The Post-Conflict Reconstruction of the Statistical System in Ukraine. Key Issues from an International Perspective[1]

Dominik Rozkrut[2], Włodzimierz Okrasa[3], Oleksandr H. Osaulenko[4], Misha V. Belkindas[5], Ronald L. Wasserstein[6]

Introduction

Statistics has accompanied the social forms of human civilization since its inception, reflecting also conflicts and wars. Statistics acts as a beacon, especially in turbulent times, capturing the most important aspects of reality, while helping decision-makers navigate key choices in the face of adversity of a radically changing situation. To this end, statisticians of a war-affected country make every effort by adapting the way statistics work to overcome methodological and organizational obstacles in everyday professional work, including innovative development of research instruments to substitute the destroyed or unavailable ones. Historical records indicate that the first statistical tables began to appear in Sumer, Egypt, ancient China, Babylon, and Assyria. Statistics continues its role with increasing scope and importance through centuries, with especially hard time during the Second World War, when conducting statistical research was prohibited in the German-occupied countries. [However, the compilation of statistics in some countries subjected to the most hostile

---

[1] Based on the presentations given by the panelists at the session ***Marshall Plan for Reconstructing National Statistical Offices After Conflict: Practical Guidance from International Principles. The role of statistical societies:*** Misha Belkindas; Ronald Wasserstein; Włodzimierz Okrasa and Dominik Rozkrut. The session was organized by Jennifer Park, Committee on National Statistics (CNSTAT); it was chaired by Dominik Rozkrut and commented by Albert Kroese (International Monetary Fund). It took place during the Federal Committee on Statistical Methodology/FCSM-2022 Research and Policy Conference, October 25–27, Washington D.C.

[2] Statistics Poland, Poland. E-mail: D.Rozkrut@stat.gov.pl. ORCID: https://orcid.org/0000-0002-0949-8605.

[3] University of Cardinal Stefan Wyszynski in Warsaw and Statistics Poland, Poland.
E-mail: w.okrasa@stat.gov.pl. ORCID: https://orcid.org/0000-0001-6443-480X.

[4] National Academy of Statistics and Accounting and Audit, Kyiv, Ukraine.
E-mail: O.Osaulenko@nasoa.edu.ua. ORCID: https://orcid.org/0000-0002-7100-7176.

[5] International Association of Official Statistics (IAOS) and ODW Consulting, USA.
E-mail: MishaBelkindas@opendatawatch.com.

[6] American Statistical Association, USA. E-mail: ron.wasserstein@amstat.org.

D. Rozkrut, W. Okrasa, O. H. Osaulenko, M. V. Belkindas, R. L. Wasserstein. Article available under the CC BY-SA 4.0 licence

occupation was conducted, including Poland, where the census was carried out in 1941, by the Underground State]. Currently, we are witnessing how Ukraine gives examples of heroism also in the sphere of official statistics, striving to fulfill its mission of constantly informing state institutions and society despite the extraordinary wartime challenges.

Remarkably, the National Statistical Office of Ukraine has continued to operate since the beginning of the Russian-Ukrainian war. This resilience is a testimony to the essential nature of objective, accurate, reliable, and timely national statistics to inform policy-making, and the steadfastness of the national statisticians behind the numbers. The healthy functioning of a national statistical office has implications for its relationships with bi-lateral and multi-lateral agreements with donor countries and organizations, and therefore, for the security of its country. This year marks the anniversary of the Fundamental Principles of Official Statistics (FPOS, 1992), championed by many esteemed thought leaders, including Józef Oleński (former President of Statistics Poland), Jean-Louis Bodin (INSEE, France), and Katherine Wallman (former, U.S. Chief Statistician and chair, UNECE CES), within the context of the United Nations Economic Commission for Europe Conference of European Statisticians as a way to support the production of national statistics among countries transitioning from centrally planned economies to market economies. The Fundamental Principles of Official Statistics subsequently was endorsed by the highest body of the UN, the General Assembly (2014). There have been additional efforts to develop aspirational and practical guidance for national statistical offices to strengthen capacity. The European Statistics Code of Practice (2005), U.S. Office of Management and Budget's Statistical Policy Directive 1 (2014) (now embedded in the Evidence Act), OECD's Recommendation of the Council on Good Statistical Practice (2015), American Statistical Association's Ethical Guidance for Statistical Practice (2022) are but a few. There have also been efforts to develop implementation guidance for these principles; notably, FPOS (2011, 2015, and 2020).

After the Second World War, the Marshall Plan was implemented to assist in the reconstruction and strengthening of nation states affected by conflict. Similarly, panelists of the FCSM-2022 session hosted by Jennifer Park, Committee on National Statistics (CNSTAT), discussed a set of the following issues:

- What would a Marshall Plan for national statistics in Ukraine look like?
- How can the parameters of the FPOS and other aspirational guidance inform practical steps of such a plan?
- What roles could various entities take to implement such a plan?
- What elements are essential in the short term? Over the longer term?

In the panel discussion summarized below, representatives of the wide spectrum of international statistical community addressed key aspects of the above questions, taking into account the current situation of statistical institutions and the circumstances in which Ukrainian statisticians try to fulfill their tasks in the conditions of war. The vast majority of the problems and challenges – along with practical ways to deal with them – are presented in the articles by Ukrainian statisticians that make up this issue. Additional information was provided by Oleksandr Osaulenko and his colleagues for the purpose of panel discussion.

Summary of the panel discussion

As an introduction to the session, its co-organizer (with J. Park of CNSTAT) and chairman, Dominik Rozkrut, characterized briefly the situation of Ukraine's state and society including information on the influx of immigrants from the war zone to Poland. He also quoted some results of the household budget surveys (recently conducted by Statistics Poland) concerning the scope and types of assistance provided to the immigrants. The extent of involvement of Polish households in various forms of help to refugees – such as hosting, food, clothes, other in-kind and in-cash assistance – seems impressive: in total, 78 percent of households, i.e. members of every three out of four dwelling units (about 11.5 out of 15.3 million) participated in one or other types of such assistance.

The international official statistics community has been filled with discussions on timeliness vs granularity of statistics for years. However, events such as a global pandemic or war become a practical test of methodological advancement and organisational solutions, a real test of the organisation's agility and readiness to meet the sudden information needs of societies. The full-scale war in Ukraine became a unique challenge for Polish statistics. The sudden influx of war refugees, a powerful economic shock through drastic increases in energy prices and disruption of value chains posed a severe challenge to Statistics Poland. This challenge is connected both with the need to provide the latest information about the rapidly changing socio-economic processes, but also with the need to anticipate future needs after the end of the war.

Statistics Poland, as was the case with the pandemic, showed a quick response. As early as April 2022, i.e. just after the outbreak of war, additional questions were introduced in the surveys currently carried out or planned, addressed to households, enterprises, social economy entities and non-profit organisations, local government units, and entities operating accommodation facilities. Even earlier (from mid-March 2022), work was started on a new pilot study addressing refugees at reception points, aimed at characterizing people fleeing the war from Ukraine to Poland. At the same time, Statistics Poland was preparing a plan to use data on Ukrainian citizens residing in Poland from public administration registers and information systems. To provide a legal basis for accessing data from information systems and official registers of public administration, Statistics Poland actively joined the work on the Act and amendments to the Act of March 12, 2022, on assistance to Ukrainian citizens in connection with an armed conflict in the territory of that country as well as preparing the regulation of the Council of Ministers amending the Statistical Work Program for 2022. As a result of these activities, Statistics Poland obtained access to newly established or amended administrative registers, including data on citizens of Ukraine who have come to the territory of the Republic of Poland, social security numbers, education, social assistance, social security, and healthcare insurance. The activities' scope was extensive and covered both the rapid development of research methodology and the enormity of organisational activities related to the implementation of large-scale mass research. As a result, a wide range of statistical research results were quickly obtained.

The statistics of aid and support for Ukraine can serve as the best example here. The invasion on February 24 this year caused, out of concern for their life and health, millions of people in Ukraine decided to leave their country and seek shelter outside its borders, mainly in Poland. Significantly extended social surveys showed that 70.2% of households in Poland granted support to the inhabitants of Ukraine from February 24, 2022, to the end of the first half of 2022, and social economy entities declared 8.0 million recipients of support. From February 24 to March 31, 2022, 28.8 thousand social economy entities (29.6%), including 28.6 thousand non-profit organisations (29.8%) and 0.2 thousand cooperatives (16.9%) (social economy sector), took additional measures to benefit those in need in connection with the war on the territory of Ukraine providing those in need with material support with an estimated value of PLN 511 million and financial support of PLN 140 million. Of 28.8 thousand social economy entities involved, 98.1% conducted activities in Poland and 7.8% in Ukraine. Natural persons were the primary recipients of aid provided by social economy entities in connection with the war in Ukraine. They were supported by 67.1% of non-profit organisations and 99.1% of cooperatives declaring their commitment to helping. One of the most basic forms of aid in connection with hostilities in Ukraine was a donation in kind (64.2% of non-profit organisations and 37.9% of cooperatives). The estimated value of in-kind support provided by non-profit organisations amounted to PLN 509.3 million, and in the case of cooperatives, PLN 1.8 million (PLN 511.1 million in total).

These are only very brief examples of timely and granular statistics provided based on extending the information scope of the surveys. In total, 17 surveys were expanded to include new questions used as direct data sources in surveys, of which 16 were modified immediately following the start of the invasion (in spring 2022). Similarly, the same happened in the case of economic surveys, where a focus was put on providing evidence to assess the economic impacts of the war, primarily through monitoring of business tendencies, employment, Inflation, and financial results.

A fine example of an ad-hoc study is a refugee health study created from scratch, designed and implemented jointly with the World Health Organisation. Statistics Poland conducted a pilot study in April and May 2022 among refugees from Ukraine who stayed at reception points in the Podkarpackie region. It was then followed up by a regular survey conducted from June to August 2022. The study covered the way and place of crossing the border; characteristics of people crossing the border due to citizenship, sex, age, and education; planned place/country of stay; intention to work in Poland, taking advantage of education, intent to return to Ukraine after the end of hostilities, health care needs in Poland, access to health care, information on vaccination against COVID-19 and vaccinations for childhood diseases, mental health, the health needs of refugees, and information on their health status in the context of WHO's planning of future assistance for this group of people

To sum up, the scope of the conducted research and the results obtained were unprecedented internationally. The actions taken and the results obtained were the subject of many international discussions, often indicated as examples of good practices for other countries in the future. In 2023, the UN Statistical Commission introduced information on developing the refugee health survey methodology, which will be further developed jointly by Statistics Poland and WHO to benefit the international community.

Along with mass relocations, also within the territory of Ukraine, institutions and statistical research centers in several areas are emptying, making it impossible to provide data on a regular basis. However, the Ukrainian system of official statistics continues to function and perform its main functions on a scale that can be achieved in wartime conditions only due to the involvement of its devoted staff in headquarters and regions.

Main features of official statistics in Ukraine[7]

The regulatory framework of the state statistical system's functioning is based on the Law of Ukraine "On Official Statistics", issued by the Verkhovna Rada (the Parliament) to be entered into force on January 1, 2023. The law harmonizes the national statistical system with European principles and standards to make it able to produce high quality statistical information about the economic, social, demographic and environmental situation in Ukraine and its regions. The law is based on the provisions of Regulation (EC) No. 223/2009 of the European Parliament and the Council dated 11.03.2009, which in turn is the basic document within the framework

[7] This section is based on the presentation by W. Okrasa and O. H. Osaulenko "Statistics in troubled times – the case of Ukraine".

of the implementation of the EU Statistical Compendium and the provisions of the Generic Law on Official Statistics. The Law contains the main provisions of the European Statistics *Code of Practice*.

The State Statistics Service/SSS of Ukraine is a central executive body in the field of statistics – its activities are guided and coordinated by the Cabinet of Ministers of Ukraine. The UA SSS also ensures the development and implementation of state policy in the field of statistics, its offices and staff: 27 regional offices; 6455 employees. According to Ukrainian authorities, national statistical system is reformed and modernized in accordance with the EU/Eurostat principles: 27 (35.1%) of the state statistical observation centers fully meet the requirements of the EU Compendium; 50 (64.9%) of the state statistical observation centers partially meet the requirements of the EU Compendium. The UA State Statistics Service (SSS) strives to fully implement the EU Statistical Requirements Compendium.

The challenges of war – the voice of Ukrainian statisticians

To illustrate the difficulties faced by Ukrainian statisticians, let us quote excerpts from some of the articles contained in this collection:

- *The inability to conduct national statistical surveys makes it difficult to estimate the size of the population due to being limited to existing sources: data from mobile operators, data from administrative registers, and from a special population sample survey*, (Volodymyr Sarioglo, Maryna Ogay).
- *Despite the current extreme situation (…), the CPI must be compiled on an ongoing basis – this is done using Big Data, especially direct cash data, expanding the sample size and improving its design while reducing the burden on respondents and obtaining more reliable transaction price data by incorporating real-time information on household expenditure,* (Tetiana Kobylynska, Iryna Legan, Olena Motuzka).
- *In order to assure operation of the official statistics in Ukraine (under the Martial Law) the involvement of alternative data sources, including Big Data, is necessary. These data should be introduced in parallel or in mix with conventional data sources, to fill the gaps in conventional data due to the war. [Ukraine has an extensive network of private digital services: e.g. Monobank, express delivery "Nova poshta"; mobile phones, social networks, Google analytics, etc. have to be considered too,* (Olha Kuzmenko, Hanna Yarovenko, Larysa Perkhun).
- *The war in Ukraine affects all forms of international economic relations, highlighting the problem of asymmetric economic interdependence in the green transition to climate neutrality, accompanied by raw materials, energy and food crises. The question arises how to minimize the impact of the crisis on the environment as*

part of getting rid of the carbon footprint of the past (Russian) energy model towards building a sustainable circular ecosystem in Ukraine, (Olga Vasyechko).

- *The war in Ukraine forced auditors to tackle new challenges due to new risks emerging that need to be recognized, systematized, and treated accordingly – including identification of persons involved in terrorist activities and the proliferation of weapons of mass destruction - while complying with the legal requirements concerning both factors associated with military aggression against Ukraine and those involving compliance with International Standards on Auditing*, (Tetyana Chala, Oleksiy Korepanov, Iuliia Lazebnyk, Daryna Chernenko, Georgii Korepanov).
- *The assessment of the scale and effects of forced external migration of Ukrainians as a result of Russian aggression – based on the data of the State Border Guard – shows that "military emigrants" are, in general, people with higher education than the national average, mainly women who easily adapt to life abroad, especially in Poland (due to the minimal linguistic and cultural differences)*, (Oleg Krekhivskyi, Olena Salikhova).

Among the hardships caused by the war, statisticians feel the following the most:

- lack of effective sampling frames and data sources;
    - production of official statistics continues using administrative data – this allows the assessment of key macroindicators like Ukraine's GDP, and to publish statistical information on foreign trade in goods, etc.,
    - regional authorities continue to register prices at the points of sale of goods, which allowed to continue producing the Consumer Price Index in Ukraine as a whole, and by regions,
- respondents are legally deprived of the obligation to provide data during Martial Law;
    - however, respondents continue to provide primary data as part of a voluntary activity – reporting rate is over 65%,
- regional offices located in temporarily occupied territories or near the military zones may perform their functions only partially, or not at all;
    - in order to ensure the continuity of the production of official statistics, a back-up system for the collection and processing of data has been established, according to which, for a regional office that is temporarily unable to perform certain statistical tasks, such tasks are delegated to be performed by another office (located in a safer place),
- surveys of household living conditions and demographic data production have been suspended;

- employees of state statistical offices in several regions were forced to migrate (to other regions or abroad);
- frequent air alarms force employees to spend a lot of time in shelters.

Research works, trainings, infrastructure base and technology

Given the extremely difficult circumstances, the tasks performed by Ukrainian statistical institutions during the past 12 months can be considered impressive, embracing 5,292,492 respondent reports processed and the numbers of products based on the statistical research (observations): 19,072 statistical information/reports – 320 Open Data sets – 3,792 press releases – 3,177 data collections – 125,223 users of the "Respondent Account" service – 139 visits to the "Search by USREOU code" service – 180,934 completed international questionnaires based on the results of the state statistical observations – 2,800,936 visits on the SSS official website.

However, drastic cuts in funding resulted in a significant reduction in the scope of the program implemented this year. To be more specific, in 2021, UAH 805,000 (approx. U$ 22,000) was allocated to two research projects: (i) methodology of conducting sample surveys of the population: "Statistics of income and living conditions in the European Union EU-SILC" (USD 10,000) and (ii) methodology for conducting an integrated survey of short-term enterprise statistics (USD 12,000). In 2022, due to budget constraints, the expenditure on the implementation of these two scientific research works was cancelled. The draft budget for 2023 does not provide for expenses for the implementation of these two projects.

Training and retraining programs cover approximately 1,000 employees per year. In 2021, 269 people were retrained; in 2022 only 15. In 2022, the cost of studying statistics students is UAH 8,150,000 (U$220,270). Nevertheless, in 2022 there were 110 students, slightly less than in 2021 (120 students).

The infrastructure base and technologies are in a deplorable state. There are practically no sources for the renovation of technical equipment – the last time the fleet of servers and computers was renewed in 2014. The challenges of martial law require an increase in the share of field work as remote work. At the same time, there are practically no laptops, etc. at the headquarters of the UA State Statistical Service and its regional branches. The most urgently needed assistance should include modern technologies for collecting data and creating analytical databases, including "alternative" new sources of information (Big Data, analysis of satellite images, smart statistics etc.)

Issues raised by panelists and the views expressed

The presentations referred below concerned on the problem of the type and scope of aid for Ukraine from the two complementary points of view – the national organization that engages in international projects, which is American Statistical Association (ASA), by Ron Wasserstein, Executive Director of ASA, and the international organization (IAOS), by Misha Belkindas, President of IAOS – towards establishing needed support and coordinate cooperation between national and international offices.

Focusing his presentation on *The role of statistical societies,* Ron Wasserstein[8] summarized what professional associations do and why it matters to the NSOs, using activities of ASA and some other societies as examples. He concluded some ideas on how associations can help.

Professional societies, such as ASA, conduct a wide spectrum of activities, each of them can provide a platform for arranging for a respective support to UASSS, and eventually other NSOs. These range from facilitating scientific gathering – conferences, networking opportunities – and collaboration in the form of workshops, colloquia interest groups, and knowledge dissemination through meetings and journals to statistical capacity building, including technical training, leadership and communication skills, accreditation/certification, and support education of future.

The indication of this type of activity is also consistent with the results of a survey conducted by ASA among its members, asking them about things they consider fundamental in the activities of a professional society: over 90% of respondents selected meetings and publications. In addition to meetings, professional societies hold smaller gatherings that bring people together to discuss research and methodological interests.

Also, many societies have local or regional groups (ASA calls them "chapters" and has 75 of them) that facilitate gatherings of statisticians. And many have groups organized around statistical topics of interest. Chapters and interest groups (called "Sections" at ASA) often function like smaller versions of the organization, offering their own meetings and networking opportunities, having a newsletter, and so on.

Professional societies serve the extraordinarily important function of disseminating knowledge through meetings and journals. For NSOs, meetings might well be the place where research and methodology can be discussed in an audience of peers in and out of government.

Professional societies serve the function of providing skills to members that they need beyond their formal education. Members provide technical training from the beginning to advanced level through seminars, webinars, workshops, etc. And there are many non-technical skills the statistician needs that NSS's can provide as well, and often there are few options for getting such training. As one example worth mentioning here

[8] ***Standard disclaimer***: ASA gives the author time and opportunity to speak at events like this. However, the views expressed here are those of the author and should not be construed as an official statement or position of ASA –R.W].

is the ASA project which provides training for individuals interested in serving as experts witnesses in the court system.

Another cluster of envisaged forms of possible assistance concerns setting standards and promoting ethical practice (develop and disseminate guidelines for ethical practice); advocating for the profession and for sound practice, and developing relationships with like societies elsewhere.

He also pointed out four key themes for which this is relevant to the NSOs, especially in the context of their creation or reconstruction: a knowledgeable base of support – an independent voice – a source of skilled workers – a source for international connections.

In conclusion, Ron Wasserstein suggested three channels through which statistical societies can help one another: (i) share structures and governance, (ii) share expertise, (iii) share resources. The NSO's staffs need to engage in research, develop as professionals, meet with other statisticians (in government and out), and so on. But there are other reasons why NSS's bring value to NSO's.

Taking the perspective of an international organization, IAOS, its President, Misha Belkindas, concentrated his presentation, *Building/Redeveloping National Statistical Offices after Conflict*, around the intertwined fundamental issues:

A. Creation of an international coalition – how to establish it and whom to approach?
B. What an international coalition should/can do in the case of Ukraine?
C. What activities are currently going on over there?

First of all, creation of an international coalition should start with identification of potential donors led most likely by the World Bank or a regional development bank, and involve others partners, such as IMF, UNSD, OECD, EU, UN regional commissions, UN specialized agencies (ILO, FAO, others). Also included should be interested countries providing funds and technical assistance (TA), private sector, providers of funds and TA, and international NGOs, such as ISI, IAOS and others. Such a type of endeavors is not unprecedented. The TA project to the countries of the former Soviet Union in the 1990s, with a broad coalition under joint leadership of IMF, WB, UNSD, OECD and Eurostat, can serve as an example. Even more relevant in this context seem to be the projects on strengthening statistical systems of Armenia during the war with Azerbaijan, or in the former Yugoslavia – the case of Bosnia and Herzegovina.

Other projects implemented under the banner of the World Bank encompass a multi-country lending facility STATCAP (started in 2004) with loans to Burkina Faso and Ukraine, followed by loans with grant elements to many countries in all the continents. And the World Bank large lending projects in the Kyrgyz Republic, Tajikistan and Uzbekistan, also regional lending programs in Africa, including such countries like war-torn Somalia. The Inter-American Development Bank has also organized similar types of projects.

In addition to the needs covered by activities of international organizations with appropriate funds, there are several areas requiring assistance for which NGOs may be suitable implementers. However, this would have to be preceded by the creation of appropriate executive structures enabling NGOs to perform their tasks at the central and regional levels.

Such an international coalition, diversified in its interests and abilities to provide need-adjusted assistance, can help with data collection, processing and production of statistical outputs, including problems related to (i) design a data collection mechanism – surveys with imperfect sampling frames, usage of administrative data, other data, other means of data capture, and (ii) procurement and installation of means/lines for data transmission, storage, manipulation and publication, and (iii) design and start of implementation of HR policies with an emphasis on training and retraining of NSO staff; (iv) initiate the development of young cadre, if needed, develop curriculum for local universities, or in neighboring countries; and (v) train data users – policy makers, journalists, civil society. It would be important to also draft, with assistance of international agencies, a new law on statistics which adheres to international standards, and create a political environment for adherence to Fundamental Principles of Official Statistics.

Taking into account the specificity of the disrupted Ukrainian SSS, and priorities for its reconstruction in accordance with the new Law on Statistics (which the Parliament passed recently), it should be mentioned that the Ukraine Government is familiar with a large-scale international aid: the first institutional building loan to the Government was approved by the World Bank in the early 1990s (with a USD 9 million funding for the NSO). In addition, there was a large-scale international Technical Assistance program TACIS, which Ukraine was one of the recipients. The second loan in the amount of USD 32 million was approved in early 2004 and addressed institutional building, HR, data collection, processing, development of specialized software, etc.

The new project will most likely include the following activities towards rebuilding/refurbishing the SSS and its regional offices: purchase of a large amount of IT equipment (servers, PCs, etc.) – attract new staff to the statistical service – launch a large staff training and retraining program – support local universities and the National Academy of Statistical Education in providing equipment, developing a curriculum and providing trainers.

As regards current activities, IAOS received a request form the Institute of Economic Forecasting of the Ukrainian Academy of Sciences, which was tasked by the Government to develop methods for the calculation of the damage caused by the war.

They want to start from the damage done to the agricultural sector, in particular small-scale farming. IAOS has so far approached the ASA Statisticians Without Borders and received a positive answer. FAO agreed to render assistance by including Ukraine in their AGRIS program. World Bank cannot finance Ukraine from their Trust Fund 50x2030 as the country is not eligible – however, IAOS will continue to help on this and will try to obtain sources from a Trust Fund, at least in-kind.

Conclusions

Ukrainians, fighting for the preservation of their state, need – apart from military means of defense – reliable information for the state, its institutions and people, for now and for the fundamental reconstruction of their country in the near future.

The above presentations and discussion provide an illustration of the type and amount of work that is expected on this line based on documentation of disruptions in data production and the general functioning of the state statistical service. The voice of Ukrainian experts is the leading for designing an effective strategy of internationally coordinated activities – including statistical capacity building at each level of the state statistical system' units.

The manuscripts submitted by Ukrainian statisticians for this Special Issue prepared jointly by *Statistics in Transition new series and Statystyka Ukraina* reflect the type and scale of the problems and challenges faced by Ukrainian statisticians in the conditions of war.

Reconstruction, modernization and strengthening of the national information infrastructure – with state statistical service as its institutional backbone – should become the goal of various initiatives and missions of the international community of statisticians. Starting with showing the areas of destruction and related needs – identified together with experts from Ukraine – is a project that requires more extensive work.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 13–28, DOI 10.59170/stattrans-2023-002
Received – 06.07.2022; accepted – 02.11.2022

Problems realating to the statistical research of the national market of logistics services in war conditions

Nataliia Hrynchak[1], Olha Yatsenko[2], Olena Bulatova[3], Olena Ptashchenko[4]

ABSTRACT

The article discusses the theoretical principles of statistical research with regard to the national market of logistics services during wartime. The system of statistical indicators of the research of the logistics services market is structured through the allocation of separate blocks of indicators by priority of their estimation in war conditions. The market of logistics services of Ukraine is characterised by selected blocks of indicators. The paper identifies the main structural changes that occurred due to the hostilities faced by the country. Moreover, the authors determine the main factors influencing the functioning of the analysed market during war, as the statistical study of these factors is considered necessary for the transformation and development of logistics services. It is substantiated that taking into account such statistical indicators as the level and availability of logistics infrastructure, security, the human factor and changes in legislation are likely to result in a new alternative direction of logistics routes and contribute to the development of the logistics services market in general. The authors proposed the construction of a centralised electronic service (or several services) with the purpose of consolidating information about the logistics infrastructure. The software system should offer the option of an operational interactive visualisation.

Key words: logistics service, logistics services market, system of statistical indicators, freight transportation, warehousing.

[1] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: gnatalia@ukr.net. ORCID: http://orcid.org/0000-0002-2046-6014.

[2] Kyiv National Economic University Named After Vadym Hetmana, Ukraine. E-mail: yacenkoolgakneu@gmail.com. ORCID: https://orcid.org/0000-0003-4399-2217.

[3] Mariupol State University, Ukraine. E-mail: olena_bulatova@yahoo.com. ORCID: https://orcid.org/0000-0001-7938-7874.

[4] Simon Kuznets Kharkiv National University of Economics, Ukraine. E-mail: helenavaleriya@gmail.com. ORCID: http://orcid.org/0000-0002-2413-7648.

© N. Hrynchak, O. Yatsenko, O. Bulatova, O. Ptashchenko. Article available under the CC BY-SA 4.0 licence

1. Introduction

The war with Russia has become a serious challenge for Ukraine, which requires general consolidation and optimization of all social processes. War is not only a battlefield, it is also an economic confrontation, where the effective functioning of the logistics services market plays a key role and to some extent is a competitive advantage over the aggressor.

David Beasley, the Executive Director of the UN World Food Programme, during the World Economic Forum in Davos, stressed that closed Ukrainian ports will have a significant impact on global food security (see: https://www.weforum.org/press/2022/05/global-food-crisis-must-be-solved-alongside-climate-crisis/). Due to the war waged by Russia against Ukraine, not only maritime but also other logistics routes and centres were affected, so optimization of logistics, reformatting of the logistics services market (LSM) are the next issues after military action that are crucial for Ukraine's survival, economic development of partner countries and eliminating the risks of the global food crisis.

To build an efficient and effective wartime economy, it is necessary to be able to comprehensively study the market of logistics services, including assessing logistics infrastructure, track the dynamics of cargo movement, fuel consumption, reputation of carriers, solve problems in the organization of loading and unloading, warehousing. An important issue is not only the survival of the population and businesses, but also the ability of the country's economy to develop in new conditions. This requires reliable and timely statistical information that will allow to analyze and structure the factors of influence that superdynamically shape the new state of the logistics services market, and optimize activities in new conditions.

The purpose of the study is to substantiate the theoretical and methodological foundations of statistical evaluation of the market of logistics services in Ukraine during the war and to identify the factors that will most influence its transformation.

2. Theoretical problems of statistical research of national markets for logistics services

Statistical study of national markets for logistics services has its own characteristics. Logistics infrastructure, the number and range of services of logistics providers in many countries differ significantly. The choice of type of transport carrying freight transportation depends on a number of factors (cost, speed, safety, etc.). National markets for logistics services are undergoing rapid transformation, in peacetime – in terms of supply and demand, during the war - also due to the lack of logistics infrastructure or limited access to it, as a result of increased security risk. Against this background, before outlining the boundaries of the national market of logistics services, its structure and potential, we define and justify the statistical classification of logistics services.

Many researchers have tried to classify logistics services using different classification criteria. Thus, S. Sremac, Ž. Stević, D. Pamučar, M. Arsić, B. Matić (2018) offer the following classification criteria for logistics services: type of service, degree of intangibility of the service, frequency of contact with the client, motive for purchasing the service, terms of service, type of service buyer. G. Rosa, M. Jedliński, U. Chrąchol-Barczyk (2017) grouped all logistics services into two categories: resource-oriented logistics services and intelligent logistics services related to supply chain planning and organization. A. Bhattacharjee (2018) distinguishes three groups of logistics services depending on the type of recipient / customer: logistics services for individual consumers, for end-users and for industrial enterprises.

Studies of classifications of logistics services confirm that there is no single universal set of logistics services. Moreover, the available classifications show the diversity of services, which is primarily related to the development of the logistics services market.

During the war, while a significant number of private providers of logistics services operating in the Ukrainian market, completely or partially suspended their activities in the occupied territories, the State Enterprise "Ukrposhta" continued to provide logistics services. Thus, in cooperation with WindRose, Ukrposhta resumed exports to the United States via Poland on March 1, 2022 to support small and medium-sized businesses. During the two months of the war (March – April 2022) the company delivered more than 926 tons of export shipments to countries around the world (see: https://www.ukrposhta.ua/ua/news).

In view of this, when assessing and analyzing the Ukrainian market of logistics services, subsection 53.1 “Activities of the national post office” should be included in the types of economic activity that determine MLS. This subsection covers the retrieval, sorting, transportation and delivery of domestic and international postal items in the form of parcels and packages by postal service, which may be carried out by one or more types of transport, both private and public, as well as the collection of letters and parcels from private mailboxes or post offices. Accordingly, statistical evaluation and analysis of the structure of the national market of logistics services is proposed, excluding passenger traffic.

The most pressing problem of assessing the market of logistics services is the complexity of its structure and the link of logistics activities with many other economic activities. Statistical research of the logistics services market should be carried out on the basis of a combination of quantitative estimates with analytical, attributive, descriptive characteristics. For studying the state of the logistics services market, the formation of a scientifically sound system of statistical indicators which characterize each element of the market structure is an extremely important and necessary task in the context of finding optimal functioning of regional transport systems in wartime and the need to recover the economy of the country in general.

3. Methodological principles of formation of a system of statistical indicators of evaluation of national market for logistics services

D. Leończuk (2016) included energy consumption, delivery time, transport speed, flexibility, reliability and load on vehicles to key indicators for assessing the efficiency of the logistics services market. E. Plambeck and B. Kalkanci (2020) noted that the efficiency of the logistics services market can be increased by harmonizing and coordinating different operations in supply chains. E. Bottani, A. Rizzi, G. Vignali (2014) outlined an integrated approach to achieving efficiency, which involves integrated management of packaging, procurement, warehousing and transport activities. Note that all these approaches have a number of disadvantages: the assessment of goods turnover is based on cost indicators, which contain the inflation component, which does not allow to correctly assess its dynamics; quantitative indicators of infrastructure development do not reflect its quality and efficiency of use; the size of warehouse space incorrectly reflects the volume of the logistics services market, as a significant part of warehouse space is used by representatives of wholesale and retail trade for their own purposes; most macroeconomic indicators do not reflect the development of new key business competencies, including in the logistics services market.

A system of statistical indicators for studies of the logistics services market needs to comply with the following set of core criteria:

- the statistical indicators must be correct, quantifiable, and must have a reliability required for practical purposes;
- the statistical indicators must be aggregated: it means that when aggregated they must allow for moving from one level to another by the use of various statistical methods;
- the statistical indicators must be politically neutral and applicable for an economic statistical analysis with identification of the causalities.

In order to comprehensively study the market of logistics services, especially given the negative effects of the war and the importance of effective functioning of this market for the recovery of the country, we propose to use a system of statistical indicators of logistics services market assessment (Figure 1). This system contains three blocks and allows you to diagnose the state and trends of the market development, taking into account the influence of other related markets and types of economic activities, as well as to predict and model the ways of development of the national market of logistics services.

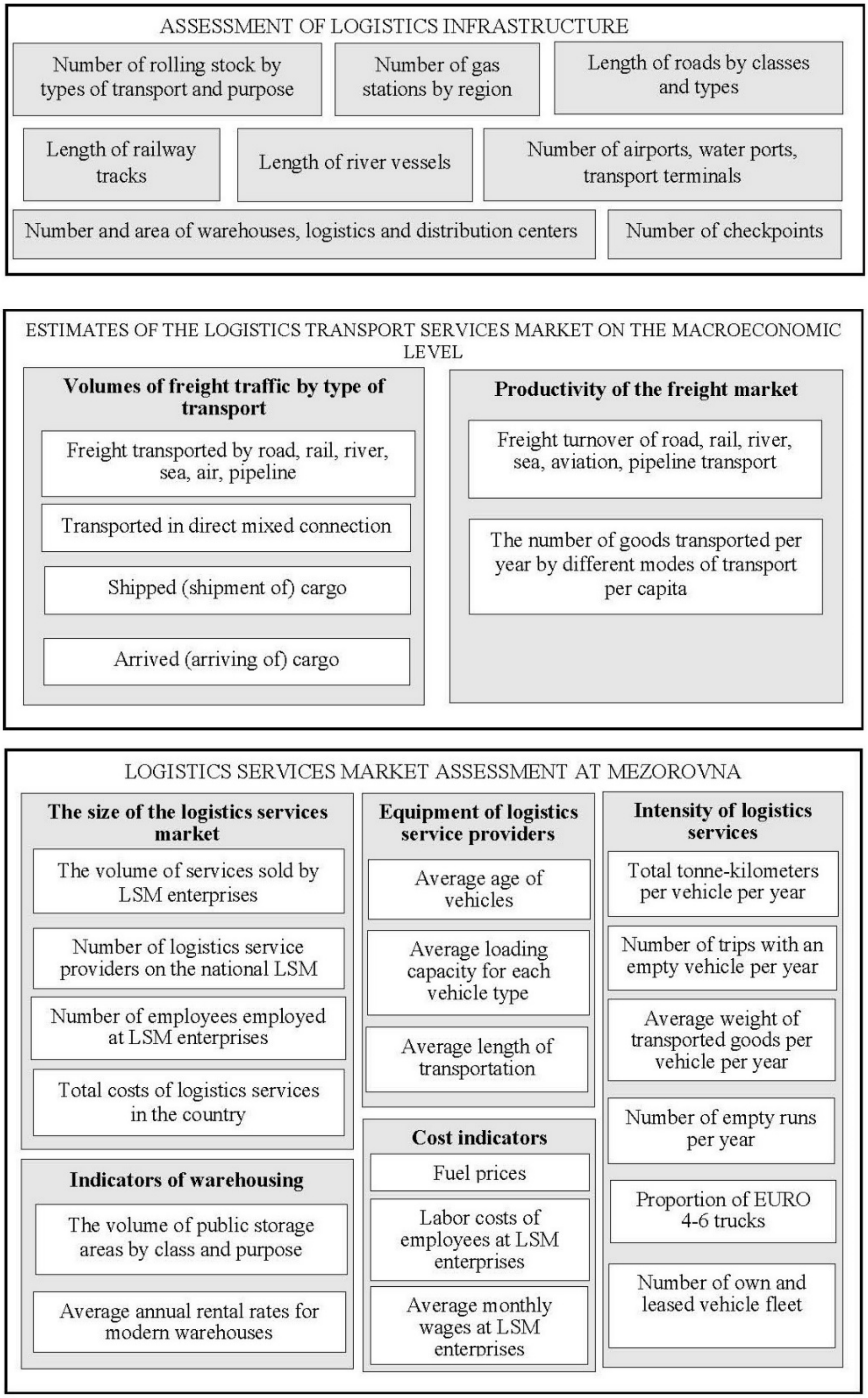


Figure 1. System of indicators for evaluating the national market of logistics services

If in peacetime the problem for logistics service providers was the economic assessment of the logistics market, then during the war the priority is to study the statistical characteristics of logistics infrastructure (density, length and accessibility of roads, number of actually functioning airports, water ports, transport terminals, warehouse size areas, the number of logistics centres, distribution centers, etc.), as the possibility of their safe use will allow to implement logistics services.

The second block is the study of transport logistics services, because transportation is a key type of these services. Socio-economic development of the country requires coordinated, highly efficient work of transport, clear cooperation with all sectors of the economy. This can be achieved through better use of vehicles, improved coordination of all types of transport, widespread use of advanced forms of transportation.

Statistical study of freight transport characterizes the transportation of goods and freight by type of public transport, including by the regions of the country, the structure of transport by type of cargo, freight and freight turnover on a commercial basis by certain types of transport. Economic indicators that characterize the market include data on freight in terms of tonnage and destination/country of origin for all types of transport, including sea, air and land.

4. Statistical study of the Ukrainian market of logistics services during the war

Problems of functioning of the national market of logistics services during the war stem from the need to ensure the security of logistics providers and the destruction of logistics infrastructure. Analysts of the analytical department of the Kyiv School of Economics KSE Institute (see: https://kse.ua/ua/about-the-school/news/zagalna-suma-pryamih-zbitkiv-infrastrukturi-vzhe-perevishhuye-105-5-mlrd/?__cf_chl_tk=_Rsabj8LNS9STF11odv8ew007nhqfaD.KoLjDWi2mpQ-1653740310-0-gaNycGzNCP0) within the project "Russia will pay", with the support of the Office of the President of Ukraine, the Ministry of Economy, the Ministry of Infrastructure and the Ministry of Communities and Territories Development, determined that a total amount of direct losses of the Ukrainian economy from the damage and destruction of residential and non-residential buildings and infrastructure already exceeds $ 105.5 billion or over UAH 3.1 trillion as of May 25, 2022. Losses of the Ukrainian economy from the damage to physical logistics infrastructure since the beginning of the war, as of May 25, 2022 amount to $ 43.5 billion (Table 1).

The provided statistical information on numerous damages and destructions is not exhaustive due to the inability of citizens and public authorities to promptly record the damage in the context of each city and town.

Table 1. Destruction of Ukraine's logistics infrastructure since the beginning of the war as of May 25, 2022

№	Infrastructure facilities	Amount	Total losses, million dollars USA
1	Roads, thousand km	23.8	29879
2	Civil airports, units	11	6817
3	Railway infrastructure and rolling stock, including:		3676
	– railway tracks, thousand tracks	6.3	
	– railway bridges, units	43	
4	Bridges and bridge crossings, units	295	1646
5	Ports and port infrastructure		471
6	Military airfields, units	12	468
7	Warehouse infrastructure, units	181	286
8	Oil depots, units	28	227

Many branches of logistics companies were literally destroyed, and most employees were forced to leave their homes. With the end of hostilities in the regions, the logistics infrastructure began to recover, but suffered severe damages. The war made it impossible for the key highways connecting the West, Center and East (Lviv-Kyiv and Kyiv-Kharkiv routes) to operate, bridges were blown up, and railway junctions were destroyed. Accordingly, logistics routes have become longer, with the use of detours over safe roads. This, in turn, increased the delivery time of parcels, cargo, etc.

Aviation and water transport logistics amount to zero. All logistics services related to freight transport are carried out by road and rail, but due to the occupation and high risks of road transportation, the delivery of goods to certain parts of the country is delayed.

Studies of the structure of cargo transportation by certain modes of transport in peacetime have shown that despite the impact on the development of the country's logistics services market factors of the political environment and the presence of temporarily occupied territories, economic activity indicators of which are not included in the country's statistics, there were no significant changes in the structure of cargo transportation within the last 10 years. In the implementation of freight, road and rail have been and will remain the main types of transport (Figure 2). The share of road transport in the total volume of all freight traffic in 2020 compared to 2011 increased by 9 points and amounted to 75%, and the share of freight by rail and pipeline decreased by 19% and 6%, respectively (in 2011 it was 25% and 8%).

Analyzing the total volume of freight traffic by all modes of transport, it should be noted that in the presence of a declining trend in 2011-2015, this figure increased by 61 million tons in 2018 (so the annual growth was 1.03%), in 2019 decreased by 64 million tons, but in 2020 almost reached the level of 2018 (Figure 2).

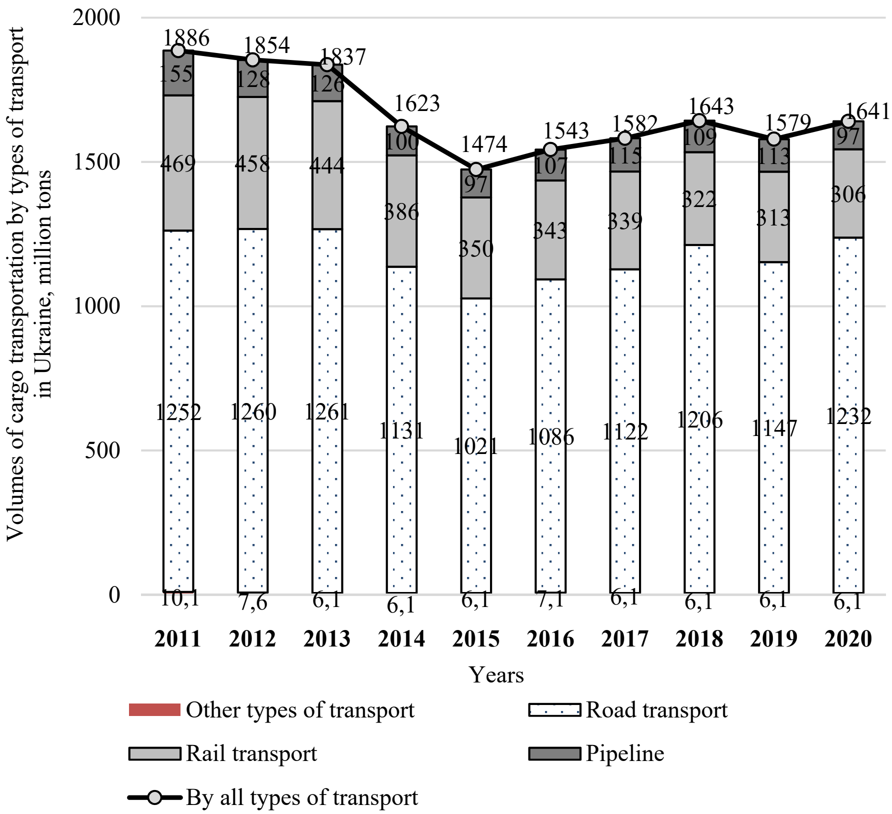


Figure 2. Volumes of cargo transportation by types of transport in Ukraine

Analysis of the productivity of the market of transport and logistics services in Ukraine based on the indicator of freight turnover by type of transport showed that in 2020 the total freight turnover by all types decreased by 30% compared to 2011 and this is primarily due to lower freight turnover by rail and pipeline (Figure 3).

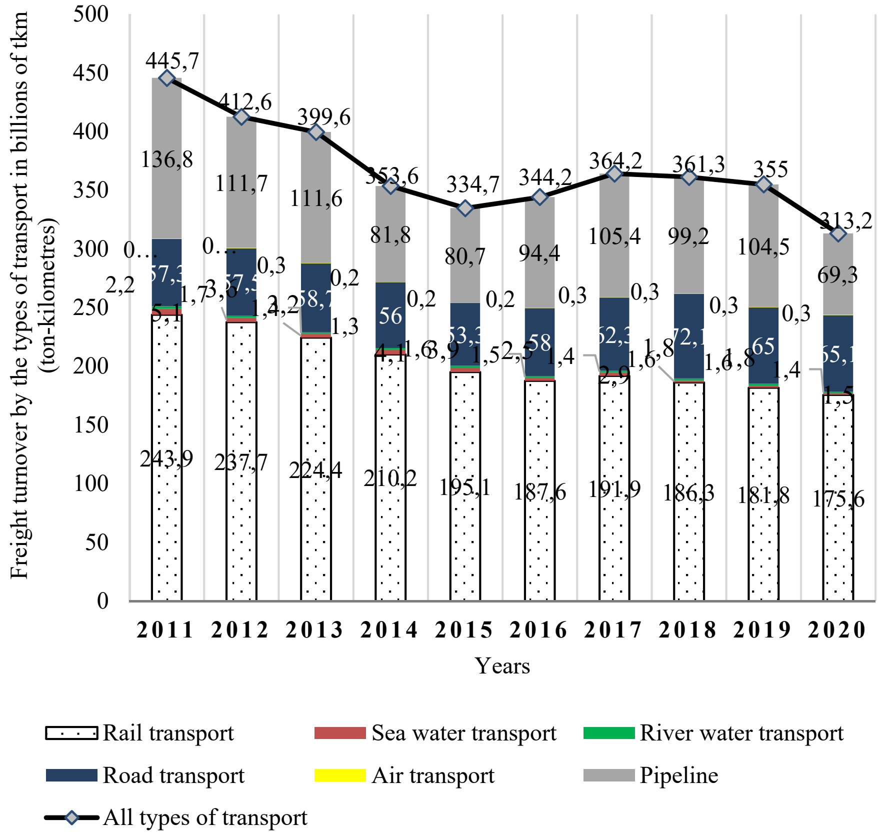


Figure 3. Dynamics of cargo turnover by types of transport in 2011–2020

In terms of cargo volumes by type of transport, even in terms of individual regions, statistical research of the transport logistics market carried out, as a rule, by all analysts in peacetime, cannot show the problems of the logistics market related to the impossibility of sea and air freight. In this case, it is more appropriate to conduct a statistical study of the volume of freight traffic by certain types of transport in terms of individual groups of goods, especially exported and imported products.

Let us take the example of transportation of grain freights. According to the information and analytical agency "APK-Inform" (see https://www.apk-inform.com/uk) and the State Statistics Service of Ukraine (see http://www.ukrstat.gov.ua/), grain exports from our country in the pre-war period was carried out mainly by sea. After the annexation of Crimea in 2014 and before the war, there were 13 seaports in Ukraine. The main volumes of grain exports in Ukraine before Russia's attack on Ukraine passed through the seaports of Mykolaiv, Odessa and Chornomorsk. In total, this is 95% of grain cargo exported by sea. Another 5% fell on Mariupol and Berdyansk. Most grain was transported to ports by rail. As a result of blocking ports due to Russian military aggression, rail traffic in this direction is expected to decline. If in February more than 96% of all grain consignments were loaded in the direction of ports, in March this figure dropped to less than 4%. Accordingly, the share of traffic in the direction of railway border crossings on the western border of Ukraine in March increased to 55% against 2% in February (Figure 4) (https://www.apk-inform.com/uk/news/1526420).

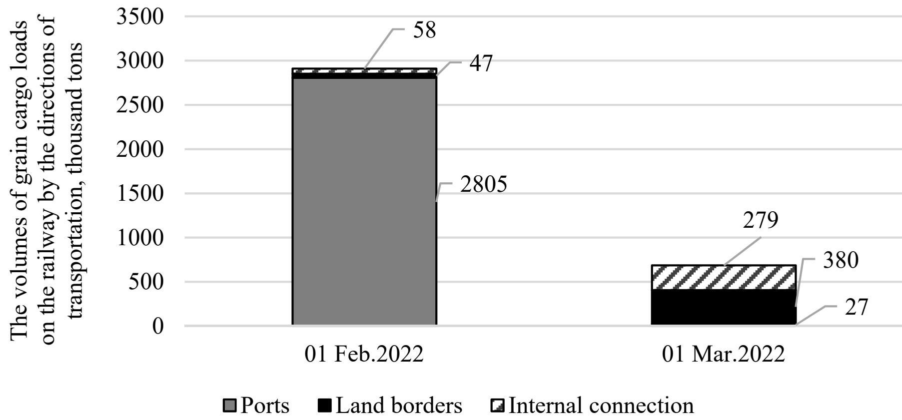


Figure 4. Structure of grain cargo loads on the railway by the directions of transportation

The global disadvantage of freight transport during the war is the lack of stable logistics chains that can provide the necessary volumes of goods, especially exports. That is, logistics operators, manufacturers, exporters need to establish new stable logistics chains, taking into account not only the danger and limitations of logistics infrastructure, but also the peculiarities of transportation of certain products. And here it is critical to pay attention to those areas of activity that were not a priority in peacetime. This requires reliable statistical information on the main problematic aspects of the implementation of logistics services that hinder the normalization of national and international freight transport, namely:

1) capacity of railway stations on the border with European countries in terms of resuming of exports / imports;
2) the number of European railway cars for the export of agricultural products from Ukraine (primarily the number of cars for grain transportation);
3) length and safety of roads by classes and types, as limited land freight transportation is not always caused by safety risks. Unfortunately, not all domestic roads allow for heavy-duty vehicles;
4) the number of vehicles for the carriage of respective goods (commodity groups);

5) the number of drivers for international road transport;
6) capacity of European logistics centres;
7) the number of railway cars for fuel import.

It is now critical to strengthen international partnerships in the field of logistics, in particular in solving the problems of finding alternative ways of delivering goods. For example, Romania has allowed Ukrainian grain to be exported through its own port, and Lithuania and Bulgaria are considering the same option. This is a difficult path, it will certainly lead to higher prices for Ukrainian products, but will preserve the export potential. According to the Minister of Agriculture and Rural Development of the Republic of Poland H. Kowalczyk, his country is currently working on the creation of a "dry port" on the border with Ukraine in order to increase transport capacity for export of Ukrainian agricultural products, including to third countries (see: https://minagro.gov.ua/news/ukrayina-ta-polshcha-nalagodzhuyut-bezperebijni-postavki-agroprodukciyi).

Another direction is road transport. With the beginning of the war, international companies suspended their work in Ukraine for various reasons. And this creates opportunities for domestic carriers. Europe has allowed our companies to carry out bilateral and transit traffic on no permits basis (see: https://delo.ua/uk/transport/bude-skladno-ale-zrestoyu-galuz-rozkvitne-logistika-pid-cas-viini-reformi-ta-maibutnje-industriyi-397214 /).

Based on the generalization of the volume of services provided by enterprises of section H by types of economic activity presented in Table 1, it is determined that the size of the logistics services market in Ukraine in peacetime, despite a slight decline in 2020 compared to the previous year, which mainly features the activity in the conditions of COVID-19, has grown 2.8 times for the last 10 years. The conducted study revealed that the rapid growth of services provided by enterprises of economic activity H "Transport, warehousing, postal and courier activities" is due to the growth in logistics services sold (Figure 5).

Given the situation that has developed in connection with the war in Ukraine, the respondents do not submit statistical reports today. According to the Law of Ukraine "On protection of the interests of the subjects of reporting and other documents during martial law or state of war" of 03.03.2022 № 2115-IX, natural persons, natural persons – entrepreneurs, legal entities shall submit accounting, financial, accounting, settlement, audit reports and any other documents within three months after the cessation or abolition of martial law or state of war for the entire period of non-reporting or obligation to submit documents (see: https://zakon.rada.gov.ua/laws/show/2115-20#Text). This makes it impossible to quickly assess the market of logistics services by the volume of logistics services sold.

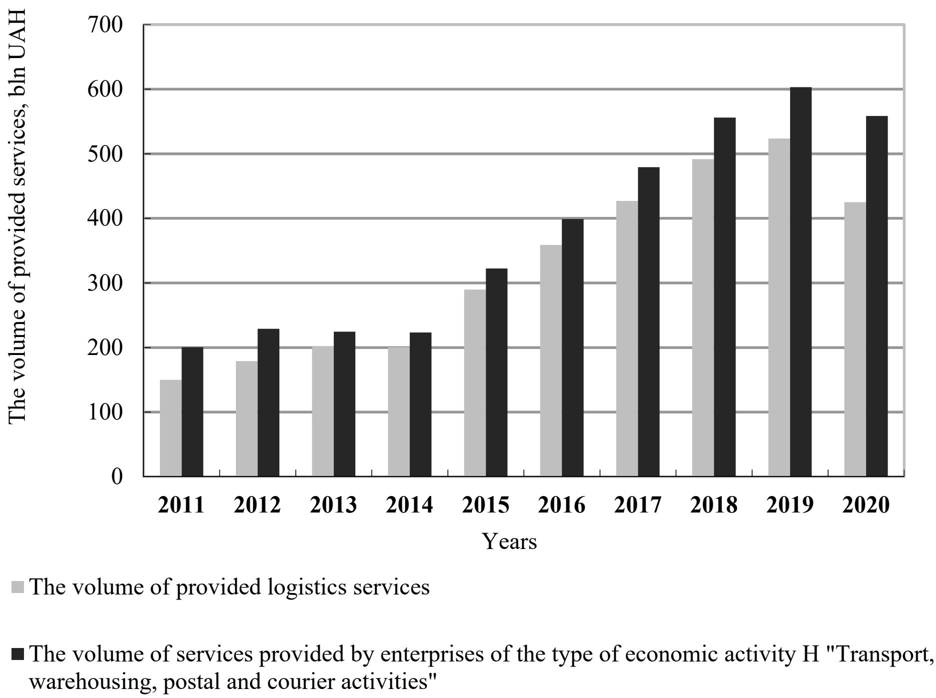


Figure 5. Dynamics of the volume of products sold by enterprises of the logistics services market for 2011–2020

It is worth focusing on the changes caused by the war in Ukraine which relate to the segment of warehousing logistics; the activity of the enterprises belonging to a kind of economic activity H 52.1 is meant. According to Alex Isachenko, the founder of digital warehousing provider WareTeka and CoreTeka, the elimination of the centre of gravity has been the first and most important change in Ukraine's warehousing logistics since the beginning of the war. The largest warehouse hub in Ukraine - 70-80% of all professional warehouse space - is located in Kyiv region, where 1.8-2.2 million square meters of professional warehouses were concentrated. According to WareTeka analysts, about 400,000 square meters of warehouses have been destroyed near Kyiv, this is about 20% of all professional warehouse real estate in Ukraine (see: https://mind.ua/openmind/20241674-zberegti-ta-zabezpechiti-yak-zminilasya-logistika-v-ukrayini-pid-chas-vijni).

With the beginning of the war, large companies, followed by medium and small ones, were forced to move their warehouse leftovers and goods to Western Ukraine, where such a large amount of warehouse space did not exist in general. The premises, which began to be used as warehouses, gave way to the former area, organization of space and level of service. All this has greatly changed the logistics processes, supply chains and the cost of the operations themselves.

The war has shown that it is unwise to place infrastructure in one place, as was the case in Ukraine. It is necessary to distribute warehouses throughout the country and create a buffer to keep them. We expect that the centre of gravity will change greatly when building new warehouses.

The situation in Ukraine is indicative of the whole world. The most important thing that can be learned about warehousing logistics is that the country must be able to quickly manage its logistics flows in emergencies. And in order to correct mistakes in the future, to invest wisely in the development of warehousing logistics at the national level, it is necessary to conduct a statistical study of the total area of warehouses not only by region but also by types (classes) of the warehouses.

Warehouse real estate companies develop their own warehouse classifications. Due to the lack of a generally accepted classification of warehouses in Ukraine, the most common is the international classification of warehouses from the brokerage agency Knight Frank. According to it, the warehouses are divided into 6 types: A +, A, B +, B, C, D (see https://content.knightfrank.com/resources/knightfrank.ru/pdf/research/ind.pdf). Knight Frank's classification is based on requirements for technical parameters, territory, equipment and communications.

It is also important to have state reserve warehouses, as it is organized in European countries: Germany, the Czech Republic, Poland. A large amount of humanitarian aid that Ukraine now receives from Europe comes from the state reserve.

In order to implement the principle of economy in freight transport, it is necessary to create logistics centres and centralized electronic services that would coordinate the actions of large carriers. Such consultation centres will help to fill any vehicle moving through the territory of Ukraine, in terms of both the commercial component and humanitarian aid. It is necessary to popularize the idea of competent logistics on the basis of reliable, up-to-date statistical data, which would unite the efforts of the state and business aimed at rebuilding the country. The war showed that these factors are also critical for work. Initiatives that can be useful for carriers, volunteers and refugees, such as the Crisis Logistics Center and VzayemoDiya, have now begun to emerge.

Another problem is that the main logistics stops working at night due to the curfew and the difficulties at checkpoints.

5. Conclusions

In the conditions of an ongoing war on the territory of Ukraine, the security risks of cargo transportation by road and rail have increased significantly. During the war, the management and coordination of logistics activities based on a statistical assessment of the availability and capabilities of the logistics infrastructure, both at the level of individual logistics service providers and at the state level in general, is of primary importance. In particular, the territorial level and potential of logistics development in the regions of Ukraine needs more statistical research.

In order to determine the criteria for the stability of the logistics services market, it is important to have statistical data which are grouped according to two aspects: logistics infrastructure and the human factor.

According to the results of the evaluation, analysis and structuring of relevant indicators and factors affecting the development of the logistics services market, their priority is determined according to the type of logistics services (warehousing, freight transportation, management and coordination of logistics activities), which makes it possible to identify new opportunities for development both at the micro- and at the macro level.

References

Bhattacharjee, C., (2018). The Effects of Broadband Service Quality on User's Value And Loyalty: A Study on Foreign Customers In Malaysia. *IOSR Journal of Business and Management,* 20(1), pp. 48–58.

Bottani, E., Rizzi, A., Vignali, G., (2014). Improving logistics efficiency of industrial districts: A framework and case study in the food sector. *International Journal of Logistics,* 18(5), pp. 1–22.

Chala, T. H., Hrynchak, N. A., (2020). Statystychne otsinyuvannya ta prohnozuvannya rozvytku rynku lohistychnykh posluh v Ukrayini. [Statistical evaluation and forecasting of the market of logistics services in Ukraine]. *Modern engineering and innovative technologies,* 3, pp. 99–106.

Fact-finding studies in support of the development of an EU strategy for freight transport logistics, Lot 1. Analysis of the EU logistics sector. (2015). Ecorys, Fraunhofer, TCI, Prognos and AUEB-RC/TRANSLOG.

Hrynchak, N., (2020). Studies of the Logistics Services Market: Information and Analytical Support and a System of Statistical Indicators. *Scientific Bulletin of National Academy of Statistics, Accounting and Audit*, (3), pp. 5–11. DOI 10.31767/nasoa.3-2020.01

Isachenko, O., (2022). Zberehty ta zabezpechyty: yak zminylasya lohistyka v Ukrayini pid chas viyny. [Save and ensure: how logistics changed in Ukraine during the war]. Retrieved from: https://mind.ua/openmind/20241674-zberegti-ta-zabezpechiti-yak-zminilasya-logistika-v-ukrayini-pid-chas-vijni [In Ukrainian]/.

Korol, O., (2022).Suvoryy eksport 2022: yak pid chas viyny prodaty ukrayins′ke zerno ta zapobihty prodovol′chiy katastrofi u sviti? (Strict exports 2022: how to sell Ukrainian grain during the war and prevent a food catastrophe in the world?). Retrieved from: https://latifundist.com/spetsproekt/969-suvorij-eksport-2022-yak-pid-chas-vijni-prodati-ukrayinske-zerno-ta-zapobigti-prodovolchij-katastrofi-u-sviti [In Ukrainian].

Leończuk, D., (2016). Categories of Supply Chain Performance Indicators: an Overview of Approaches. *Business Management and Education,* 14(1), pp. 103–115.

Plambeck, E. L., Kalkanci, B., (2020). Reveal the Supplier List? A Trade-off in Capacity vs. Responsibility. *Manufacturing & Service Operations Management*, 22(6), pp. 1107–1286/.

Rosa, G., Jedliński, M., Chrąchol-Barczyk, U., (2017). Marketing usług logistycznych. Warszawa, Poland: C. H. Beck.

Services sectoral classification list, Word trade organization. Retrieved from: https://www.wto.org/english/tratop_e/serv_e/mtn_gns_w_120_e.doc

Sremac, S., Stević, Ž., Pamučar, D., Arsić, M., Matić, B., (2018). Evaluation of a Third-Party Logistics (3PL) Provider Using a Rough SWARA–WASPAS Model Based on a New Rough Dombi Aggregator. *Symmetry,* 10, 305, DOI 10.3390/sym10080305.

Statistical Classification of Economic Activities in the European Community. Retrieved from: https://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=NACE_REV2&StrLanguageCode=EN&IntPcKey=&StrLayoutCode=HIERARCHIC.

Ukrayina ta Pol′shcha nalahodzhuyut′ bezperebiyni postavky ahroproduktsiyi (Ukraine and Poland establish uninterrupted supplies of agricultural products). Retrieved from: https://minagro.gov.ua/news/ukrayina-ta-polshcha-nalagodzhuyut-bezperebijni-postavki-agroprodukciyi [In Ukrainian].

Ukrposhta za 2 misyatsi viyny dostavyla ponad 926 tonn eksportu (Ukrposhta delivered more than 926 tons of exports during the two months of the war). Retrieved from: https://www.ukrposhta.ua/ua/news/57644-ukrposhta-za-2-misjaci-vijni-dostavila-ponad-926-tonn-eksportu [In Ukrainian].

V berezni 2022 roku obsyahy perevezen′ zernovykh vantazhiv zaliznytseyu zmenshylysya v 4,2 razy (In March 2022, the volume of grain transportation by rail decreased by 4.2 times). Retrieved from: https://www.apk-inform.com/uk/news/1526420 [In Ukrainian]

Zahal′na suma pryamykh zbytkiv infrastruktury vzhe perevyshchuye $105.5 mlrd (The total amount of direct infrastructure damage already exceeds $ 105.5 billion). Retrieved from: https://kse.ua/ua/about-the-school/news/zagalna-suma-pryamih-zbitkiv-infrastrukturi-vzhe-perevishhuye-105-5-mlrd/?__cf_chl_tk=_Rsabj8LNS9STF11odv8ew007nhqfaD.KoLjDWi2mpQ-1653740310-0-gaNycGzNCP0 [In Ukrainian].

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
Joint Special Issue: A New Role for Statistics
Vol. 24, No. 1 pp. 29–43, DOI 10.59170/stattrans-2023-003
Received – 01.07.2022; accepted – 12.10.2022

Using Big Data by Ukrainian official statistics when martial law applies: problems and solutions

Oleksandr H. Osaulenko[1], **Olena Horobets**[2]

ABSTRACT

The article is focused on issues of the secure operation of official statistics in Ukraine during the application of martial law. It was found that the gaps in conventional sources of statistical data caused by the war needed to be filled with data from alternative sources, including Big Data. The level of digitalisation in Ukraine as the basis for using Big Data was analysed by the proposed indices of internetisation, social progress and digital transformation. Thanks to our research, several problems (methodological, legal, financial, and managerial) were identified as vital for statistical offices on their way to the implementation of Big Data in statistical processes. Our proposals concern tools for Big Data processing, such as Data Hypercube as a way for presenting Big Data for their visualisation, applications of Web scraping in estimating the consumer prices index, analyses of labour and real estate markets, and the applications of specialised software for the collection, processing and analysis of Big Data sets

Key words: official statistics, statistics during war, Big Data, digitalisation.

1. Introduction

In connection with the large-scale invasion of the territory of sovereign Ukraine by the Russian Federation, the President of Ukraine signed the Decree “On the Imposition of Martial Law” and a series of other legal acts, including the Law of Ukraine “On the Protection of Interests of the Entities Submitting Reports or Other Documents in the Period of Martial Law or the State of War”, which specifies that the submission of any category of reports in paper or electronic form, including statistical and financial reports, shall be suspended for the period of martial law and three months following its abolition.

[1] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: O.Osaulenko@nasoa.edu.ua. ORCID: https://orcid.org/0000-0002-7100-7176.

[2] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: babutska@ukr.net. ORCID: https://orcid.org/0000-0003-1762-2140.

© Oleksandr H. Osaulenko, Olena Horobets. Article available under the CC BY-SA 4.0 licence

Apart of from irreversible losses for Ukraine and each of its residents, the war has hit its public administration system severely, including official statistical bodies, which had to suspend the collection of primary statistical data and publishing of statistical information.

In fact, the official statistics of Ukraine now is devoid of all the main sources for statistical information: primary data obtained through statistical observations and surveys (coming from business entities, physical persons, households, as well as government organizations, public entities, etc.), secondary data from administrative sources, and information that used to be collected by specialized government agencies.

In view of the above, it can by suggested that the operation of the Ukrainian system for primary statistical data collection has actually been stopped. Given that the availability of reliable statistical data constitutes an integral component of the efficient work and progress in all the walks of life, public administration included, with the official statistical information being part of the global information space, the official statistics of Ukraine urgently needs the involvement of alternative data sources and applications of innovative methods for their processing, in order to fill the gaps in the conventional data environment and preserve the national statistical system.

2. Methodology

When writing this article, the authors used methods of scientific theory and statistics, such as theoretical generalization, systematization and index analysis (to assess the digital transformation process).

3. Big Data: a promising source for statistical data in time of war

Although Big Data have gained increasing popularity in the recent years, this notion had long been in use before it became conventional. This term was used in the context of visualizing Big Data sets as early as 1997 (Cox and Ellsworth, 1997). It was applied in the context of data analysis in 1998 (Weiss and Indurkhya, 1998), and in the context of statistics in 2003 (Diebold, 2003).

The Oxford English Dictionary (2022) defines Big Data as "sets of information that are too large or too complex to handle, analyse or use with standard methods"; the Gartner Dictionary (2022) – as "high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation". It follows that Big Data refer to big sets of data on diverse social and economic phenomena and processes, produced on a continuous basis.

Big Data are characterized by specific components, which are six in number as distinguished by Microsoft Company. These components are given by C. Wu, R. Buyya, K. Ramamohanarao (2016):

1. Volume stands for scale of data.
2. Velocity denotes the analysis of streaming data.
3. Variety indicates different forms of data.
4. Veracity focuses on trustworthiness of data sources.
5. Variability refers to the complexity of data sets. In comparison with "Variety" (or different data format), it means the number of variables in data sets.
6. Visibility emphasizes that you need to have a full picture of data in order to make informative decision.

As argued by Bentein A. (2021), "the amount of digital data globally is doubling approximately every three years. The 74 zettabytes we have today stand for 79 billion terabytes, which is more than 10 terabytes for each living human on earth". Wiltshire D. & Alvanides S. (2022) emphasize that "Big data holds great potential for research and for society, large volumes of varied data can be produced and made available to researchers much faster compared to 'traditional' data… Big data is generated in higher frequencies than other forms of data, such as from social surveys and national censuses that can take months even years to be made available to researchers". It follows from the above that Big Data can provide a potential source for more relevant and timely statistical information compared with its conventional sources.

The global statistical community has realized the great significance of Big Data. The crucial step towards it was made by the Conference of European Statistical System on Big Data Problems (ESS Big Data event, 2014) in Rome, where the issue of statistical uses of Big Data was defined as a core one for the official statistics.

Later on, the European Commission clearly articulated its vision of Big Data sets as an important resource for economic growth and social progress on the whole. In particular, the European Strategy for data (2020) states that "data is at the centre of this transformation and more is to come. In a society where individuals will generate ever-increasing amounts of data, the way in which the data are collected and used must place the interests of the individual first, in accordance with European values, fundamental rights and rules. At the same time, the increasing volume of non-personal industrial data and public data in Europe, combined with technological change in how the data is stored and processed, will constitute a potential source of growth and innovation that should be tapped".

On February 23, 2022, the European Commission approved the Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act) (2022), with the aim of ensuring fairness in the allocation of value from data among actors in the data economy and to foster access to and use of data.

Big Data types are classified according to a definition that is mostly based on data sources, (United Nations, 2013), as follows: ”Data and information sources arising from the administration of a programme, be it governmental or not, e.g., electronic medical records, hospital visits, insurance records, bank records and food banks; Commercial or transactional sources arising from the transaction between two entities, e.g., credit card transactions and online transactions (including from mobile devices); Sensor network sources, e.g., camera data, satellite imaging, road sensors and climate sensors, such as those pertaining to Remote Sensing data sources; Tracking device sources, e.g., tracking data from mobile telephones and the Global Positioning System (GPS); Behavioural data sources, e.g., online searches (about a product, a service, or any other type of information) and online page views; Opinion data sources, e.g., comments on social media; Geographic information system (GIS) data and information of various sources and types“.

Following the adoption of the Sustainable Development Goals (SDGs), National Statistics Offices worldwide are requested to undertake a data revolution, which implies a titanic challenge to their mandate, namely to populate the SDGs matrices in all their economic, social and environmental dimensions (United Nations, 2021).

In spite of the abovementioned, a standard algorithm for collection, grouping, refining and analysis of Big Data still does not exist, a fact entailing distortions of produced results and formulating erroneous conclusions when processing these data sets. Apart from due consideration to the scopes, velocity, diversity, authenticity and variability of Big Data flow, its complex ecosystem has also to be born in mind. It follows that another characteristic of Big Data should be added, the hierarchical complexity caused by the full reliance on ICTs, whose advancement stimulates the occurrence and diversification of unique data in the continuous process. It involves the creation of a single set of primary (raw) data that can be roughly divided into three chains, with each one containing several interrelated groups, thus creating the data ecosystem.

The first chain contains the groups of data from general administrative data to data reflecting individual or public opinions or sentiments. The second chain, being more granulated and personalized, contains, inter alia, confidential data that have synergetic effects for each other. The third chain in the ecosystem, operating as a data filter, contains the already accurate and harmonized data formed in the four global groups: Shallow data, Deep data, Micro-data, and Nano-Data. It should be added that each type of the data has the so-called Dark data (Gartner, 2022a) of its own, which remain after mining. Dark data are not meant for repeated use because of high cost for their handling, which puts in question the reasonability of these data storage.

In this context, Henriques A. (2022) emphasizes that analysts are becoming more and more concerned with the occurrence of structured, non-structured and semi-structured data of a new type in the ecosystem, which are the data of experience, accumulating the data on education, life activities, behaviour or performance. Thus, the data on life activities reflect, for the most part, the specifics of routine life of a human and his/her daily interactions; the data on education contain evidence on formal or informal education, gained skills or competencies, etc.; the data on behaviour inform on the efficiency of a human in the society, helping reveal his/her potentials or threats; the data on performance give evidence of success on a job place (with sources of these data being ratings, results of interviews, etc.). Although analysis of data on experience still remains at the primary phase, their exclusive importance has already been understood: laying grounds for organization of analytical effort covering all the experiences accumulated by the humanity, for formulating logical conclusions and using results for various practical purposes, including the elimination of the warfare impact on the official statistics and the society on the whole.

When elaborating methodologies for the implementation of Big Data as alternative sources for the official statistics, some problems pertaining to this environment should be put in focus.

Cai et al. (2019) emphasize the problems related to data collection, storage and analysis queries. Kenglung H. (2022) prioritizes the access to reliable information in time of Big Data collection and analysis.

Making Big Data applicable for statistical problems requires the elaboration of a radically new policy on data management and protection, relations with respondents and the process of professional training of statistical staff.

With respect to the methodology, the problem involved in Big Data use in statistics is how to ensure their representativeness. Given the lack of a formalized general population as such, difficulties in defining a targeted population on the whole and sampling can occur.

It is known that conventional statistical observations are based on censuses and registers allowing to define all the types of populations required for conducting a statistical observation. However, with Big Data formed spontaneously, their random, unstructured and dynamic nature complicates the statistical process that is supposed as stable for the official statistics concept. Besides, the existing statistical methods are designed for a consistent, deep and long-lasting analysis of data in small sized samples, which is admitted to be a strong impediment for the statistical production process.

Another methodological problem in using Big Data for statistical purposes is how to ensure the relevance of technical support for statistical studies with respect to data quality measurement, the overall limitation of applications of data from external sources, the complexity of integrating information from different external origins in statistical databases to have a reliable end product.

It should be added that there also exist legal problems related with the regulation of access and use of Big Data, first and foremost with protection of personal data of residents and individual data of business enterprises, financial problems of optimizing the ratio of costs and gained advantages, administrative problems in policy-making with respect to the principles of data management and protection.

In the context of this article's theme, it should be noted that the warfare in Ukraine features the intensive involvement of information and communication technologies, especially ones based on artificial intellect, operated by the use of effectively designed and customized algorithms for Big Data processing.

In view of the above, it should be emphasized that smart technologies based on artificial intellect are popular in Ukraine, whose data are incorporated in one of the above mentioned global groups of Big Data ecosystem. Smart technologies are largely used by Ukrainians at household level, being installed in houses, cars or house territories. These technologies[3] have been applied by the Ukrainian Ministry for Digital Information to collect infinite numbers of video and photo evidence of violations of laws and customs of war, fixed in the agreements of Geneva Convention (Geneva, 1949).

Another essential source for Big Data is social networks. Their popularity, together with technologies based on artificial intellect, have allowed to find and identify servicemen of the Russian armed forces engaged in brutal massacres of civilian population in Ukraine in the towns of Bucha, Hostomel, Irpin in the Kyiv region of Ukraine (Ministry of Digital Transformation of Ukraine, 2022).

This successful experience in Big Data applications in the conditions of warfare gives evidence of the value and timeliness of this data type, in spite of the persisting "problematic" issues pertaining to their implementation, including in official statistics.

4. Digital framework for using Big Data in Ukraine

The implementation of Big Data in any kind of life activity calls for a certain level of technological readiness of a country as a whole and its sectors, with the adequate level of digitalization in the first place. Digitization is a current megatrend, meaning that digital technologies are integrated into our everyday life. The use of digital technologies enables the connection of different services and automation of many processes.

According to the OECD report (2017), the modern society has been a witness of the gradual expansion of digitalization that has impact on all the activities, it is on the way to the digital economy. This process began nearly 50 years ago, with the recent years marked by an incredibly higher pace at which digitalization penetrates the life.

[3] Titles and components of used technologies based on artificial intellect are not indicated by the authors for security considerations.

When reference is made to future scenarios of the social development, it is emphasized in the report that nowadays the world has plunged in the digital megalomania. First and foremost, it is caused by the rapid technological advancement in information technology sector and the rapid diffusion of Internet, entailing the increasing scopes of information, a kind of non-stop data. Mayer-Schönberger and Cukier (2013) labelled this phenomenon as datafication.

It should be noted in view of the above that Ukraine could create a positive self-image in the digital space. It was confirmed by estimations of Internetization Index proposed by the authors (Table 1) and Digital Transformation Index (Table 2).

Internetization Index is based on nine indicators: number of Internet subscribers, contention ratio, scope of provided Internet services, availability of the following ICT devices in households: personal PC, notebooks, Ipads, mobile phones, ratio of technological development, ratio of the availability of advanced technologies in Ukraine.

Because the data for these indicators are stimulators, their normalization was made by the maximal value over the period by the formula:

$$z_{ij} = \frac{x_{ij}}{x_{i,max}},$$

where x_{ij} is the value of i indicator in j year, $x_{i,max}$ is the maximal value of i indicator in the period under study.

Internetization Index was estimated by the formula:

$$I_j = \frac{\sum_i z_{ij}}{n},$$

where I_j is the value of Internetization Index in j year, n is the number of indicators determining Internetization Index, $n = 9$.

Table 1. Normalized indicators and Internetization Index

Year	z1	z2	z3	z4	z5	z6	z7	z8	z9	$\sum_i z_{ij}$	Internetization index, %
2016	0.642	0.663	0.725	1.000	0.771	0.789	0.990	0.850	0.956	7.387	82.1
2017	0.881	0.788	0.815	1.000	0.771	0.789	0.990	0.900	0.956	7.890	87.7
2018	0.907	0.894	0.893	0.974	1.000	1.000	1.000	0.950	0.978	8.595	95.5
2019	1.000	1.000	1.000	0.974	1.000	1.000	1.000	1.000	1.000	8.974	99.7

Data in Table 1 show that in the pre-war Ukraine there were favourable conditions for the penetration of digitalization.

But a good level of Internet penetration is not sufficient for the assessment of the society's readiness for global digitalization. Bearing this in mind, the authors proposed Social Progress Index to be used as a balance indicator, which is based on a set of indicators measuring the social development of a country in an integrated way. SDI used in this study consists of the indicators characterizing three areas of the social development: basic human needs, foundations of well-being, and opportunities (Global Index).

Basically, Social Progress Index can cover data on 60 indicators, including: access to information and communications, freedom of choice and personal freedom, access to basic education, access to higher education, personal security, tolerance, nutrition, basic medical service, quality of environment, etc.

A comprehensive description of the society's readiness for digitalization processes was made by the authors with use of Digital Transformation Index that was estimated as the arithmetic mean of the two sub-indices: Internetization Index and Social Progress Index.

Table 2. Sub-indices and Digital Transformation Index, %

Year	Internetization Index	Social Progress Index	Digital Transformation Index
2016	82.1	66.4	74.3
2017	87.7	68.4	78.1
2018	95.5	69.1	82.3
2019	99.7	66.9	83.3

Data in Table 2 show that the society's readiness for digital transformation in Ukraine was increasing year by year in the pre-war period, giving evidence that the country's ability to face the challenges of the modern digital world.

Thus, the EasyWay service (see: https://www.eway.in.ua) informing on all the routes and stops of public transport in 60 cities of Ukraine, as well as Poland, Moldova, Bulgaria. Uzbekistan, Serbia, Croatia, and Kazakhstan, was developed in Ukraine as early as in 2011. The main feature of this service is online monitoring of the movement of buses, trolleybuses, trams and urban trains via GPS.

It was in 2019 that the Ministry of Digital Transformation of Ukraine presented a mobile application and web portal under "Diia" brand. This application allows one to keep driver identification, internal and external passports, vaccination certificates (Covid-19) and other documents in one's smartphone, and to transmit their copies when receiving bank or post services, in hotel check-in or in other circumstances of daily life. Also, this application opens free access to some categories of public services: one can receive the comprehensive service in time of child birth (eMaliatko (eBaby)), register business online, pay taxes and submit statements, sign any category of documents, change the place of registration, etc.

It is expected that 100 percent of public services will be provided via “Diia” (see: https://diia.gov.ua/) until 2024. At the end of 2021, the application and portal was used by more than 12 million persons. The portal has provided access to 72 services, and the application – to 9 services and 15 digital documents. An important feature of the application is that it does not collect and store individual data of Ukrainians, being designed only for displaying the data filled in the official registers.

The brand has incorporated the newly created and already implemented public project “Diia City”: it is a fiscal and regulatory space for IT companies, designed for providing tax privileges and opportunities to apply advanced tools for attracting venture investments or protecting intellectual property. Earlier this year “Diia City” was included in the short list of Emerging Europe Awards 2022 in “Partnership” category, nomination “Modern and Future-Proof Policy”.

Also, Ukraine has successfully implemented the program for digitalization of health protection through designing and introducing a two-component system by which a user can interact with the central database via the online medical information system eHealth (see: https://ehealth.gov.ua/), which enables to computerize the operation of business entities in the heath protection sector, to create and review information, and to exchange information in electronic form.

Mobile operators in Ukraine have already launched mass-scale applications of Big Data for providing assistance to business: creating a client portrait, look-alike audience, target, trigger emails, geoanalitics, and financing scoring.

Four days after the beginning of the war, following the request of the Minister of digital transformation, Elon Musk opened access to high-speed Internet for Ukraine: his Starlink is providing assistance for the militaries and supporting the operation of critical infrastructures and services in medical care, finance and energy supply.

The abovementioned allows to admit that the expansion of digital processes across the Ukrainian territory lay a firm ground for using Big Data that have been gradually penetrating various walks of life.

5. Proposals

The main effort of Ukrainian statisticians now needs to focus not only on building the Big Data capacity or exchange of knowledge at international scales, but, first and foremost, on developing partnerships with the private sector for practical applications of gained knowledge and skills with the final end of involving and processing these new category of data. In the conditions of war, Big Data should become a supplement to the conventional statistical data or their alternative with an even higher level of detailing.

There is a large number of proposals on using Big Data in the official statistics. The process of implementing Bid Data in the official statistics must necessarily conform to the Fundamental Principles of Official Statistics (2014). A useful approach will be the reliance on the principles for Big Data analysis (Chen and Zhang, 2014) and 12 rules of Codd (1985), which have already been classical in building database management systems.

In view of the abovementioned and considering the necessity of using Big Data by the official statistics, we propose the following principles of Big Data handling:

1. Big Data collection procedure must guarantee their security and confidentiality, and must be backed by tests for reliability of quality of data.
2. Big Data collection and analysis, publishing and dissemination of the conclusions based on the results of their processing must have intellectual support.
3. The norms and rules pertaining to Big Data collection, processing and analysis, publishing and dissemination of the conclusions based on the results of their analysis must be agreed and publicized in the format of quadripartite group "government-science-business-statistics".
4. The guaranteed free access to official and private Big Data repositories for members "government-science-business-statistics" group.
5. Statistical departments must elaborate methodologies and recommendations on Big Data collection, processing and analysis considering the specific features of each Big Data group.
6. The algorithms for main processes of Big Data collection, processing and analysis must be made consistent and open to broader public.

Once the above principles are followed, it will allow for further elaboration and unification of methods on the way from "unordinary proposals on using Big Data" towards "adequate management decisions on the implementation of Big Data in various walks of life".

A greatly important factor in decision-making pertaining to statistical uses of Big Data is the choice of proper tools for their handling, first of all in IT technologies with applications of mathematical modelling methods.

A method of Big Data presentation that seems appropriate for statistics is Data Hypercube (Marchand, P., Brisebois et al., 2003), a unique organization of data in the form of a multidimensional set, the so-called cube of data.

The technique for this organization is a component of On-Line Analytical Processing (Codd et al., 1993), a technology for analytical data processing by the use of methods for collection, storage and analysis of multidimensional data for the support of decision-making processes. The cube structure allows for making a quick and multifaceted analysis from various perspectives, which can be fit for Big Data presentation for their further visualization.

An important step in the implementation of Big Data in the conventional statistical processes is the choice of appropriate computerized tools, applications and software required for these data handling.

A quite effective tool for extracting the necessary data is Web scraping. According to Dogucu M. & Çetinkaya-Rundel M. (2020), "Web scraping is the process of extracting data off the web programmatically and transforming it into a structured dataset. Web scraping allows for larger amounts of data to be collected in a shorter span of time and in an automated fashion that minimizes errors". Web scraping can be efficiently used in estimating the consumer price index (using websites of retail trade stores), analysis of the real estate market, etc.

However, this tool has weak sides. Summa D. et al. (2019) mention some problems that may be faced by statisticians in time of web scraping. They argue that online price data (web scraped) from the e-commerce platforms may not cover the full list of goods or services that the NSOs rely on in the compilation of the CPI. Technical issues include frequent changes in the website structure, the need to update crawlers or develop separate crawlers for different websites, the possibility of automatic blockage of high frequency web scraping, which calls for collaboration and partnership with web site owners. Another important challenge is the quality of data.

In spite of this, the experience of Italian National Institute of Statistics (Polidoro et al., 2015) in using Web scraping for the elaboration of new strategies for integrating data from conventional surveys, administrative bodies and innovative sources like Big Data, give evidence of positive results in reducing the statistical burden on respondents, allowing to create relevant databases and new formats of statistical registers.

It should be noted that the rapid expansion of digitalization processes stimulated many software developments for data analysis, which can be used for statistical purposes: Actian (a tool allowing to store raw data and their preparation for further analyses), Ambari (a tool for cluster management), Avro (a system of data serialization), Apache Kafka (a platform for data processing in the real-time format), Hive (an infrastructure for data storage, enabling for their aggregation), etc.

One of the high performing tools for analysis of Big Data with the open code is a collection of open-source software utilities that facilitate using a network of many computers to solve problems involving massive amounts of data and computation. It provides a software framework for distributed storage and processing of Big Data using the MapReduce programming model – Apache Hadoop. A notable advantage of Hadoop is easy storage and dissemination of large data sets on servers that can be operated in parallel, with data management by the use of clusters, thus enabling a method for storage based on shared file systems.

6. Conclusions

The results of this research could demonstrate that the needs of the official statistics in primary data can be effectively met in time of the martial law through using alternative sources, Big Data in the first place. These data should be introduced in parallel or in mix with conventional data sources, to fill the gaps in conventional data due to the war.

Nowadays, the important Internet sources of Big Data in Ukraine are applications designed by the government sector: “Diia”, eHealth, the platform of public procurement (Prozorro), the system of road motion sensors, consumption meters. Also, Ukraine has an extensive network of private digital services: online banking services (e.g. Monobank, the first neobank in Ukraine), a logistic application for express delivery (“Nova poshta”), market places (Prom, Rozetka, Bigl), etc. More “familiar” sources of Big Data (mobile phones, social networks, Google analytics, etc.) have to be considered, too.

Acknowledgement

The authors express sincere gratitude to the defenders of Ukraine for the opportunity to write this article in conditions that are safe for life.

References

Bentein, A., (2021). Data is the New Gold. Retrieved from: https://www.qad.com/blog/2021/04/data-is-the-new-gold#:~:text=Are%20the%20phrases%20%E2%80%9Cdata%20is,spend%20more%20money%20on%20technology%3F

Cai, L., Qi, Y., Wei, W., Wu, J., Li, J., (2019). mrMoulder: A recommendation-based adaptive parameter tuning approach for big data processing platform Future Gener. Comput. Syst., 93, pp. 570-582.

Codd, E. F., (1985). Is Your DBMS Really Relational? ComputerWorld.

Codd, E. F., Codd, S. B. and Salley, C. T., (1993). Providing OLAP to User-Analysts: An IT Mandate. Retrieved from: https://web.archive.org/web/20170808214004/https://www.minet.uni-jena.de/dbis/lehre/ss2005/sem_dwh/lit/Cod93.pdf

Cox, M., Ellsworth, D., (1997). Application-controlled demand paging for out-of-core visualization. Proceedings of the 8th conference on Visualization'97, pp. 235–247.

Chen, C. L. Philip and Zhang, C.-Ya, (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Science, 275, pp. 314–347.

Diebold, F. X., (2003). Big Data' Dynamic Factor Models for Macroeconomic Measurement and Forecasting. (Discussion of Reichlin and Watson papers), in M. Dewatripont, L.P. Hansen and S.Turnovsky (Eds.), Advances in Economics and Econometrics, Eighth World Congress of the Econometric Society. Cambridge: Cambridge University Press, pp. 115–122.

Dogucu, M., Çetinkaya-Rundel, M., (2021). Web Scraping in the Statistics and Data Science Curriculum: Challenges and Opportunities. Journal of Statistics and Data Science Education, Vol. 29, pp. 112–122.

ESP Rome, (2014). Collaboration in Research and Methodology for Official Statistics (CROS). Retrieved from: https://ec.europa.eu/eurostat/cros/content/esp-rome-2014_en.

European Commission, (2022). Data Act: Proposal for a Regulation on harmonised rules on fair access to and use of data. Retrieved from: https://digital-strategy.ec.europa.eu/en/library/data-act-proposal-regulation-harmonised-rules-fair-access-and-use-data.

European Commission, A European Strategy for data. Retrieved from: https://digital-strategy.ec.europa.eu/en/policies/strategy-data.

European Commission. (2020). Communication from the Commission to the European Parliament, the Council, the European economic and social committee and the Committee of the regions a European strategy for data. COM/2020/66 final. Retrieved from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066.

Gartner Glossary, (2022). Big Data. Retrieved from: https://www.gartner.com/it-glossary/big-data/.

Gartner Glossary, (2022a). Dark Data. Retrieved from: https://www.gartner.com/it-glossary/dark-datahttps://www.gartner.com/it-glossary/dark-data.

Geneva, (12 August 1949). Convention (IV) relative to the Protection of Civilian Persons in Time of War. Retrieved from: https://ihl-databases.icrc.org/applic/ihl/ihl.nsf/Treaty.xsp?documentId=AE2D398352C5B028C12563CD002D6B5C&action=openDocument.

Global Index: Overview. Social Progress Imperative. Retrieved from: https://www.socialprogress.org/index/global.

Henriques, A., (2022). Mind the gaps: how experience data can help fight climate change. World Economic Forum. Retrieved from: https://www.weforum.org/agenda/2022/02/climate-change-experience-data/.

Kenglung H., (2022). Big data analysis and optimization and platform components. Journal of King Saud University – Science. Vol. 34. Issue 4. Retrieved from: https://www.sciencedirect.com/science/article/pii/S1018364722001264#b0015.

Marchand, P., Brisebois A., Bédard Y. and Edwards, G., (2003). Implementation and evaluation of a hypercube-based method for spatio-temporal exploration and analysis. Workshop ISPRS, Québec, Canada.

Mayer-Schönberger, V., Cukier, K., (2013). Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt. Retrieved from: https://psycnet.apa.org/record/2013-17650-000/.

Ministry of Digital Transformation of Ukraine, (2022). Mykhailo Fedorov explains about recognizing the faces of the Russian occupiers. [Mykhaylo Fedorov poyasnyuye pro rozpiznavannya oblych rosiys′kykh okupantiv]. Retrieved from: https://www.kmu.gov.ua/news/mihajlo-fedorov-pro-rozpiznavannya-oblich-rosijskih-okupantiv [in Ukrainian].

OECD, (2017). Going Digital: Making the Transformation Work for Growth and Well-Being: Meeteng of the OECD Council at Ministerial Level. Retrieved from: https://www.oecd.org/mcm/documents/C-MIN-2017-4%20EN.pdf.

Oxford Learner's Dictionaries, (2022). Big Data. Retrieved from: https://www.oxfordlearnersdictionaries.com/us/definition/english/big-data.

Polidoro, F., Giannini, R., Conte R. L., Mosca, S. and Rossetti, F., (2015). Web scraping techniques to collect data on consumer electronics and airfares for Italian HICP compilation. Statistical Journal of the IAOS, 31(2), pp. 165-176.

Summa, D., Bianchi, G., Consalvi, M., Gentili, B., Pancella, F. and Scalfati, F., (2019). Using Big Data for Official Statistics: Web Scraping as a Data Source for Statistical Business Registers (SBRs). NTTS2019. Retrieved from: https://coms.events/ntts2019/data/abstracts/en/abstract_0007.html.

United Nations, (2013). Big Data and modernization of statistical systems. United Nations Economic and Social Council. Report of the Secretary-General. Forty-fifth session of the UN Statistical Commission, NY, 4-7 March 2014, doc. E/CN.3/2014/11, December, New York, pp. 1–16.

United Nations. Economic and Social Commission for Western Asia, (2021). Use of Big Data in Compilation of SDG Indicators in the Arab Region, Challenges and Opportunities. Retrieved from: https://www.unescwa.org/sites/default/files/pubs/pdf/big-data-compilation-sdg-indicators-arab-region-challenges-opportunities-english.pdf.

United Nations Statistics Devision, (2014). Fundamental Principles of Official Statistics (A/RES/68/261 from 29 January 2014). Retrieved from: https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.

Weiss, S. M., Indurkhya, N., (1998). Predictive data mining: a practical guide. Morgan Kaufmann Publishers, 1997, pp. 44–48.

Wiltshire, D., Alvanides, S., (2022). Ensuring the ethical use of Big Data: lessons from secure data access, Heliyon, Vol. 1. Issue, 2. pp. 1–6.

Wu, C., Buyya R., Ramamohanarao, K., (2016). Chapter 1. Big Data Analytics = Machine Learning + Cloud Computing. Big Data Principles and Paradigms, Elsevier, pp. 3–38.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 45–62, DOI 10.59170/stattrans-2023-004
Received – 06.07.2022; accepted – 19.10.2022

The impact of the Russian-Ukrainian war on the green transition and the energy crisis: Ukrainian scenario of circular economy development

Nataliia Reznikova[1], Iryna Zvarych[2], Roman Zvarych[3], Oksana Ivashchenko[4]

ABSTRACT

The aim of the article is to minimise the impact of the energy crisis on the environment as one of the ways of getting rid of carbon footprints resulting from the growth of the russian energy and building a circular sustainable ecosystem in Ukraine. Therefore, the study proposes a classification of green inflation and singles out the following forms of its manifestation: 1) as a reaction to irresponsible management; 2) as a reaction to an attempt to implement large-scale projects aimed at decarbonisation within strict time limits; 3) as a reaction to the strengthening of business standards, which the end consumer will have to pay for. The paper determines the impact that the war has on the practice of applying resource nationalism associated with a wide variety of modern global problems. Discussions focus particularly on the following topics: the 'resource curse', the growing concern of the OECD countries. The paper identifies the dominant diversification tendencies in the EU in terms of the circularity of the economy. The proposed concept of a global inclusive circular economy can be considered as a complex multidimensional system, whose main components are based on the economic, sociological, environmental and circular aspects of life.

Key words: circular ecomomy, green transition, energy crisis, russian -ukrainian war, humanitarian crisis, global inclusive circular ecomomy.

Main part of research paper. Despite a number of internationally initiated large-scale climate agreements defining the commitments to be made by both the public and private sectors in both developed and developing countries, the world community has

[1] Educational and Scientific Institute of International Relations Taras Shevchenko National University of Kyiv, Ukraine. E-mail: r_nv@ukr.net. ORCID: https://orcid.org/0000-0003-2570-869X.
[2] West Ukrainian National University, Ternopil, Ukraine.
E-mail: irazvarych@gmail.com. ORCID: http://orcid.org/0000-0001-5155-540X.
[3] West Ukrainian National University, Ternopil, Ukraine. E-mail: romazvarych@yahoo.com.
ORCID: http://orcid.org/0000-0003-3741-2642.
[4] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: oaivashchenko@nasoa.edu.ua.
ORCID: https://orcid.org/0000-0002-8490-778X.

© N. Reznikova, I. Zvarych, R. Zvarych, O. Ivashchenko. Article available under the CC BY-SA 4.0 licence

not come close to achieving the zero greenhouse gas emissions planned for 2050. Most of these commitments have not yet been accompanied by proper planning and financing, in part due to the global pandemic, which has already increased global public debt, reduced global GDP by 3.3% and led to rapid inflation around the world. The debate over who will pay for the green transition – governments, companies or consumers – and how to avoid increasing "green inflation" is getting hotter. The term "green inflation" explains the "price vicissitudes" that may occur on the way to the environment in the future (as a result of increased demand for selected resources), as well as the costs associated with reducing carbon emissions as part of "green" packages investment and financing of relevant projects that increase the money supply, or as a result of limited supply of goods due to climate catastrophes or environmental degradation. In other words, green inflation can occur: 1) as a response to irresponsible management (where economic agents have actively created a reality in which environmental degradation and climate problems lead to food, raw materials, energy and other crises that reduce supply and provoke rising prices); 2) as a reaction to the attempt to implement large-scale projects aimed at decarbonisation in a limited time (where economic agents and consumers become "victims" of monetary and fiscal policies of governments); 3) as a reaction to the strengthening of business standards, which will have to be paid by the final consumer (the price of minerals needed for "green" technologies – from wind turbines and solar panels to electric vehicles – will require a "green allowance" (greenium) as a relatively higher demand for them with limited supply, and because the extraction of minerals is usually associated with higher environmental costs, which will be "punished" by taxes). The relationship between climate change and political risks is reflected in growing geopolitical challenges, especially in finding a compromise between the need to fund climate neutrality and the demands of developing countries to give them more leeway in achieving their emission reduction targets, maintaining the pace of economic development needed to lift their populations out of the poverty trap.

Raw materials and energy crises, disruptions in value chains and supply chains have not contributed to progress towards the green transition (Melnyk et al., 2020) exacerbating the risks of deglobalization (Ivashchenko and Reznikova, 2018), despite governments' awareness of their inability to find solutions on their own, including for food problem. The war in Ukraine has exacerbated all these trends, both nationally and globally, by shifting priorities and shifting the focus of world leaders from climate neutrality to national defence, energy, food and humanitarian security. For example, transforming the energy sector and reducing emissions by a third, compared to the BAU scenario (business as usual), may increase India's GDP by 3.9%, China's by 1.4% and Indonesia's by 2.4%, low-income African countries – by 2% and ASEAN countries – by 1.6% (Reznikova, 2021).

The Sixth Expert Report, published a few days after Russia's invasion of Ukraine by a UN government group on climate change (IPCC, 2022) noted that climate change, which has a significant impact on human and natural systems, could in itself increase the risk of future conflicts, within and between countries due to differences in access and distribution of resources such as energy, food and water. As the effects of climate change will scale at unpredictable rates and non-linearly, the possibility of avoiding the catastrophic effects of climate change is rapidly closing. At the same time, the prospect of aggravation of energy shortages is estimated to accelerate the development of the green energy market. Along with this, stimulating investment in renewable energy, energy efficiency and decarbonization, some experts traditionally continue to attribute to the factors of rising energy prices and instability, while others see them as the key to solving these problems. By the end of the decade, global clean energy markets are expected to reach $ 23 trillion through the development of new technologies to support deep decarbonisation, energy storage, semiconductors and nuclear energy.

The energy crisis is creating new fault lines in the world economy. To the usual division of countries according to the level of availability of fossil fuel deposits and available energy resources, the division was added: 1) according to the degree of supply of "green" minerals; 2) the degree of susceptibility to rising energy prices (which depends not only on the level of diversification of energy sources and suppliers, but also on the structure of the economy and the country's specialization in the international division of labour); 3) the degree of transition to production with zero emissions; 4) the degree of participation in the chain of creating the cost of energy transition.

Some countries rich in green minerals, in their quest to move up the value chain of the energy transition, restrict exports of relevant raw materials and minerals, requiring value added to remain in the country. For example, Indonesia's ban on export sales of nickel ore as part of plans to expand the domestic metallurgical industry may be adopted by other countries. Some governments may go even further to establish state control over the production and export of green minerals. Mexico plans to give the state exclusive control over lithium production, and China has announced a state-owned enterprise that will control about 70% of domestic production, which depends on rare earth metals and elements. The fact that the extraction of "green" minerals can lead to significant deterioration of the environment on a global scale is already being used by countries importing such minerals as an argument against the introduction of resource nationalism.

The concept of resource nationalism can easily be linked to a wide variety of contemporary global issues, including: discussions of the "resource curse"; growing concern from OECD countries about the economic importance of sovereign wealth funds; "Good governance" in middle- and low-income countries; trends in risk management in the political sphere; practical efforts to develop a methodology for measuring the contribution of socio-economic development of extractive industries and a wider range of issues regarding "favourable conditions" for responsible business practice. The context in which resource nationalism is becoming more pronounced includes: high food prices combined with increasing pressure on land fertility in the process of intensifying biofuel production; increasing the economic (and, consequently, political) importance of national oil companies; increasing the problems of security of energy and natural resources; rapid economic growth and domestic consumption in a number of middle-income countries, including India and China.

Given that countries rich in the resources needed for the green and digital transition are essentially the strategic resource for a new technological revolution, but the intermediate technologies for providing it are in capital-rich countries, a web of trade-offs that needs to be unravelled by market participants, only exacerbates the uncertainty of green markets.

According to McKinsey's expert estimates, the transition to zero balance is impossible without taking into account the nine key positions, which are grouped into three categories: 1) the necessary physical infrastructure or the so-called "Structural units" (technological innovations; the ability to create large-scale supply chains and ancillary infrastructure; the availability of the necessary natural resources); 2) economic and social change (effective structures of redistribution of capital and financing; management of changes in demand and short-term increase in unit costs; compensation mechanisms to eliminate socio-economic consequences); 3) governance, institutions and commitments (effective institutions, standards, monitoring system, effective market mechanisms; commitment and cooperation of leaders at public and private levels to support citizens and consumers, including public support for progress in reducing greenhouse gases) (Krishnan et al., 2021).

Awareness of the potential impact of the war on each of these positions will assess the prospects for the transition to zero emissions. Due to the suspension of production processes, destruction of existing industrial facilities and the deployment of large-scale deindustrialization of the Ukrainian economy in the short term, the availability of the necessary physical infrastructure and provision of the above structural units in Ukraine will be reduced. There are signs and consequences of the introduction of economic sanctions against Russia and the blocking of economic cooperation between a number of countries. At the same time, the growth of consumption of energy-intensive goods in the near future will stimulate the demand for high-tech innovations that will be able to compensate for increased emissions through carbon capture and sequestration. Since the beginning of the war on a global scale, there has been a significant influx of capital into the renewable energy sector, which has changed the downward trend in the pandemic period. Along with this, although in the short term the desire to expand zero-emission infrastructure may increase, its implementation may be hampered by logistical stresses related to market reorganization (sanctions) and rising energy prices (due to industrial and green inflation), which may lead to shocks in the complex hierarchical multinational transport-intensive supply chains of “zero emission” technologies (Samandari et al., 2022).

However, the impact of energy shortages on the level of investment attractiveness of zero-emission technologies cannot be unambiguously assessed, as investments in renewable energy require large one-time capital expenditures but minimal operating costs, which means that rising production costs may affect the energy sector less than sustainable growth prices for energy resources. And while Europe has traditionally been more receptive to rising fossil fuel prices, the current resource supply shocks will be felt in both Europe and the United States, where energy prices are increasing production costs and forcing price and supply chains to be revised. It is noteworthy that some large zero-emission technology countries are not involved in sanctions against Russia and have been able to maintain access to supplies, which affects the costs of their trading partners and the position of competitors who have joined the sanctions.

The transition to zero emissions in the short term depends on such a factor of production as land. Unfavourable climate forecasts for 2022 indicate that yields may fall on their own due to natural causes, which will reduce the supply of a number of agricultural products. Supply shortages and rising prices in agricultural markets could lead to increased deforestation in countries most sensitive to food security.

In the short term, the economic and social changes needed to achieve more orderly climate neutrality (zero emissions) depend on managing changes in demand and unit costs, introducing compensatory mechanisms to address the socio-economic consequences of the green transition and ensuring efficient capital allocation and financing (Zhong and Zvarych, 2022). As a result, rising energy costs are approaching the break-even point of carbon footprint reduction solutions for many industries. At the same time, the supply shock - the shortage of goods - stimulates the movement towards increased processing, and thus the demand for a circular economy and the green transition can move into the practical field of their implementation at a faster pace.

The war in Ukraine will negatively affect the movement of international capital flows from developed countries to countries with the highest level of debt burden. In fact, the movement of capital served as a kind of compensatory mechanisms in the global asymmetries of development. Given that even before the war capital inflows to developing countries were almost 20% lower than previously developed countries' promises of $ 100 billion in annual aid by 2020, further cuts in aid are no longer in doubt.

Given that the preparation and commissioning of renewable alternative energy sources takes time, and rising energy prices are provoking an economic and political crisis that threatens to provoke mass unrest and destabilization around the world, according to McKinsey experts, growth is expected in the short term funding for fossil fuel extraction to control further price increases. In Europe, rising energy prices will lead to increased short-term investment in fossil fuel production and consumption, especially through existing or recently decommissioned assets. This will help diversify the sources of mineral imports, both in terms of pricing and energy security policy. The abandonment of Russian gas involves finding ways to overcome logistical barriers, develop import capacity and time to sign new contracts to circumvent the limitations of existing pipeline infrastructure. After the invasion of Ukraine, the EU countries have already signed contracts for the purchase of Russian gas worth more than 46 billion dollars. In a situation where lower prices are not possible due to increased domestic production or diversification of sources, a return to cheaper but more emitting fuels, such as coal, is already happening in Germany and Poland. In the short term, the United States will also maintain a trend of increasing fossil fuel production to address rising domestic prices and support diversification of supplies to Europe. At the same time, in the medium term, the United States has the opportunity to reduce fossil fuel consumption through energy efficiency policies. Asia, which has time to enter the markets directly to build alternative pipelines, has already announced its readiness to increase the use of cheap coal, which is less regulated than oil.

The war could jeopardize the further development of international standards, institutions and institutions necessary for a green transition. At the same time, many large economies, including China, use the transition to zero emissions as an excuse to increase the state's presence in the economy given their large-scale investment in the production of “green” technologies and components. For example, China's leadership, which already produces a third of the world's wind turbines, 70 percent of the world's solar photovoltaic power and 75 percent of the world's lithium-ion battery capacity, could be significantly strengthened by recession and unprecedented industrial inflation in Europe.

Thus, according to McKinsey, the war in Ukraine will delay the transition to climate neutrality, and a broken transition, in turn, could exacerbate the catastrophic effects of global warming and lead to new conflicts between countries that will affect the energy crisis and climate change. Exit from the post-conflict transition is possible only if there is a unity of perception of the climate threat and the price of energy security.

Depending on the scale and duration of market failures, the timeliness and effectiveness of government support and changes in central bank policies, consumer and business responses, McKinsey identified two most likely potential scenarios for Russia's war against Ukraine on EU developments (White et al., 2022):

Moderately optimistic scenario – in the event of the end of hostilities within a few weeks and moderate sanctions against the aggressor country was expected in 2022 to restore GDP growth to pre-pandemic levels, controlled inflation by reducing monetary incentives ECB, recovery of European investment and consumer already in the second quarter;

Pessimistic scenario – describes a set of strong market failures due to prolonged hostilities and strong sanctions that will lead to the cessation of oil and gas exports from Russia to Europe and includes, in particular, the following: migration crisis in Europe; reduction of economic growth and post-pandemic recovery; inflation over 7%; return to coal use and suspension of green initiatives; creation of a new infrastructure for export/import of liquefied gas; reduction of real incomes and consumer spending.

As we can see, the moderately optimistic scenario was not justified, so, according to McKinsey experts, the war in Ukraine will cause 12 shocks that will change the world (White et al., 2022).

The invasion of Ukraine is causing a large-scale humanitarian crisis.

According to UN estimates, by the end of 2022 more than 8.3 million Ukrainians may be refugees. Given the capacity of governments that receive forced migrants from Ukraine to provide financial assistance and humanitarian support, including accommodation, food, access to basic health services, and employment promotion, there is a growing demand for new international humanitarian assistance programs, such as the United Nations Regional Refugee Response Plan (RRP). The RRP consolidates measures to support countries' efforts to protect and assist refugees coming from Ukraine and includes the initial financial needs of 12 partners (including UN agencies, national and international NGOs and civil society) for six months (UNOCHA, 2022). In the long run, Ukrainians who integrate into society and the labour market will be able to enjoy the benefits of immigration.

Vulnerable groups will suffer the most from the war.

It is estimated that the food price index compiled by the United Nations Office on Food and Agriculture may increase by 45% in 2022. Typically, rising prices of this magnitude have historically pushed millions of people into low- and middle-income countries into poverty. These countries will be the first to feel the effects of the war, such as the slowdown in world trade, currency devaluation and debt service problems.

Energy policy will provide for the diversification of energy sources and ensuring uninterrupted access to them.

Governments will work to build infrastructure for liquefied natural gas and biofuel production, while diversifying supplies by phasing out Russian gas. Industrial consumers may face a reduction in gas supplies earlier than households. There will be a reduction in the use of gas in the production of electricity, and private consumers are invited to get used to energy asceticism.

Food security on the global agenda.

Ukraine and Russia account for about a third of the world's exports of the main fertilizer ingredients, ammonia and potassium, supplying about 30% of world exports of wheat and barley, 65% of oil and 15% of corn. In the first weeks after the invasion, prices for fertilizers and some foodstuffs rose by 20-50%. Ukraine and Russia provided about 20% of the total volume of food purchased under the UN World Food Program in 2020. As a result, food aid recipient countries will experience food shortages. Global fertilizer shortages can also hamper food production. This will force governments to implement programs to increase supplies to countries that may be on the brink of starvation, as well as to look for ways to increase regional agricultural production. Governments are expected to use the tool to subsidize consumers and control the prices of agricultural products.

Increasing competition for the most important materials, equipment and goods.

The war accelerated the growth of prices for dozens of goods exported by Ukraine and Russia (e.g. coal, steel, nickel), which ranged from 10 to 50% on various commodity items (in particular, 48% of world trade in palladium). Automakers expect spot prices to increase by 15–25% due to higher prices for key materials such as aluminium, copper and steel. It is believed that iron ore will fall in price, while the price of anthracite will increase significantly. Prices for automotive products will not return to pre-war levels.

A new era of supply chain management.

Supply chain managers have begun to shift their focus from optimizing deliveries on time to preparing for unforeseen circumstances just in case. Reorganizing supply chains to increase their sustainability can mean managing a large number of relationships between manufacturers, distributors and consumers. For example, due to the interdependence between the participants in the supply chain, a change in the source of supply or manufacturer may cause fluctuations throughout the chain. There is a demand for two different suppliers – **double sourcing**, although both in terms of supply management and planning, it will not only complicate the supply process, but also increase the cost of goods. A new phenomenon is emerging – **"friendshoring"**, which is a process of cooperation, which is based on the commitment to work with countries that are firmly committed to democratic norms and values.

Coexistence of different and sometimes competing technological standards.

Many countries that have used the tools of digital neoprotectionism, the conscious use of the Internet to control e-commerce to manage consumer demand, will continue to implement excellent technical standards (Internet protocols), but security issues are now at the forefront. The mass exodus of many leading Western companies from Russia means that Russia is effectively excluded from much of the global high-value chain. About 80% of Western technology companies have left Russia or are reducing their presence. The introduction of excellent technology standards will mean more expensive services for consumers and lower productivity growth worldwide.

Uncertain impact on the financial system.

The war exacerbates the risks to the financial system that emerged in 2021, such as: the recession caused by inflation; growth in dollar-denominated borrowing by emerging market countries; speculative bubble in China's real estate sector; increasing risk of non-repayment of loans circulating in the shadow banking sector. However, the recession caused by inflation may be the greatest threat to the stability of the financial system.

Rising defence spending.

Fifteen NATO countries and Sweden have announced increased defence spending since the invasion of Ukraine, and five (including Denmark, Germany, Italy, Spain and Sweden) will exceed the 2% target set at the 2014 NATO summit in Wales. McKinsey's analysis shows that the increase in costs in many countries is likely to go to equipment, as many weapons programs have been reduced. If that happens, countries will have to choose between immediate or long-term investment in armaments. If they decide to buy immediately, this could be a problem for manufacturers, who will have to deal with supply chain disruptions that will only intensify due to the need to scale production, as well as the potential lack of materials such as titanium, platinum group metals and more.

Cyber conflicts.

The risk of cyberwar between countries (attacks on energy systems and telecommunications systems) and cyber attacks in the commercial sector are forcing cybersecurity budgets to increase the availability and value of a range of services and goods.

The reaction of the corporate sector to the war as a manifestation of consolidated resistance to Russian aggression.

Of the 281 Fortune 500 companies that worked in Russia before the war, about 70 percent either cut or shut down after the war. Almost 85% of companies headquartered in Europe, the United Kingdom or the United States have left or reduced their operations in Russia, compared to 40% of companies based in other regions.

Volatility as a new norm.

Although instability in energy supplies and rising energy prices could have serious consequences for the global economy, according to the McKinsey Global Survey, the war has led to significant volatility in the risks that business leaders see for economic growth. In March 2022, geopolitical risk was identified as a major threat to growth, suppressing the pandemic and inflation.

Green Recovery and Circular Economy For World and Ukraine

The transition from a linear to a circular economy is determined by the change in the positioning of global risks from year to year, which determines the vectors of such a change. Thus, the main risks for 2020 are those related to the environment and occupying the first positions in the rating for the last 3 years (in particular, in 2019, the risk of extreme weather events). The gradual increase in awareness of this risk has led to a change in the sentiment of both producers and consumers. Experts have assessed climate change as a major risk in 2019, outpacing cyberattacks, financial instability and terrorism. Thus, to mitigate this risk in 2020, the global business community has been proposed to implement circular "designs", reducing resource use and prioritizing low carbon materials. The potential effects of the transition to a circular economy on greenhouse gas emissions are significant, mainly achieved by improving resource efficiency, increasing the useful life of buildings and assets, increasing recycling and reuse, and completely reducing primary raw material use. Thus, the circular economy can be seen as an effective strategy for promoting climate change mitigation.

The poly-paradigmatic nature of economic-theoretical knowledge from the standpoint of the existential nature of the imperatives of economic development in the context of responsibility to the global future causes certain paradigm shifts, and greening is the mainstream and imperative that reflects the heterodox beginning of this methodological approach. Multidisciplinary epistemological perspective for the analysis of economic phenomena from the standpoint of the new pragmatism in the context of the "triad" of sustainability (economic, social and ecological components), distinguishes the humanitarianism of economic science and positions "in the foreground" the role of values in the economic activity of people and society.

Acceleration of exploitation of natural resources; climate change; the formation of a new environmental order; environmental and food security) and the dominant paradigm formation and (exceeding the environmental limit; Paris agreement; changing public sentiment of fossil producers and businesses; global economic losses; UN sustainable development goals with strong circular practices; WTO involvement in supporting new technologies, minimizing waste production and promoting circular trade). The methodological features of the paradigm of a global inclusive circular economy from the standpoint of postmodernism are the strengthening of socio-humanistic orientations (reflecting its ideology and creating the basis for an inclusive-oriented society); ontological nonlinearity (emphasizes circularity); consensus (the need for a global consensus to achieve a goal) emulated using the economic-mathematical method.

The conceptual provisions of the theoretical and methodological model of global inclusive circular economy are substantiated in the work. System characterises by the sustainable development; stability; inclusive growth; expansion of opportunities of creation of own capital; equality of access to resources and distribution of goods.

This approach gave the opportunity to substantiate the basic concepts of paradigm formation of the global inclusive circular economy: global chains of value creation (supply of recyclables), sharing platforms, circular trade, circular product life cycle, circular cities and circular cores.

The theoretical and methodological model of global inclusive circular economy is developed based on: system characteristics (sustainable development; stability; inclusive growth; expansion of opportunities of creation of own capital; equality of access to resources and distribution of goods) and principles of circular self-organization (conservation of resources and conservation of resources) rationalization of their use; counteraction to management dysfunction), determined by the critical constraints of corporate culture, communications in global chains.

The paper identifies dominant tendencies of diversification of the European Union in terms of circularity of the economy. In particular, the highest level of household waste recycling is identified in Germany. Belgium, Denmark, France, Italy have consistently high positions. In general, the EU (28) is characterized by an average level of waste recycling (50–60%). The highest value is characteristic of Luxembourg (90%), Belgium (80%), Slovenia (60–70%), the Netherlands (70%), Italy (60–70%). The lowest rates are found in Greece (10–20%), Bulgaria (15–20%), Estonia (20–22%). The countries with the highest positive dynamics are distinguished: Croatia (25 → 40%), Latvia (40 → 61%) and Slovenia (50 → 75%).

For the most comprehensive consideration of the full range of global inclusive circular economy opportunities, an integrated index of global inclusive circular economy (*Igice*) development by ecological, economic, social and circular components with isolated weakly correlated individually ranked indices is developed and analyzed accordingly, on the basis of which the circular cores are separated: the social component (Belgium, Czech Republic, USA, China, France, Greece, Austria, Australia); environmental component (Japan, Denmark); economic component (Germany, China), which became the basis for the model of formation of global inclusive circular chains. The conducted cluster analysis based on the components of the indexes of the global inclusive circular economy confirmed the formation of a large circular gap in environmental and circular components. Graphical data visualization clearly reflected the circular circle that the index forms.

The systematic analysis of the spatial-component structure of the waste and scrap trade has been carried out and the leading countries, development priorities in the sectors (pharmacy, clinical, household, rubber, polymers, silk and cotton waste) have been identified. The main problematic aspects are illegal waste trading and growing smuggling, which cause serious negative social consequences and actualize the inclusive component in justifying the paradigm of a global inclusive circular economy (Shnyrkov et al., 2019). It has been researched that waste reduction coupled with rational use of resources has the potential to close the gap due to scarcity of natural resources and global population growth or consumption. It is substantiated that the formation of circular trade will contribute to: determination of priority materials for trade and the required level of processing capacity; harmonization of quality standards of materials; promoting the demand for used goods and recyclables; eliminating unnecessary regulatory barriers and avoiding environmentally damaging activities such as non-compliance, poorly regulated nature and informal recovery.

According to the proposed concept of a global inclusive circular economy, it can be considered as a complex multidimensional system, the main components of which are economic, sociological, environmental and circular aspects of life. To fulfil this task, the indicators of the relevant countries and statistics of the Organization for Economic Cooperation and Development for 28 member countries from 1995 to 2017 were used: Australia, Austria, Belgium, Great Britain, Greece, Denmark, Estonia, Israel, Canada, China, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, Germany, New Zealand, South Africa, South Korea, Poland, USA, Turkey, Hungary, Ukraine, Finland, France, Czech Republic, Japan. The analysis is based on the GNU Regression, Econometrics and Time-series Library - an application software package for econometric modelling, part of the GNU project.

Thus, Table 1 shows the results of the economic component of the indicator, in particular, the regression functions and uncorrelated indicators, respectively, for each country studied. Therefore, based on the analysis for Australia, the eigenvalue of the correlation matrix (i.e. the characterizing number at which a solution can be obtained to describe the global inclusive circular economy), which characterizes the absolute contribution (significance) of the main principal component, is $\lambda_y (\vec{x}_1) =$ 3.6582 and describes the solution of the problem by 73% of all processes, which indicates its sufficient adequacy.

Table 1. The results of the economic component of the indicator (the regression functions and uncorrelated indicators)

Country	Regression function% comp.	% comp.	Low-correlated indicators[5]
Australia	$y(\vec{x}_1) = 0.485 \cdot x_{11} + 0.475 \cdot x_{12} + 0.5 \cdot x_{13} - 0.454 \cdot x_{15} + 0.284 \cdot x_{16}.$	73%	x_{14}
Austria	$y(\vec{x}_1) = -0.5 \cdot x_{11} - 0.364 \cdot x_{12} - 0.493 \cdot x_{13} - 0.494 \cdot x_{15} + 0.36 \cdot x_{16}.$	76%	
Belgium	$y(\vec{x}_1) = 0.483 \cdot x_{11} + 0.473 \cdot x_{12} + 0.5 \cdot x_{13} + 0.362 \cdot x_{14} + 0.4 \cdot x_{15}.$	76%	x_{16}
Greece	$y(\vec{x}_1) = 0.494 \cdot x_{11} + 0.483 \cdot x_{12} + 0.382 \cdot x_{13} + 0.437 \cdot x_{14} + 0.431 \cdot x_{15}$	73%	x_{16}
Denmark	$y(\vec{x}_1) = 0.508 \cdot x_{11} + 0.504 \cdot x_{12} + 0.499 \cdot x_{13} + 0.177 \cdot x_{15} + 0.455 \cdot x_{16}$	76%	x_{14}
Great Britain	$y(\vec{x}_1) = 0.517 \cdot x_{11} + 0.52 \cdot x_{12} + 0.521 \cdot x_{13} + 0.437 \cdot x_{16}$	89%	$x_{14}. x_{15}$
Estonia	$y(\vec{x}_1) = 0.513 \cdot x_{11} + 0.467 \cdot x_{12} + 0.514 \cdot x_{13} + 0.316 \cdot x_{14} + 0.393 \cdot x_{16}$	72%	x_{15}
Israel	$y(\vec{x}_1) = -0.587 \cdot x_{11} - 0.52 \cdot x_{12} - 0.52 \cdot x_{13} - 0.276 x_{14} + 0.339 x_{15}$	67%	x_{16}
Canada	$y(\vec{x}_1) = -0.538 \cdot x_{11} - 0.551 \cdot x_{13} + 0.558 \cdot x_{15} + 0.307 \cdot x_{16}$	97%	$x_{12}\ .x_{14}$
China	$y(\vec{x}_1) = -0.452 \cdot x_{11} + 0.443 \cdot x_{12} - 0.462 \cdot x_{13} + 0.455 \cdot x_{14} + 0.423 \cdot x_{15}$	92%	x_{16}
Latvia	$y(\vec{x}_1) = 0.482 \cdot x_{11} + 0.482 \cdot x_{12} + 0.483 \cdot x_{13} + 0.16 \cdot x_{14} + 0.327 \cdot x_{15} + 0.411 \cdot x_{16}.$	66%	
Lithuania	$y(\vec{x}_1) = -0.516 \cdot x_{11} + 0.514 \cdot x_{12} + 0.520 \cdot x_{13} - 0.428 \cdot x_{15} + 0.122 \cdot x_{16}.$	72%	x_{14}
Luxembourg	$y(\vec{x}_1) = -0.476 \cdot x_{11} - 0.4 \cdot x_{12} - 0.477 \cdot x_{13} - 0.054 \cdot x_{14} + 0.459 \cdot x_{15} + 0.416 \cdot x_{16}$	72%	
Mexico	$y(\vec{x}_1) = 0.493 \cdot x_{11} + 0.461 \cdot x_{12} + 0.499 \cdot x_{13} + 0.483 \cdot x_{15} + 0.248 \cdot x_{16}.$	77%	x_{14}
Netherlands	$y(\vec{x}_1) = 0.435 \cdot x_{11} + 0.444 \cdot x_{12} + 0.432 \cdot x_{13} + 0.467 \cdot x_{14} + 0.457 \cdot x_{15}.$	88%	x_{16}
Germany	$y(\vec{x}_1) = -0.446 \cdot x_{11} - 0.423 \cdot x_{12} - 0.442 \cdot x_{13} - 0.299 \cdot x_{14} + 0.344 \cdot x_{15} + 0.38 \cdot x_{16}.$	83%	
New Zealand	$y(\vec{x}_1) = -0.462 \cdot x_{11} - 0.499 \cdot x_{12} - 0.425 \cdot x_{13} - -0.427 \cdot x_{15} + 0.417 \cdot x_{16}.$	75%	x_{14}
South Africa	$y(\vec{x}_1) = -0.556 \cdot x_{11} + 0.132 \cdot x_{13} + 0.576 \cdot x_{14} + 0.585 \cdot x_{15}$	67%	$x_{12}\ .x_{16}$
South Korea	$y(\vec{x}_1) = 0.516 \cdot x_{11} + 0.524 \cdot x_{12} + 0.521 \cdot x_{13} + 0.433 \cdot x_{15}.$	87%	$x_{14}\ .x_{16}$
Poland	$y(\vec{x}_1) = 0.453 \cdot x_{11} + 0.455 \cdot x_{12} + 0.451 \cdot x_{13} + 0.435 \cdot x_{14} - 0.438 \cdot x_{15}.$	79%	x_{16}

[5] x_11 - GDP, million $;
x_12 - employment, %;
x_13 - real GDP per capita, $;
x_14 - GDP growth, %;
x_15 - environmental taxes, % of total tax revenues;
x_16 - subsidies for the development of environmental technologies related to the environment, % of total aid.

Table 1. The results of the economic component of the indicator (the regression functions and uncorrelated indicators) (cont.)

Country	Regression function% comp.	% comp.	Low-correlated indicators[6]
Turkey	$y(\vec{x}_1) = -0.536 \cdot x_{11} - 0.529 \cdot x_{12} - 0.531 \cdot x_{13} + 0.389 \cdot x_{15}.$	85%	x_{14} . x_{16}
Hungary	$y(\vec{x}_1) = -0.468 \cdot x_{11} - 0.474 \cdot x_{12} - 0.466 \cdot x_{13} - 0.474 \cdot x_{14} + 0.34 \cdot x_{15}$	76%	x_{16}
Finland	$y(\vec{x}_1) = 0.536 \cdot x_{11} + 0.492 \cdot x_{13} + 0.478 \cdot x_1 + 0.379 \cdot x_{15} + 0.315 \cdot x_{16}.$	62%	x_{12}
France	$y(\vec{x}_1) = -0.522 \cdot x_{11} + 0.112 \cdot x_{12} - 0.514 \cdot x_{13} - 0.497 \cdot x_{14} - 0.014 \cdot x_{15} + 0.451 \cdot x_{16}$	79%	
Czech Republic	$y(\vec{x}_1) = -0.433 \cdot x_{11} - 0.435 \cdot x_{12} - 0.436 \cdot x_{13} - 0.364 \cdot x_{14} + 0.436 \cdot x_{15} + 0.334 \cdot x_{16}$	80%	
Japan	$y(\vec{x}_1) = -0.489 \cdot x_{11} - 0.503 \cdot x_{12} - 0.476 \cdot x_{13} + 0.205 \cdot x_{14} + 0.49 \cdot x_{15}$	76%	x_{16}

Similarly, a study was conducted for all proposed countries. The following indicators turned out to be weakly correlated indicators for the economic component:

GDP growth, %:

Australia, Denmark, the United Kingdom, Canada, Lithuania, Mexico, New Zealand, South Korea and Turkey.

Environmental taxes, % of total tax revenues:

Great Britain, Estonia.

Subsidies for the development of environmental technologies related to the environment, % of total aid:

Belgium, Greece, Israel, China, the Netherlands, South Africa, South Korea, Poland, Turkey, Hungary, Japan.

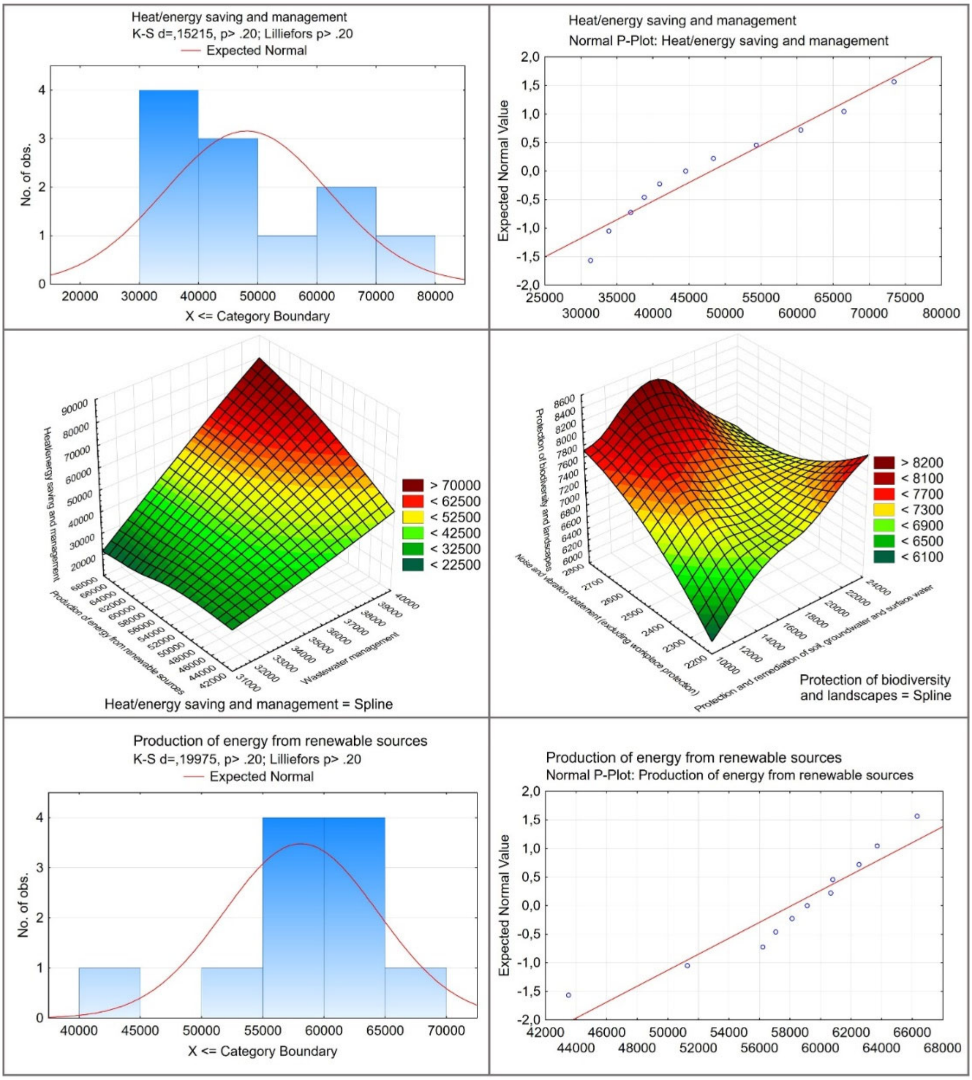
Employment, %:

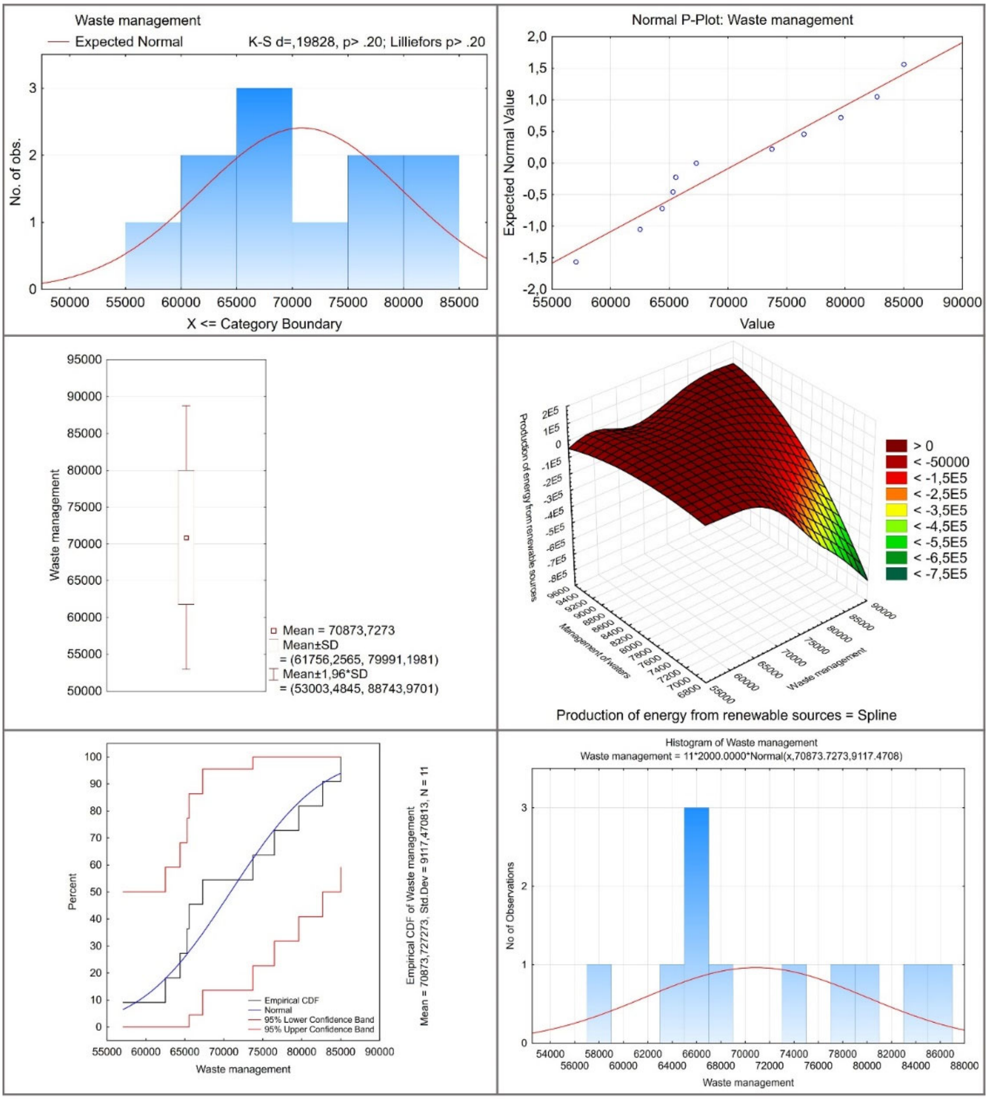
Canada, Finland.

The following indicators lead to a decrease in the impact of the economic component in the overall integrated indicator of the global inclusive circular economy in the following countries.

[6] x_11 - GDP, million $;
x_12 - employment, %;
x_13 - real GDP per capita, $;
x_14 - GDP growth, %;
x_15 - environmental taxes, % of total tax revenues;
x_16 - subsidies for the development of environmental technologies related to the environment, % of total aid.

That is, the increase in the weight of reflected indicators affects the main component among the many indicators of the economic component, which generally frees up space for other components in the integrated indicators of a global inclusive circular economy, such as circular, environmental or social (or inclusive) components.





The link between the Sustainable Development Goals and the opportunities for the formation of a circular environment in Ukraine for the period up to 2030 is apparent. Emphasis is placed on addressing the issue of circular circuit consumption in the areas of local collection, establishment of food hubs and recycling outsourcing. The role of extended responsibility for accepting returned products and wastes remaining after using such manufacturer's products as a necessary element in shaping Ukraine's circular policy has been established.

Conclusion

The war in Ukraine will affect all forms of international economic relations, highlighting the problem of asymmetric economic interdependence in the green transition to climate neutrality, accompanied by raw materials, energy and food crises. The trap of green development is that stimulating investment in renewable energy, energy efficiency and decarbonisation, some experts have traditionally continued to attribute to rising energy prices and instability, while others see them as the key to solving these problems. Since the beginning of the war on a global scale, there has been a significant influx of capital into the renewable energy sector, which has changed the downward trend in the pandemic period. At the same time, although in the short term the desire to expand zero-emission infrastructure may increase, its implementation may be hampered by logistical stresses related to market reorganization (sanctions) and rising energy prices (due to industrial and green inflation), which may to lead to shocks in complex hierarchical multinational transport-intensive supply chains of zero-emission technologies green transition. Some large zero-emission technology countries are not involved in sanctions against Russia and have been able to maintain access to supplies, which affects the costs of their trading partners and the position of competitors who have joined the sanctions. According to McKinsey, the war in Ukraine will cause twelve upheavals that will have a fundamental impact on international economic development: the invasion of Ukraine is causing a large-scale humanitarian crisis; vulnerable populations are most affected by the war; energy policy will include diversification of energy sources and ensuring uninterrupted access to them; food security will be high on the global agenda; increased competition for critical materials, equipment and goods will exacerbate global inequality; a new era of supply chain management will lead to higher prices for goods and services; the coexistence of different and sometimes competing technological standards will exacerbate the technological asymmetries of international economic development; uncertain impact on the financial system will be a trigger for financial shocks; rising defence spending will exacerbate budget deficits and require a revision of spending, which will affect social policy; cyberconflicts will be a challenge for digital platforms and governments; sanctions against Russia have affected the corporate sector; volatility will become a new norm in the world economy.

Reorganizing supply chains to increase their sustainability can mean managing a large number of relationships between manufacturers, distributors and consumers. There is a request for double sourcing. A new phenomenon is emerging – “friendshoring”, which is a process of cooperation, which is based on the commitment to work with countries that are firmly committed to democratic norms and values.

References

IPCC, (2022). Climate Change 2022: Mitigation of Climate Change. Retrieved from: https://www.ipcc.ch/report/sixth-assessment-report-working-group-3.

Ivashchenko, O. and Reznikova, N., (2018). Projections of Deglobalization in the Contemporary International Economic Relations in the Context of the Paradigm of Globalism. Bulletin of Mariupol State University. Series: Economics,15, pp. 98–106.

Krishnan, M., Nauclér, T., Pacthod, D., Pinner, D., Samandari, H., Smit, S. and Tai, H., (2021). Solving the net-zero equation: Nine requirements for a more orderly transition. Retrieved from: https://cutt.ly/QH8yheh.

Melnyk, T., Reznikova, N. and Ivashchenko, O., (2020). Problems of statistical study of "green economics" and green growth potentials in the sustainable development context. Baltic Journal of Economic Studies, 6(3), pp. 87–98.

Reznikova, N., (2021). Economic situation of developed countries, Kyiv: Agrarian Media Group.

Samandari, H., Pinner, D., Bowcott, H. and White, O., (2022). The net-zero transition in the wake of the war in Ukraine: A detour, a derailment, or a different path? Retrieved from: https://cutt.ly/VH8yVWI.

Shnyrkov, O., Reznikova, N., Zvarych, R. and Zvarych, I., (2019). Global circular e-chain in overcomingthe global waste. Procedia Environmental Science, Engineering and Management, 6(4), pp. 641–647.

UNOCHA, (2022). Regional Refugee Response Plan of Ukraine for 2022. Retrieved from: https://cutt.ly/tH8oiGu.

White, O., Buehler, K., Smit, S., Greenberg, E., Mysore, M., Jain , R., Hirt, M., Govindarajan, A. and Chewning, E., (2022). War in Ukraine: Twelve disruptions changing the world. Retrieved from: https://cutt.ly/0H8uMTx.

Zhong, D.and Zvarych, I., (2022). Benchmarking of the circular economy in the world. Economic scope, 179, pp. 164–171.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 63–87, DOI 10.59170/stattrans-2023-005
Received – 06.07.2022; accepted – 19.10.2022

A statistical study of climate change in Ukraine under martial law

Tetiana Kobylynska[1], Iryna Legan[2], Olena Motuzka[3]

ABSTRACT

The article concentrates on the development of theoretical and methodological foundations of statistical research in the field of national environmental and economic accounting, which forms the basis for the development of indicators of climate change under martial law and shapes the adaptation to these changes. Given Russia's military aggression against Ukraine, the national statistical system is faced with an urgent task of ensuring the continuity of the production of complete and high quality statistical information.

The issue of forming a methodology for the production of statistical information in the field of ecology, i.e. climate change statistics, remains unresolved. Outlining ways to address this issue is an important condition for the improvement of modern state statistics of Ukraine and the development of statistical theory and practice, as this largely determines the quality of the information-related support for environmental management at all levels.

The statistical estimation of the dynamics and structure of greenhouse gas emissions on the basis of the economic activity of enterprises was performed. It was pointed out that among the various related branch methods of obtaining information, the system of environmental and economic accounting is the most relevant and source of high quality data. One of the aspects of solving the complex problem of forming methodological bases for collecting statistical environmental information is the creation of a national system of indicators of climate change and preparing statistical informational products using administrative data sources.

The paper studies issues of producing ecological information relating to Ukraine according to statistical data. Moreover, the main problems which arise during the construction of national environmental accounts were characterised. Finally, the paper identified the key factors which influence to the largest extent the quality of statistical data and calculations, and which are necessary for the transformation and development of the statistical estimation of climate change under Russian military aggression.

Key words: climate change statistics, environmental and economic accounting, natural environment, statistical evaluation, ecological processes.

[1] Main Department of Statistics in Zhytomyr Region, Zhytomyr, Ukraine. E-mail: TVstat@i.ua. ORCID: https://orcid.org/0000-0001-8376-9656.

[2] State University "Zhytomyr Polytechnic", Ukraine. E-mail: legan.i@ukr.net. ORCID: 0000-0003-2933-4971.

[3] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: ommotuzka@nasoa.edu.ua. ORCID: https://orcid.org/0000-0001-9028-6994.

© Tetiana Kobylynska, Iryna Legan, Olena Motuzka. Article available under the CC BY-SA 4.0 licence

1. Introduction

So far, the war in Ukraine has led to large-scale destruction and mass deaths. Of course, there will be much more of them, as the military conflict continues. Any war causes enormous damage to the environment and the current situation is no exception. Although the fighting is still ongoing, it has already led to a significant deterioration in the environment and agricultural land, as the destruction of the chemical industry, oil refineries, fuel storage facilities and fossil fuel consumption during hostilities has led to huge greenhouse gas emissions (hereinafter - GHG) into the atmosphere of the planet. But in addition to emissions of hazardous substances, we should expect long-term losses due to the collapse of global efforts and measures to slow down the pace of climate change on the planet.

Due to the ongoing hostilities, it is now difficult to assess the damage done to the country's natural environment, so improving the organizational and methodological framework for statistical production of climate change indicators to assess existing losses is one of the main tasks today.

An integrated approach is needed to effectively produce statistical information on environmental processes during a military conflict. Today, the priority issue is to ensure the ability of the environmental and economic system of the country to develop in conditions of war, and this requires high-quality statistical information that will enable a full and reliable assessment of environmental losses.

As a result, there is an increasing need to modify the system of collecting and processing statistical information in order to assess climate change and to increase the requirements for existing statistics, including those regarding natural capital. Appropriate changes to metadata should be made using a proper set of rules. Faced with reality, we need to adapt to changes and their consequences, primarily through the rational management of water resources, reducing the use of other natural resources. But we will be able to control all this only when we have high-quality statistics to make informed decisions in a state of war and to develop an effective environmental policy. The next step should be the development of new technologies that will facilitate both adaptation to climate change and the implementation of effective mitigation measures.

Thus, our task is to conduct a qualitative statistical assessment in today's conditions, to develop and improve existing approaches to building a system of indicators of natural capital accounting, which, in turn, characterize climate change and adaptation to these changes. Improving the methodological and methodical foundations of statistical accounting and assessment of environmental indicators of climate change requires a clear and comprehensive statistical approach. Statistical analysis and evaluation can provide a complete picture of the functioning of such a complex phenomenon, help develop measures to reduce the negative impact on the environment, and, consequently, mitigate the negative effects of climate change on our planet.

The purpose of the research is to substantiate the basics of national statistical production and evaluation of information on climate change under martial law, the formation of environmental and economic governance mechanisms and their proper application.

2. Theoretical and methodological problems of statistical study of the national climate change statistics

Despite the extreme urgency of the problems of ecology and environmental protection, today both in Ukraine and abroad there is no unambiguous definition of the concept “natural capital”. This term belongs to the economist E.F. Schumacher, who introduced this concept in 1973 in the book “Small is Beautiful: A Study of Economics As If People Mattered”. There are many definitions of this term. A typical example was developed by the Natural Capital Coalition after a lengthy consultation process: natural capital is another term for a stock of renewable and non-renewable resources (e.g. plants, animals, air, water, soil, useful minerals) which, if combined, benefit people (SEEA).

One of the authoritative sources in this field is the Glossary of Statistical Terms of the Organization for Economic Co-operation and Development (OECD). It defines natural capital as natural assets which act as natural resources and environmental services for economic production (Glossary of Environment statistics). The UN Glossary of Environmental Statistics defines natural capital as natural assets in their role in providing natural resources and environmental services for economic production (Environment glossary).

Climate change is defined in Article 1 of the United Nations Framework Convention on Climate Change (UNFCCC). According to it, the term “climate change” means a change of climate which is attributed directly or indirectly to human activity that alters the composition of the global atmosphere and which is in addition to natural climate variability observed over comparable time periods. The term “climate system” is interpreted by the Convention as “the totality of the atmosphere, hydrosphere, biosphere and geosphere and their interactions”. In addition, the UNFCCC defines the term “adverse effects of climate change” as “changes in the physical environment or biota resulting from climate change which have significant deleterious effects on the composition, resilience or productivity of natural and managed ecosystems or on the operation of socio-economic systems or on human health and welfare” (United Nations). In turn, NASA experts believe that climate change will affect people around the world. Rising global temperatures are expected to raise sea levels and change both precipitation and other regional climatic conditions. Regional climate change will affect forests, crop yields, drinking water supplies, as well as human and animal health and almost all types of ecosystems. Deserts will expand to existing ridges, and the areas of national parks and national forests will be constantly changing (GlobalMaps; Global climate change).

Today, the world community is facing an acute problem of increasing the concentration of hazardous substances, especially GHG, in the surface layers of the atmosphere. Scientists around the world are conducting numerous studies to identify effective ways to implement alternative energy sources, calculate quantitative indicators of GHG emissions, develop various scenarios and algorithms to combat the negative impact of human economic activity on the natural environment (Khvesyk, 2010; Bun, 2004). The leading role in reducing the level of anthropogenic GHG emissions in national plans and programs is given to energy, energy conservation and energy efficiency, development of renewable energy sources, and, in the long run, replacing fossil fuels in power plants, in the process of combustion of which the significant amounts of carbon dioxide (CO2) are released.

The main sources of carbon dioxide emissions are extraction, processing, storage, transportation and consumption (combustion) of carbon-containing fuels (coal, oil and natural gas), internal fermentation of animals and livestock waste, agriculture (rice cultivation, which produces significant amounts of methane), the use of nitrogen fertilizers (with the subsequent formation of nitrogen oxides) (SEEA).

In recent decades, even before the onset of hostilities, climatic conditions in Ukraine have been changing significantly, causing increased risks to human health and life, natural ecosystems and sectors of the economy. Thus, warming leads to: sudden changes in weather; increase in frequency and intensity of dangerous natural hydrometeorological phenomena both in a warm season (showers, thunderstorms, squalls, hail, long hot periods), and in cold (heavy snowfalls, ice, difficult deposits); increasing the frequency and intensity of droughts and expanding the areas covered by them; reduction of river runoff in the south and south-east of Ukraine, increase in the intensity of floods on the rivers of the west of the country (Prykarpattia and Zakarpattia (Subcarpathia and Transcarpathia)), especially in the Dniester river basin; changes in the annual distribution of river runoff.

Combating climate change is a global challenge, which requires broad international cooperation, the consensus of which is reflected in a number of successive international agreements: UNFCCC (United Nations, 1996), the Kyoto Protocol (Kyoto Protocol) and the Paris Agreement (Verkhovna Rada of Ukraine, 2016). Given that the tools provided for in the first two agreements failed to achieve a significant reduction in GHG emissions, 195 countries signed the Paris Agreement, which entered into force on

November 04, 2016 (Verkhovna Rada of Ukraine, 2016). According to the goal of keeping global average temperature increase below 2°C of pre-industrial level, the Paris Agreement parties must balance sources of GHG emissions and removals in the second half of this century or in fact achieve net zero global GHG emissions by 2100.

Ukraine's goals are consistent with the global goals of the Paris Agreement, and joint international action will be crucial to achieving them. International cooperation will significantly reduce decarbonisation costs and create economic opportunities for people and enterprises, while reducing the risks and consequences of climate change. In addition, in order to improve the current climate policy, the Cabinet of Ministers of Ukraine by orders of 07.12.2016 № 932-r and of 06.12.2017 № 878-r approved the Concept for the Implementation of State Policy in the Field of Climate Change until 2030 and the Action Plan for its implementation (Cabinet of Ministers of Ukraine, 2017). This Concept defines the tasks in the following areas: increasing the institutional capacity of formation and implementation of state policy in the field of climate change; prevention of climate change by reducing anthropogenic emissions with increasing GHG absorption and ensuring a gradual transition to low-carbon development of the state; adapting to climate change, increasing resilience and reducing the risks associated with climate change.

Climate change is a threat and poses a huge challenge to society. Better understanding of all aspects of it through monitoring, modelling, analysis and reporting is crucial to identify best practices. Demands for environmental information in modern society are high, so government statistics bodies play an important role as major producer of high quality data.

In the long run, adaptation to climate change in Ukraine will have the same priority degree as climate change prevention, and will cover aspects such as proper policy planning and improvement in the context of expanding the knowledge base, scientifically based determination of needs and costs, implementation of innovative approaches and creating conditions for attracting foreign investment (Khvesyk, 2013; Stepanenko, 2016).

Scientific research often gives a picture of a particular issue at a particular point in time. Such research is certainly important for understanding this issue and ways to address it, but it does not always provide an understanding of how this phenomenon is developing in dynamics. This is where official statistics are of particular value, given that the development of long consecutive time series is the main competence of the national statistical system (NSS) of the country (Osaulenko, 2008; Osaulenko, 2020).

Official statistics involves the development of methods for adjusting data for comparability of changes over time, for example, by taking into account seasonality, price variation, temperature, etc. When such adjustments are made, the influence of a known source on data change is eliminated and only changes caused by unknown sources of impact and natural changes remain. This type of data processing is important in the study of climate change because by definition it is an issue that requires assessment over time, and therefore benefits from the use of long time series available in the NSS. The statistical system also has methodologies, techniques and procedures to develop statistics that are internationally comparable. This is done through a global institutional infrastructure that harmonizes definitions, classifications and data collection processes between countries. In addition, the main goal of the statistical system is to create quality statistics for the needs of society. Official systems of the countries of the world have developed their own methods of data collection to improve their reliability and timeliness, so statistical agencies can help, in particular, to satisfy requests for timely information on climate change (SEEA; Eurostat, 2021).

Scientific research emphasizes the importance of statistical assessment of socio-economic aspects of climate change and its impact for sustainable development (Verkhovna Rada of Ukraine, 2010). In this context, the NSS of Ukraine provides statistics for the creation of GHG emission inventories (cadastres) in order to provide a broader and more reasonable assessment of the socio-economic aspects of climate change. Recently, official statistics has increasingly focused on the issues of the environment and climate, giving preference to the development and improvement of methodologies and processes for compiling environmental accounts for the economic measurement of natural resources, energy, waste, water resources and emissions to the atmosphere, as well as environmental protection costs (State Statistics Service of Ukraine, 2018, 2019, 2020). However, despite the many competencies of the statistical system for measuring climate change, existing official statistics does not always use its potential to assess climate change.

The UN Statistical Commission has adopted a system of environmental and economic accounting (hereinafter – SEEA) at the global level. This new standard provides important support for the development of statistics on climate change by strengthening work on environmental and economic accounting. The SEEA contains internationally agreed concepts, definitions, classifications, accounting rules and tables for the production of comparable statistic data on the environment and its interrelation to the economy (SEEA).

Considering climate change statistics as a tool for influencing decision-making in this area, we will determine how to improve statistical accounting relative to climate and climate change. A change from a statistical standpoint is the difference between two observations (usually an absolute or relative difference between the values of indicators at different points in time). Qualitative measurement of change requires consistent and comparable data collected over a long period of time. In general, existing official statistics are not developed to analyse climate change. Therefore, the NSS needs to be transformed to better respond to the need for data and assessment regarding climate change. This will certainly require new ways of disseminating official statistics, collecting data, and also may lead to organizational changes in the statistical system.

The statistical system of Ukraine currently offers a lot of data on different areas of socio-economic and environmental life of the country, but often does not pay much attention to preventing duplication of data and measuring the relationship between these areas of statistics. Defining and modelling interactions is the work of economists, scientists, analysts, experts of state ministries, etc., who are outside the statistical system (Mishchenko, 2007; Lavryk et al., Dorohuntsov, 2007). The main elements of these definitions are: human-caused climate change (not related to natural causes); climate as a system consisting of several components (atmosphere, hydrosphere, etc.), their interaction; taking into account the fact that climate change ultimately affects both the natural environment and the socio-economic aspects of human life.

The main cause of climate change due to human activities, which is the basis of development and technological progress, is GHG emissions into the atmospheric air. Thus, Article 2 of the UNFCCC, outlining the main purpose of the Convention, clearly states this reason: "The ultimate objective of this Convention and all related legal instruments ...is to achieve, in accordance with the relevant provisions of the Convention, stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system. Such a level should be achieved within a time-frame sufficient to allow ecosystems to adapt naturally to climate change, to ensure that food production is not threatened and to enable economic development to proceed in a sustainable manner" (United Nations, 1996).

How can a statistical system, which has many particular qualities of accounting and data collection that are important in the context of climate change (Cabinet of Ministers of Ukraine, 2001), help to measure it? Official statistics adhere to international principles and standards in this area, which are formulated as a result of numerous discussions on climate change and are characterized by a number of important features. First of all, official statistics mainly measure the activities of legal entities (enterprises), individuals and households. Monitoring the state of the natural environment, climate changes and adaptation to these changes is usually carried out outside the statistical system. However, when data on the environment and climate are related to sectors of the economy and households that affect the environment, these data may fall into the scope of official statistics. The latter also often contains links to the geographical region. In some countries, statistics receive exact spatial coordinates using geo-links. This practice needs to be actively developed further, as the spatial dimension is too important in the context of climate change.

The development of new statistics to support the analysis of climate change within the statistical system, such as monitoring progress towards the ecological structure of consumption and production, will be a challenge for the NSS. As climate change statistics cover many sectors, it is necessary to outline and study the field of statistics related to these changes, as well as to create a structure in the statistical bodies which will be responsible for this area.

Intensification of extreme weather and climate events will cause an increase in negative impact on ecosystems, economic sectors, health and welfare of population. As a result, actions to adapt to climate change must be paramount and tailored to specific circumstances in different areas of human activity. It is now clear that the issues of the quality of our habitat cannot be approached only in any one of aspects, but must be comprehensive. The current model of environmental development is futureless for humanity in terms of security, as it exacerbates the environmental and economic problems of various spheres of activity. The ecological condition of the environment continues to deteriorate, and therefore the preservation of nature, in particular the reduction of emissions of hazardous substances from all sources of pollution, should become a priority for every state and society.

The formation and implementation of the country's socio-economic development strategy and state environmental policy must be coordinated and interconnected, as the health of the nation, environmental and social growth are closely and inextricably linked. The modern national economy is impossible without taking into account the environmental component. Over the past year, Ukraine has approved a number of important measures in the field of environmental policy, including the Action Plan for the Implementation of the Concept of the State Climate Change Policy; the Low Carbon Development Strategy of Ukraine until 2050 was worked out, and work on legislation to monitor greenhouse gas emissions has started (Ministry of Energy and Environmental Protection of Ukraine, 2018). High-quality statistical data will provide the basis for effective decision-making regarding climate action. State statistics bodies play a key role in national efforts to promote the development of official statistics for climate assessment and national reporting. Thus, the main task of this development is the scientific substantiation of the following statistical estimates:

- GHG emissions;
- drivers of climate change (emissions in the industrial sector, energy, agriculture, etc.);
- social, economic and environmental after-effects of climate change;
- adaptation to the obtained consequences.

The outlined direction of the research is related to the 13th Sustainable Development Goal (hereinafter referred to as the SDG) on climate change mitigation. To do this at the national level it is necessary to develop a methodology for compiling the main list of indicators related to climate change, approved by the Conference of European Statisticians in 2017 (Supporting countries to achieve the SDGs). The developed national list of indicators of SDG 13 does not fully correspond to the global indicators of SDG (G), and therefore, the search for available sources of information is still ongoing (Figure 1).

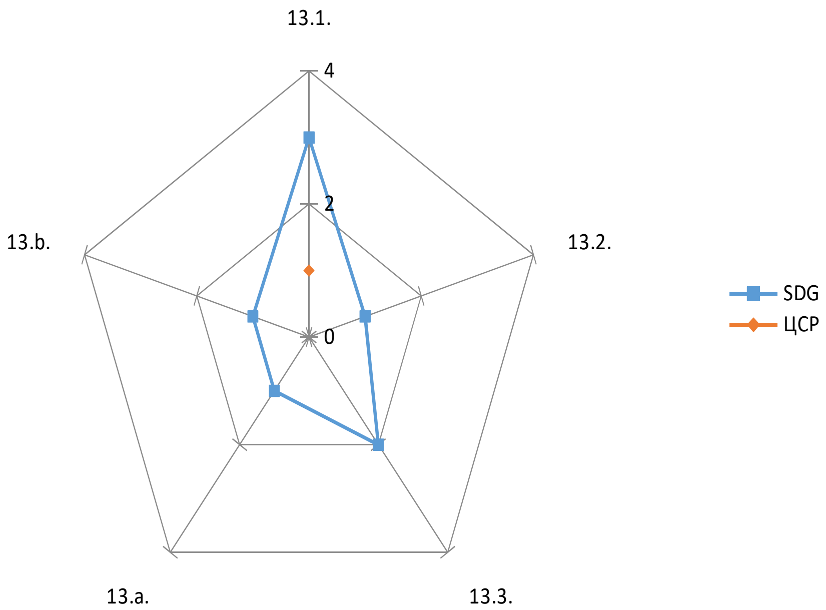


Figure 1. Comparison of national and global SDG

Source: developed by the author.

As climate change affects the natural environment and society directly through complex interactions and cause-effect relationship, the conceptual basis for the production of climate change statistics should cover the widest possible range of data and be as comprehensive as possible. Climate change, on the one hand, is due to, on the other hand, affects a wide range of anthropogenic activities. In fact, only a few human activities are not related to climate change in any way or on some aspects do not contribute to positive climate change.

As a result, the scope of statistics related to climate change is:

1) environmental, social and economic data measuring the anthropogenic causes of climate change (Figure 2);
2) the impact of climate change on anthropogenic and natural systems;
3) society's efforts to prevent effects;
4) society's efforts to adapt to these effects.

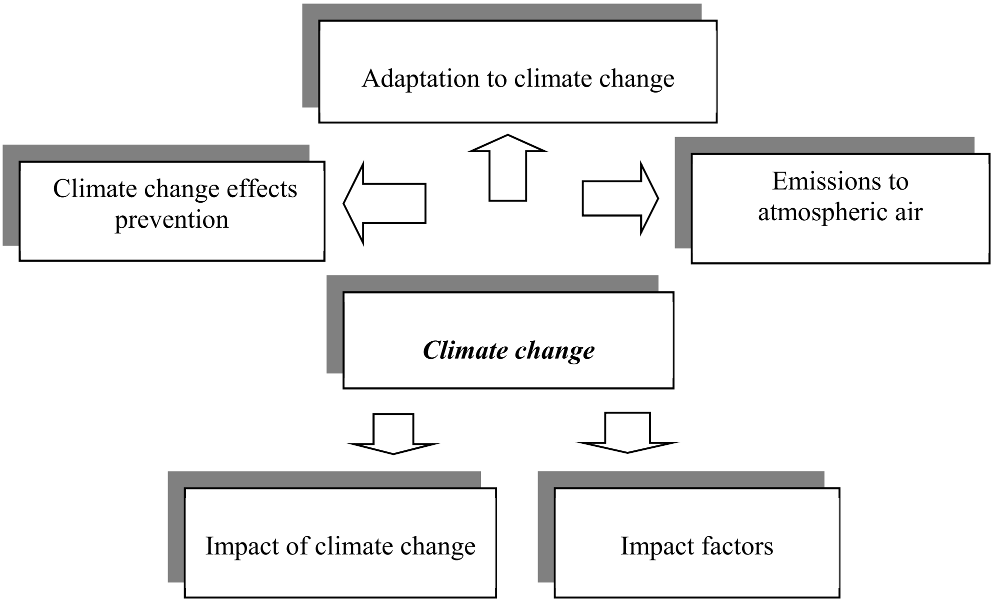


Figure 2. Coverage of statistics related to climate change

Source: developed by the author.

3. National approaches to statistical assessment of climate change indicators under martial law

A key element in building national climate change statistics is to define basic responsibilities for quality assurance and the availability of climate change statistics, namely, the development of statistics for GHG inventory and building relationships with key users and producers of climate information, including users and producers within the NSS itself.

Exploring the capabilities of the statistical system will help to identify the subject areas related to climate change, and will provide coverage of all relevant areas in the work of the NSS. In current conditions, the following approaches to statistical accounting can be used:

- Drivers, Pressures, State, Impact and Response model of intervention, (hereinafter DPSIR);
- United Nations Environmental and Economic Accounting System (SEEA);
- United Nations Framework for the Development of Environmental Statistics (UNFDES);
- Approach to natural capital accounting;
- Impact, mitigation and adaptation.

The purpose of creating national approaches to statistical accounting is to identify the area which is best suited to structuring climate change statistics. It should be recognized that each area has its own purpose and benefits, and the areas are not independent of each other. Thus, the approach to the accounting of natural capital is closely related to the system of environmental and economic accounting, which is a statistical implementation of many, but not all areas (Osaulenko, 2020). SEEA, in turn, covers the concept of natural capital. At the same time, the structures of impact, mitigation and adaptation are identical, the main difference between them is the direct use of terminology of climate change policy.

It should be emphasized that an important feature of the SEEA is its approval by the UN Statistical Commission, which puts this system in the same category as the widely used system of national accounts (hereinafter – SNA). The SNA is an international statistical standard and system, which uses the NSS to calculate gross domestic product and other macroeconomic variables, with SEEA linking environmental statistics to economic statistics. This status makes the SEEA an attractive basis for use in measuring climate change statistics, as the NSS of European countries have already agreed to use it in the development of environmental reports (Osaulenko, 2020).

Climate change statistics can be studied and structured according to the so-called model of the DPSIR. This model is used, for example, by the European Environment Agency to structure the relationship between the environment and socio-economic activities. The DPSIR is based on a systematic approach to the assessment of environmental phenomena. Although the model seems quite simple, it may also involve a more complex accounting system. The organization of climate change statistics, based on the DPSIR model, will help to structure information from different areas, which will allow to analyse various aspects of climate change, including a wide range of socio-economic events, GHG emission sources, actions relating to adaptation and mitigation of climate change.

In turn, the impact of climate change can be seen as direct or indirect. In the first case, the impact is analysed from the standpoint of natural phenomena, for which data on water, air, land resources, biodiversity and ecosystems are relevant. The second should take into account the socio-economic impact of physical change, such as the impact on agriculture, forestry, human health and the economy. In this case, the role of the NSS is more important, as environmental data are often provided not by it, but by environmental institutions and organizations.

The United Nations Environmental and Economic Accounting System (SEEA) is a multi-purpose system for measuring the share of the environment in the economy and the impact of the economy, human activities on the environment. In practice, the SEEA covers an integrated set of basic statistics which can be regrouped into various indicators useful for climate change analysis. This system is flexible in the sense that its implementation can be adapted to policy priorities and needs, while providing a common framework between countries and coherence with the concepts, structures, rules and principles of the SNA. Although the SEEA is not primarily intended for climate change analysis, it contains a number of components related to its various aspects, namely:

1) accounting for physical flows of materials and energy;
2) accounting for stocks of environmental assets and changes in them;
3) accounts of economic activities and operations related to the environment.

With regard to climate change, the SEEA accounts of emissions to atmospheric air contain data on the origin and purpose of these emissions, including GHGs. Thus, energy flow accounts provide relevant information on the supply and use of energy (both renewable and non-renewable). Water flow accounts are especially important for understanding the impact of climate change on changes in water availability and use. Solid waste stream accounts contain data on waste sources, their incineration and methane emissions in landfills, which is one of the powerful sources of GHG. Currently these modules are still under development and most countries are unable to provide full statistics and time series useful for environmental policy purposes.

The introduction of SEEA into the national statistical system will help to create a basis for the collection and analysis of indicators related to climate change, as those which characterize, among others:

- energy use and GHG emissions per unit of GDP by industry;
- actions regarding GHG emissions calculated on the basis of consumption;
- actions to separate the economy and the environment.

The main SEEA accounts measure the size of environmental assets (ecosystems, land and natural resources) and the annual changes of these reserves. Asset accounting is important for quantification of the impact of climate change on various environmental assets. Accounts are valued both physically and (where possible) in value terms and are compiled for energy, land, soil, timber and water resources that may change as a result of climate change. For example, as precipitation regimes vary, the availability of water resources and their geographical distribution will change.

SEEA accounts measure activities and operations the main purpose of which are:

1) prevention, reduction and elimination of pollution and other forms of environmental degradation;
2) conservation and maintenance of natural resources, and therefore protection against depletion;
3) influence on the behaviour of producers and consumers in relation to the environment.

The SEEA provides recommendations on measuring environmental taxes, subsidies, permits and licenses related to energy production and use, GHG emissions and ecological innovations. These data are the basis for analysing the relationship between GHG emissions, energy use and emissions permits, as well as for monitoring GHG emissions trading. SEEA limits the scope of activities that are considered environmental, environmental protection and resource management activities.

National environmental and economic statistical accounting is represented by two SEEA accounts. They are: 1) the account of emissions into the atmosphere; 2) auxiliary (satellite) cost account for the protection of the national natural environment.

First of all, it should be noted that the Decision No 1578/2007/EC of the European Parliament and of the Council of 11 December 2007 on the Community Statistical Programme 2008 to 2012 determines the need to produce high quality statistical information and accounts on the environment and to supplement existing indicators with data, including environmental and social aspects, in order to implement a coherent and comprehensive environmental policy (State Statistics Service of Ukraine, 2019, 2020). That is why environment and economic accounting (hereinafter - EEA) has been introduced as a statistical tool in order to monitor the pressure exerted by the economy on national natural environment, and to study ways to reduce it.

Taking into account the existing international and European experience, the organization of creation, analysis and dissemination of environmental and economic accounting is entrusted to the State Statistics Service of Ukraine. The purpose of the introduction of the account of emissions into the atmosphere is the formation of statistical information on the amount of pollutants and GHG entering the air from the production activities of residents within the economic territory of the country to provide data for environmental and economic analysis.

The introduction and maintenance of a national EEA ensures complete and high-quality accounting of the condition and use of the national natural environment. The environmental account of emissions into the atmosphere registers the flows of gaseous and dispersed materials, describing these emissions by their sources and types of economic activity in terms of both production and consumption. Atmospheric emissions include emissions of GHGs and air pollutants, taking into account the principle of residency. This principle can be formulated as follows: an economic unit is considered to be a resident unit of a country if the centre of its economic interest is located in the economic territory of that country, i.e. if this unit carries out economic activity in this territory for a long period (one or more years).

It should be noted that the first experimental calculations in the official statistics of Ukraine were conducted in 2015 on the basis of data for 2013 in compliance with international standards for accounting and dissemination of data on emissions of pollutants and GHG into the air in a way compatible with SNA. The development of this account provided data on emissions of pollutants and GHG in the atmosphere air from enterprises and households activities.

The basic standard for creating an account is the International Standard of the System of Environmental-Economic Accounting 2012 - Central Framework, (SEEA) Regulation (EU) No 691/2011 of the European Parliament and of the Council of 6 July 2011 On European Environmental Economic Accounts (European Commission, 2019). The legal basis for the creation of the account on emissions is the Law of Ukraine "On State Statistics", the annual plan of state statistical surveys, which is approved by the order of the Cabinet of Ministers of Ukraine, and the procedure defined by metadata, which are approved by the orders of the State Statistics Service (Cabinet of Ministers of Ukraine, 2017). This account reflects data on 11 types of pollutants and GHGs released into the atmosphere by stationary and mobile sources of pollution by type of economic activity (KVED-2010 (Classification of economic activities). In addition to data on emissions, the account reflects economic indicators relative to output, gross added value and the number of employed population according to KVED-2010. The data sources for compiling the above account are:

- the State Statistics Service data "Protection of atmospheric air" on emissions of pollutants and solid suspended particles into the atmosphere from the production activities of stationary sources;
- the SNA data on final consumption expenditures of households, output, gross added value according to KVED-2010;
- the State Statistics Service data "Labor Force Survey" on the number of employed population according to KVED-2010.

For the purpose of maintaining the quality of statistical information, the environmental account for atmospheric air protection (hereinafter – EAAAP) is compared with national emission inventories. The data provided on a disaggregated basis may correspond to emission sources by type of economic activity and supplement national emission inventories (State Statistics Service of Ukraine, 2020).

In accordance with international and European requirements, information on emissions of pollutants into the atmospheric air is presented in terms of types of economic activities for the following substances:

- greenhouse gas emissions (CO2, N2O, CH4, HFC, PFC and SF6);
- emissions of pollutants (NOx, CO, NMVOC, SO2, NH3);
- emissions of solid suspended particles (less than 10 μm and 2.5 μm).

In national inventories, GHG air emissions and pollutants are broken down by process and source, classified by technologies. Unlike environmental accounts, national emission inventories are largely based on the principle of territoriality, i.e. they cover emissions from the country's geographical area. It should be noted that the national accounting of pollutant emissions is conducted in terms of 109 economic activities

(KVED-2010) in accordance with Regulation (EU) № 691/201 of the European Parliament and the Council (Compilation Guide, 2013).

The generally recognized at the international level and adapted for Ukraine features of the calculations are the following:

- statistical data on air emissions are recorded at the time of emissions;
- emissions from cultivated plants, soils and forests are not taken into account in the calculations;
- landfill emissions are taken into account only in part of the actual entry of pollutants into the air (captured gases are not taken into account);
- emissions of non-economic origin, which occurred, for example, in wetlands, due to forest fires, volcanic eruptions, etc., are not taken into account;
- emissions from livestock due to fermentation (methane production), as well as from manure processing are taken into account if they relate to the economic activities of residents;
- repeated emissions due to natural processes in the atmosphere are not taken into account;
- CO2 emissions from biomass are included as a separate item.

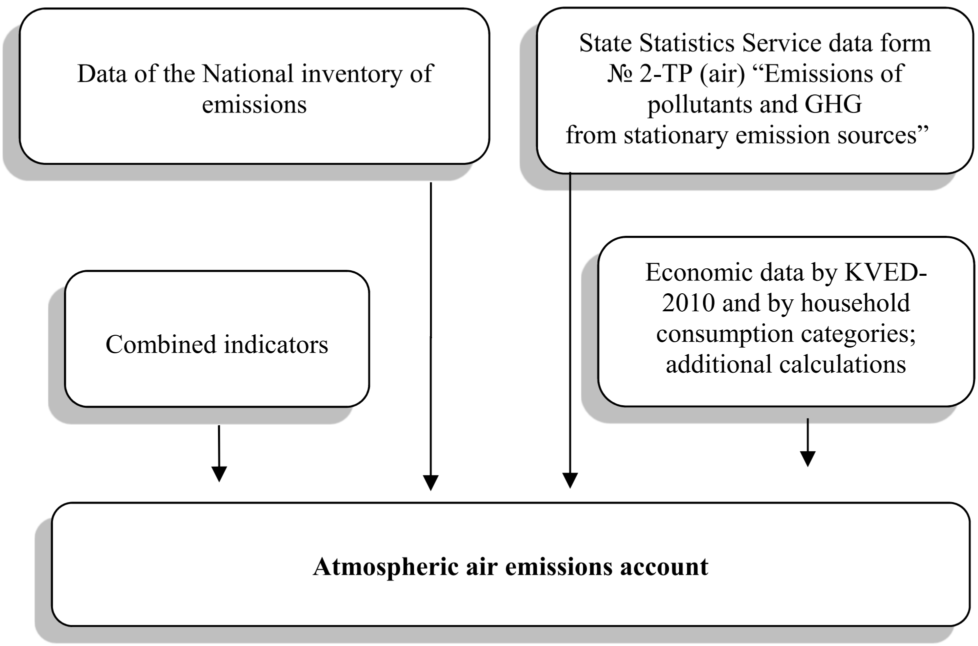


Figure 3. Chart for collecting national information to compile an account of emissions into the atmospheric

Source: developed by the author.

It should be noted that the national results on air protection differ from the results defined in the emission inventories. These differences are due to differences between the principles of presidency and territoriality and the definition / coverage of national results in the emission inventory. For example, emissions from international shipment are excluded from the results of inventories consistent with the structure and principles of accounting for SEEA (United Nations, (2014).

In the context of the quality of statistical information, given the national differences in the list of pollutant emissions, it is necessary to emphasize the differences between the national system of indicators and the accounts of air pollutant emissions collected by the United Nations Economic Commission for Europe. For this purpose, within the framework of the air emissions account, an appropriate section entitled "Combined Indicators" has been created (see Figure 3). In addition, it should be noted that in all accounting systems, emissions are usually determined by estimation (i.e. calculation) rather than measurement. At the same time, all accounting systems consider anthropogenic, i.e. human-caused, emissions into the atmosphere, and, as a rule, do not take into account emissions from natural sources.

Concepts for coverage of emission sources are defined: national inventories cover GHG emissions and their absorption within the national territory and coastal areas over which the country has jurisdiction. In this case, emissions from the use of fuel by road transport are included in the emissions of the country in which the fuel is sold, not the country in which the vehicle is used, as fuel sales statistics are widely available and usually much more accurate.

In turn, the UNFCCC inventory covers six types of GHGs, for which certain quantitative targets have been set: CO2; N2O; CH4; HFC; PFC; SF6. The first three are individual substances, and the rest are groups of substances called F-gases (fluorine-containing gases). In 2015, another substance was added which is not currently included in the emissions accounts: nitrogen trifluoride (NF3).

Emissions of all GHGs can be summed up taking into account their global warming potential, and for each type of GHG emissions are recalculated using specific global warming coefficients that reflect the gas potential in CO2 equivalents. CO2 emissions from biomass are recorded as a "Reference" item in the UNFCCC inventory. It should be noted that the UN inventories for information purposes also cover emissions of NOx, CO, NMVOC and SO2, but this information is not used to monitor the achievement of environmental policy objectives. In order to evaluate it, the relevant registration of these substances is carried out in the inventory of the Convention on Long-range Transboundary Air Pollution (United Nations Statistics Division).

According to the data of the State Statistics Service of Ukraine, during 2020, 109.1 million tons of carbon dioxide were released into the country's air due to the production activities of enterprises, which is 12.2 million (or 12.2%) less than in 2019 (Figure 4), according to (State Statistics Service of Ukraine, 2021).

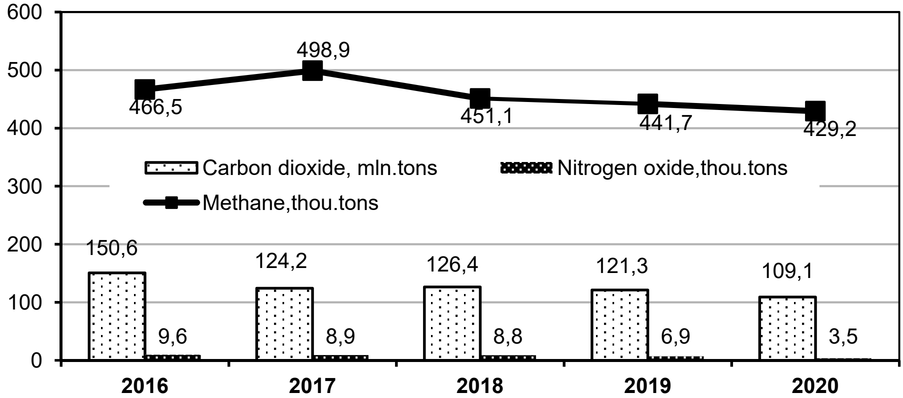


Figure 4. Dynamics of emissions of the main types of greenhouse gases into the atmospheric air of Ukraine

Source: built by the author according to (State Statistics Service of Ukraine, 2021).

CO2 emissions from the production activities of stationary sources of pollution occur in the following main sectors of the economy (sections KVED-2010):

- supply of electricity, gas, steam and conditioned air;
- process industry;
- mining industry and quarrying;
- agriculture

The largest volumes of CO2 were emitted by enterprises of section D “Supply of electricity, gas, steam and conditioned air” – 56.2 million tons, which is 52% of all emissions volume (Figure 5, according to (State Statistics Service of Ukraine, 2021).

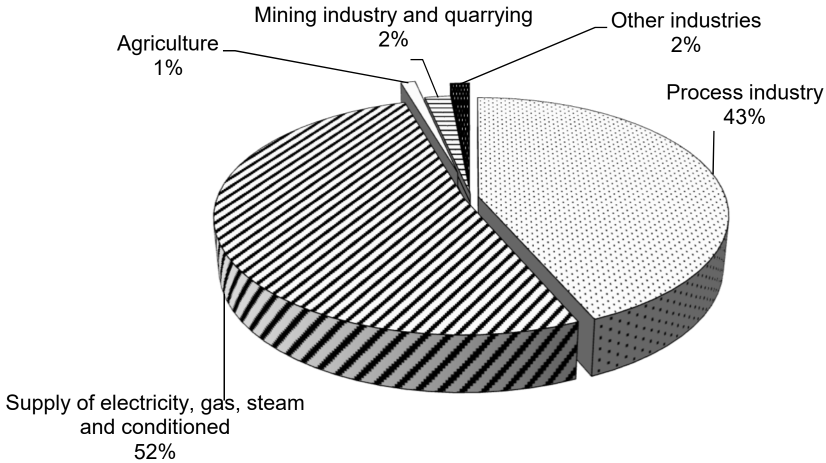


Figure 5. Structure of carbon dioxide emissions by main sections of KVED-2010, Ukraine, 2020, million tons

Source: built by the author according to (State Statistics Service of Ukraine, 2021).

The largest amount of 9470.2 thousand tons of CO2 emissions (8.6% of total emissions in the country) came into the air in 2020 from the metallurgical plant "Arcelor Mittal Kryvyi Rih", as well as from the Zaporiuzhzhia metallurgical plant "Zaporizhstal" - 6631.8 thousand tons (respectively 6.1%). It should also be noted that the five largest industrial enterprises, which are among the 10 largest polluters in Ukraine, are located in Donbas area, where active, fierce fighting is currently taking place. They are:

1) Kurakhiv thermal power plant;
2) Mariupol Metallurgical Plant named after Ilyich;
3) Azovstal metallurgical plant;
4) Slavic thermal power plant "Donbasenergo";
5) Vuglegirsk Centrenergo Coal Mine Thermal Power Plant.

During 2020, the above-mentioned enterprises generated a total of 19,361.5 thousand tons of CO2 into the country's air, which is 17.7% of all national emissions.

As a result of the study, it was found that the largest GHG emissions, namely, of carbon dioxide, per capita were generated in the Zaporizhzhia region – 7740.0 kg, Ivano-Frankivsk – 7479.9 kg, Dnepropetrovsk – 6480.7 kg (Figure 6).

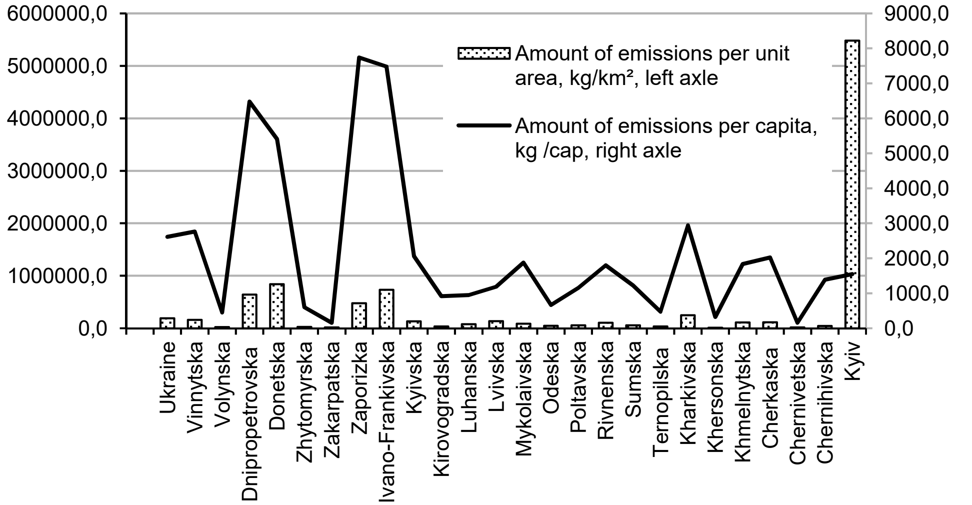


Figure 6. GHG emissions by regions (oblasts) per unit area and per capita, 2020

Source: built by the authors according to (State Statistics Service of Ukraine, 2021).

It should be noted that the national SEEA is represented by another account: that is the account of the natural environment protection costs, which is formed in accordance with the statistical classification of environmental protection activities and costs (CEPA 2000), which identifies nine environmental areas for its description:

- protection of atmospheric air and problems of climate change;
- waste water treatment;
- waste treatment;
- protection and rehabilitation of soil, groundwater and surface water;
- reduction of noise and vibration exposure (except for labour protection measures);
- conservation of biodiversity and habitat;
- radiation safety (except for measures to prevent accidents and catastrophes);
- research works in the field of environmental protection;
- other areas of environmental protection activities.

Typically, environmental protection costs are described by the structure of environmental protection costs according to the classification of environmental measures by type of costs: capital investment, including overhaul costs, and running costs. However, these indicators do not determine the place and role of the natural environment in the economic life of the country. This goal can be achieved only with the help of cost indicators, the methods and methodology of which will be comparable to those used for valuation of other sectors of the economy. Such indicators form the basis of the natural environment protection cost account, providing a detailed study of the environment as an economic component of the SNA.

In national accounts, environmental activities are seen as a system of consistent and interconnected accounts that characterize the operations regarding the production of environmental services, the creation, distribution and use of income, the accumulation of capital and the acquisition of financial instruments. Institutional residents are grouped into the sectors of non-financial corporations, financial corporations, general government bodies, private households, non-profit institutions which serve households, with their further breakdown by subsector. In the framework of the implementation of strategic directions for further improvement of environmental statistics and in accordance with international standards, in particular the EU Regulation №691 / 2011 on European environmental and economic accounts module "Environmental Protection Expenditure Accounts" (EPEA), countries are encouraged to introduce national auxiliary (satellite) accounts of costs for the protection of natural environment.

Data in the EPEA account regarding the protection of the natural environment costs are presented by the following institutional sectors of the economy: state government, corporations, private households and other world. At the same time, non-profit organizations which serve private households are not separately identified and are included in general government bodies. The corporation sector is divided between corporations as the main producers of environmental protection services and producers of the market services of the natural environment protection as "ancillary activities" and "other corporations", the production activities of households are included in corporations as market producers.

To facilitate monitoring and reporting on climate change, a working group under the auspices of the Conference of European Statisticians (CES) of the United Nations Economic Commission for Europe has developed a set of relevant indicators [30]. The creation of such a set for measuring and monitoring changes over time will be based on the core competence of national statistical offices to ensure the high quality of a consistent dynamic series.

The purpose of the use of the global set of statistics and indicators of climate change by UN is to create a framework with relevant indicators that will guide countries in preparing their own sets, especially for countries with less developed statistical systems. The global set is designed to meet the needs of all countries and contains an exhaustive list of metadata-related indicators (including definitions, input variables, aggregations, measurement categories and data links). In total, the CES set covers five main areas; these are 44 indicators that affect climate change and are distributed as follows:

- emissions of pollutants: 9 indicators;
- drivers: 9 indicators;
- impact of climate change: 13 indicators;

- climate change mitigation: 8 indicators;
- adaptation to climate change: 5 indicators.

The CES set of indicators provides a basis for the development of national sets of indicators related to climate change, which aim to show the general issues of climate change, the most acute policy issues and help to meet the immediate information needs. It contains indicators that cover the above factors and offers relevant indicators which help to interpret the core set in national and global contexts.

4. Conclusions

Thus, climate change research is a complex and complicated process, in which qualitative statistical data on emissions of hazardous substances, including greenhouse gases, as well as other inter-sectoral data and statistical calculations, are absolutely important. Improving and modifying the collection, processing of data and additional calculations is a requirement of today and should provide up-to-date and timely information for environmental policy decisions.

On the whole, in order to further achieve the effectiveness of environmental impact assessment, the implementation of the following methods and measures deserves attention:

1) Engaging the necessary experts through partnerships.
2) Familiarizing the staff with GHG emission inventory methodologies and explaining the need for appropriate cooperation between inventory compilers and official statisticians.
3) Expanding knowledge, developing methodologies and tools for obtaining and using geographic data throughout the whole statistical system, including through the organization of appropriate cooperation with specialists in geospatial information.
4) Ensuring the effective international exchange of knowledge and skills between NSSs.

Reference

Bun, R. A., Husti, M. I., Dachuk, V. S. et al., (2004). Informatsiini tekhnolohii inventaryzatsii parnykovykh haziv ta prohnozuvannia vuhletsevoho balansu Ukrainy [Information technologies for inventory of greenhouse gases and forecasting of carbon balance in Ukraine]. The Ukrainian Academy of Printing, Lviv [in Ukrainian].

Cabinet of Ministers of Ukraine, (2001). Pro zatverdzhennia Poriadku vedennia derzhavnoho obliku v haluzi okhorony atmosfernoho povitria [Approval of the

Procedure for the official accounting in the field of protection of ambient air]. URL: https://zakon.rada.gov.ua/laws/show/1655-2001- %D0%BF#Text [in Ukrainian].

Cabinet of Ministers of Ukraine, (2017). Pro zatverdzhennia planu zakhodiv shchodo vykonannia Kontseptsii realizatsii derzhavnoi polityky u sferi zminy klimatu na period do 2030 r. [Approval of the action plan on the implementation of the Guidelines on realization of the policy on climate change till 2030] [in Ukrainian]. URL: https://zakon.rada.gov.ua/laws/show/878-2017-%D1%80#Text.

Dorohuntsov, S. I., (2007). Ekoseredovyshche i suchasnist [The environment and the modernity]. Ekoseredovyshche ta yevrointehratsiini protsesy [The environment and Eurointegration processes], in 8 Vols. Kyiv [in Ukrainian].

Environment glossary. URL: https://unstats.un.org/unsd/environmentgl/

European Commission, (2019). Communications from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. URLhttps://eur-lex.europa.eu/resource.html?uri=cellar:b828d165-1c22-11ea-8c1f-01aa75ed71a1.0002.02/DOC_1&format=PDF

Eurostat, (2013). Compilation Guide (2013) for Eurostat's Air Emissions Accounts (AEA) revision of part B of the Manual for Air Emissions Accounts (AEA) – 2009 edition, version: April 2013.

Eurostat, (2021). Guide to statistics in European Commission development cooperation. Vol. 4: Environment and climate change. URL: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-21-015

Global Climate Change. URL: https://climate.nasa.gov//

Global Maps. URL: https://earthobservatory.nasa.gov/global-maps

Glossary of Environment Statistics. URL: https://stats.oecd.org/glossary/detail.asp?ID=1730

Khvesyk, M. A., (2010). Formuvannia instytutsionalnoho seredovyshcha pidpryiemnytskoi diialnosti u sferi pryrodokorystuvannia: investytsiino-innovatsiinyi aspekt [Building up the institutional environment for business activities in the sphere of nature use]. National Academy of Sciences of Ukraine, Council on Studies of Production Forces of Ukraine. Lutsk: Tverdynia [in Ukrainian].

Khvesyk, M. A., Stepanenko, A. V., Obykhod, H. O., (2013). Innovatsiino-investytsiina i tekhnolohichna bezpeka transformatsii rehionalnykh ekonomichnykh system

[Transformation of regional economic systems: innovation, investment and technology aspects of security]. Institute of Environmental Economics and Sustainable Development of the National Academy of Sciences of Ukraine. Kyiv: Naukovadumka [in Ukrainian].

Kyoto Protocol to United Nations Framework Convention on Climate Change [in Ukrainian]. URL: https://zakon.rada.gov.ua/laws/show/995_801#Text

Lavryk, V. I. Boholiubov, V. M., Polietaiev, L. M., Yurasov, S. M., (2010). Modeliuvannia i prohnozuvannia stanu dovkillia [Modelling and forecasting of the environmental performance]. Kyiv: Academia [in Ukrainian].

Ministry of Energy and Environmental Protection of Ukraine, (2018). Stratehiia nyzkovuhletsevoho rozvytku Ukrainy do 2050 roku [The strategy for low carbon development of Ukraine till 2050]. URL: https://mepr.gov.ua/files/docs/Proekt/LEDS_ua_last.pdf

Mishchenko, V. S., (2007). Ekonomichni priorytety rozvytku y osvoiennia mineralno-syrovynnoi bazy Ukrainy [Economic priorities of the development and utilization of mineral resources base of Ukraine]. Kyiv: Naukovadumka [in Ukrainian].

Natsionalnyi kadastr antropohennykh vykydiv izdzherel ta absorbtsii pohlynachamyparnykovykh haziv v Ukrainiza 1990–2018 rr. [National Cadaster of Anthropogenic Emissions from Sources and Absorption of Greenhouse Gases by Absorbers in 1990–2018] [in Ukrainian] URL: http://surl.li/cgfif

Osaulenko, O. H., (2008). Natsionalna statystychna systema: stratehichne planuvannia, metodolohiia ta orhanizatsiia [The national statistics system: strategic planning, methodology and organization]. Kyiv [in Ukrainian].

Osaulenko, O. H., Kobylynska, T. V., (2020). Statystyka zminy klimatu: yevropeiskyi dosvid ta natsionalna otsinka [Statistics of climate change: European practices and national assessment]. Kyiv: Avgust-Treid Ltd [in Ukrainian].

Plan of official statistical observations, approved by a directive of the Cabinet of Ministers of Ukraine, which procedure is based on the metadata approved by orders of the State Statistics Service of Ukraine). URL: http://www.ukrstat.gov.ua/

State Statistics Service of Ukraine, (2018). Metodolohichni polozhennia z skladannia dopomizhnoho (satelitnoho) rakhunku vytrat na okhoronu navkolyshnoho pryrodnoho seredovyshcha [Methodological guidelines on the compilation of the support (satellite) account on the expenditure on environmental protection] [in Ukrainian]. URL: http://www.ukrstat.gov.ua/

State Statistics Service of Ukraine, (2019). Metodolohichni polozhennia z orhanizatsii derzhavnoho statystychnoho sposterezhennia shchodo vykydiv zabrudniuiuchykh rechovyn ta parnykovykh haziv v atmosferne povitria vid statsionarnykh dzherel zabrudnennia [Methodological guidelines on organization of the official statistical observation on the emission of pollutants and greenhouse gases to the ambient air from stationary pollution sources] [in Ukrainian]. URL: http://www.ukrstat.gov.ua/norm_doc/2019/454/454.pdf

State Statistics Service of Ukraine, (2020). Metodolohichni polozhennia zi skladannia rakhunku vykydiv v atmosferu [Methodological guidelines on the compilation of the account on air emissions] [in Ukrainian]. URL: http://www.ukrstat.gov.ua/norm_doc/2020/268/268.pdf

State Statistics Service of Ukraine, (2021). Dovkillia Ukrainyza 2020 rik [The Environment of Ukraine in 2020]. URL: http://www.ukrstat.gov.ua/druk/publicat/kat_u/2021/zb/11/Dovk_20.pdf

Stepanenko, A. V., (2016). Ekolohichna modernizatsiia v systemi okhorony atmosfernoho povitria v rehionakh Ukrainy [Environmental modernization as part of the ambient air protection in Ukrainian regions]. Kyiv [in Ukrainian].

Supporting countries to achieve the SDGs. URL: https://unece.org/unece-and-sdgs-2

System of Environmental-Economic Accounting. URL: https://seea.un.org/content/frequently-asked-questions

United Nations Framework Convention on Climate Change. International agreement ratified by Ukraine 29.10.1996. URL: https://zakon.rada.gov.ua/laws/show/995_044#Text [in Ukrainian].

United Nations Statistics Division. Manual on the Basic Sets of Environmental Statistics. URL: https://unstats.un.org/unsd/envstats/fdes/manual_bses.cshtml

United Nations, (1979). Convention on long-range transboundary air pollution. URL: https://treaties.un.org/Pages/ViewDetails.aspx?src=IND&mtdsg_no=XXVII-1&chapter=27&clang=_en

United Nations, (2014). System of Environmental Economic Accounting 2012 – Central Framework. URL: https://unstats.un.org/unsd/envaccounting/seearev/seea_cf_final_en.pdf

Verkhovna Rada of Ukraine, (1996). Pro ratyfikatsiiu Ramkovoi konventsii OON pro zminu klimatu [Ratification of the United Nations Framework Convention on Climate Change] [in Ukrainian]. URL: https://zakon.rada.gov.ua/laws/show/435/96-%D0%B2%D1%80#Text

Verkhovna Rada of Ukraine, (2010). Pro osnovnizasady (stratehiiu) derzhavnoi ekolohichnoi polityky Ukrainyna period do 2020 roku [Basic principles (of the strategy) of the environmental policy of Ukraine till 2020] [in Ukrainian]. URL http://zakon3.rada.gov.ua/laws/show/2818-17

Verkhovna Rada of Ukraine, (2016). Ratification of Paris Agreement [in Ukrainian]. URL: https://zakon.rada.gov.ua/laws/show/995_l61#Text.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
Joint Special Issue: A New Role for Statistics
Vol. 24, No. 1, pp. 89–103, DOI 10.59170/stattrans-2023-006
Received – 01.07.2022; accepted – 02.11.2022

Method of auditing in conditions of martial law

Olha Lubenchenko[1]**, Svitlana Shulga**[2]**, Halyna Pavlova**[3]

ABSTRACT

In the article there are considered methodical recommendations on the actions of auditors during martial law. They relate to such stages of the audit as the preparatory phase, the planning phase, the task implementation and the final phase. The preparatory stage requires the identification of the client and the conclusion of an audit agreement. Under martial law, new risks are emerging, systematized by the authors and related to the identification of persons involved in terrorist activities and the proliferation of weapons of mass destruction. The systematization of risks and the use of IT technologies to identify and verify the client allows auditors, as specially designated entities of primary financial monitoring, to identify and freeze the assets of such persons in a timely manner and stop providing any services. At the client acceptance and assignment stage, auditors assess ethical threats. As a result, a working paper has been developed to assess ethical threats in the light of martial law. At the planning stage, special attention should be paid to reviewing risks, namely how military aggression affects the continuity of any business. Clarification of risk factors for termination allows you to plan audit procedures to gather audit evidence and further determine the auditor's opinion on continuity (opinion with an explanatory paragraph, opinion with a reservation, negative opinion, disclaimer of opinion). One of the effective procedures for collecting audit evidence at the stage of the task is inventory.

Key words: war, risk, business entities.

1. Introduction

The largest military conflict in human history – World War II – ended on September 2, 1945. Military action in Europe virtually ceased in May 1945. But on February 24, 2022, the continent was again on the brink of war. Ukraine, a country

[1] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: olga.lubenchenko@gmail.com. ORCID: https://orcid.org/0000-0002-4209-8929.

[2] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: StShylga@ukr.net. ORCID: https://orcid.org/0000-0002-2396-2507.

[3] Dnipro State Agrarianand Economic University. Dnipro, Ukraine. E-mail: halynapavlova@ukr.net. ORCID: https://orcid.org/0000-0002-1400-7348.

© O. Lubenchenko, S. Shulga, H. Pavlova. Article available under the CC BY-SA 4.0 licence

in central Europe, called for an invasion by the aggressor, the Russian Federation, which cynically called the war a “special operation”. Fierce fighting is taking place in Ukraine, some cities and towns are under occupation, there are already a significant number of civilian and military casualties, destroyed cities and businesses.

According to the Kyiv School of Economics, the amount of direct damage to Ukraine's infrastructure during the war reached $ 63 billion or UAH 1.8 trillion as of March 24 (Table 1). Total losses of Ukraine's economy due to the war are projected, taking into account both direct and indirect losses (fall in gross domestic product, cessation of investment, outflow of labor, additional spending on defense and social support, etc.) range from $ 543 billion to $ 600 billion (https://kse.ua/ua/about-the-school/news/zbitki-naneseni-infrastrukturi-ukrayini-v-hodi-viyni-skladayut-mayzhe-63-mlrd/, 2022).

Table 1. Losses of the economy of Ukraine from damage to physical infrastructure (in case of complete destruction of facilities), from the beginning of hostilities until March 24, 2022

Infrastructure facilities	Number of units	Total losses, million US dollars
Roads	8265	27 546
Residential buildings	4431	13 452
Civil airports	8	6 816
Factories and enterprises	92	2 921
Healthcare facilities	138	2 466
Nuclear power plants	1	2 416
Railway infrastructure and rolling stock	н/д	2 205
Bridges and bridge crossings	260	1 452
Ports and port infrastructure	2	622
Institutions of secondary and higher education	378	601
Administrative buildings	35	574
Military airfields	10	390
The AN–225 “Mriia” plane	1	300
Shopping and entertainment centers	11	188
Religious buildings	44	150
Cultural buildings	42	144
Kindergartens	165	133
Thermal and hydroelectric power plants	7	101
Other	x	412
Total	x	62 889

Source: (Kyiv School of Economics).

The authors' surveys of audit firms show a significant loss of audit clients, which ranges from 60 to 80 percent of the total amount of contracts.

Such horrific events have created new challenges for the audit community – the provision of services under martial law. Auditors face many ethical and technical issues regarding the methodology of auditing financial statements and providing other audit services. International Standards on Auditing do not provide specific guidance on the conduct of auditors during martial law. There are already clarifications from the Public Oversight Audit Authority (Audit Public Oversight Body of Ukraine, 2022) the Audit Chamber of Ukraine (Audit Chamber of Ukraine) on the actions of auditors in martial law (Council of the Chamber of Auditors, 2022), but as soon as possible the auditors need methodological developments taking into account the new risks caused by the war.

2. Purpose and methodology of the article

The purpose of the article is to provide auditors with practical advice on accepting a client under martial law, planning an audit, risk assessment, including risks related to business continuity, collection of audit evidence, their assessment, the formation of audit opinion.

Conventionally, the audit process can be divided into several stages, as presented in table 2.

Table 2. Stages of audit and actions of the auditor

№	Audit stage	Auditor's actions
1	Preparatory stage	Implementation of procedures for accepting the client and the task, agreeing on the terms of the task, drawing up a letter - agreement, contract, preliminary assessment of risks and ethical requirements, control measures
2	Planning	Development of strategy and plan, clarification and assessment of risks, determination of materiality limits both for financial reporting as a whole and for performance of audit procedures, definition of audit procedures, control measures
3	Completion of the task	Execution of audit procedures, collection of audit evidence, their evaluation, control measures
4	The final stage	Formation of the auditor's opinion, control measures

Source: (International standards of quality control, audit, inspection, other assurance and related services, 2018).

At each stage of the audit, certain principles are applied, such as the principle of assessing the risks of the client and the audit, business continuity, internal standardization; methods, in particular, deduction, induction, analogy, abstraction, classification, synthesis, analysis, observation, measurement, comparison, consolidation and grouping, comparative law methods, elimination, balance method, professional judgment, procedures for collecting and summarizing audit evidence, formation of audit thoughts. The audit procedures are presented in Table 3. The set of principles, methods and procedures form the basis of the audit methodology. The methodology is correlated based on the specifics of the tasks, defining the purpose, features of the presentation of audit results and is the algorithm of the auditor's actions.

Table 3. List of audit procedures used in the audit stages

Auditprocedures	Preparatory stage	Planning	Completion of the task	Final stage
Identification	x	x	x	x
Analyticalprocedures	x	x	x	x
Request	x	x	x	x
Observation	-	-	x	-
Inspection	x	x	x	x
Rating	x	x	x	x
Audit	x	-	x	x
Confirmation	x	x	x	x
Recalculation	x	x	x	x
Repeatexecution	-	-	x	-

Source: International standards of quality control, audit, inspection, other assurance and related services, 2018.

3. The impact of military aggression on the stages and risks of the audit

Each stage of the audit was affected by events directly related to the military aggression. Therefore, today it is necessary to clarify the algorithm of the auditor's actions.

The preparatory stage is primarily concerned with identifying the client and concluding an audit agreement. Under martial law, together with the Law of Ukraine "On Auditing Financial Statements and Auditing Activities" of December 21, 2017 № 2258 – VIII (the Law № 2258) auditors, as specially designated subjects of primary financial monitoring, are obliged to apply provisions of the Law "On Prevention and Counteraction to Legalization (Laundering) of Proceeds from Crime, Financing of Terrorism and Financing of Proliferation of Weapons of Mass Destruction" of December 6, 2019 № 361-IX (the Law № 361). Law 361 requires a risk-based approach to assessing audit clients (Verkhovna Rada of Ukraine, 2022).

The State Financial Monitoring Service provided a list of new risk indicators caused by the war (Table 4).

Table 4. Risk assessment by auditors in the financial monitoring system

№	Type of risk	Procedures
1.	Clients are - residents of Ukraine who have shares in the authorized capital, formed by individuals and legal entities registered in the Russian Federation and the Republic of Belarus and in uncontrolled territories, clients or founders of clients - citizens of Russia and Belarus, as well as citizens of Ukraine, registered in the temporarily occupied territories of Ukraine, who opened accounts in banking institutions of Ukraine/received transfers during the year before the war; - individuals - non-residents; - law enforcement officers of the Russian Federation and the Republic of Belarus, who took/are taking part in the criminal activity of the aggressor country on the territory of Ukraine and Europe;	In conditions of limited use of open state registers and other Internet resources, when accepting a client during martial law: 1. to request and examine statutory documents and / or their copies (charter, extract or extract of the Unified State Register of Individuals - Entrepreneurs and Legal Entities; for non-resident legal entities. authorized body of a foreign state on the registration of the relevant legal entity; 2. in the letter-agreement, or in the initial request to the client it is obligatory to provide responsibility of the client for identity of the original of the provided copies of documents; provide for the possibility of providing such copies via e-mail or through the organization of remote access to customer databases, which prevents unsanitary delivery of information (automated), electronic communication, information and communication systems, electronic communication networks of the client.
1.1.	- legal entities - residents of Ukraine, which are official representative offices of Russian and Belarusian companies;	
1.2.	- residents of Ukraine who have permanent business relations with Russian/Belarusian individuals or legal entities;	1. in the request for financial monitoring and the client's questionnaire, which are developed within the framework of the Law № 361, provide for the provision of information on business relations with certain persons; 2. in the future, to provide audit procedures for the study of primary documents for certain contracts and transactions.
1.3.	- non-residents with hidden ultimate beneficial owners (controllers) who have open accounts in Ukrainian banks;	1. to request and examine documents and their copies: for non-resident legal entities, copies of the legalized extract from the trade, banking or court register on the registration of the relevant legal entity; 2. by requesting to find out whether the client cooperates with prominent entities; 3. in the future, to provide audit procedures for the study of primary documents for certain contracts and transactions.
1.4.	- non-residents registered in offshore jurisdictions and jurisdictions with weakened tax/currency controls that have open accounts in Ukrainian banks.	

Source: Verkhovna Rada of Ukraine, 2001; 2022.

Table 5. Transfer of funds outside Ukraine on the basis of pretended additional agreements

№	Type of risk	Procedures
1.	Transfer of funds outside Ukraine on the basis of alleged additional debt assignment agreements, offsetting of claims under import contracts concluded with residents of Russia and Belarus, in order to withdraw funds to other enterprises, including other jurisdictions.	1. In the request for financial monitoring and client questionnaires, which are developed under the Law № 361 to provide information on the conclusion of debt assignment agreements, offsetting on import contracts, securities transactions, etc. 2. In the future, to provide audit procedures for the study of primary documents for certain contracts and transactions
1.1.	transfer of funds outside Ukraine on the basis of alleged additional debt assignment agreements, offsetting of claims under loan agreements (loans, credits, financial assistance) concluded with residents of Russia and Belarus, in order to return borrowed funds to other enterprises, including other jurisdictions (lack of complete information on the total amount of the loan, which allows repayment of loans, interest payments after many years);	
1.2.	carrying out any transactions with securities issued by residents of Russia and Belarus or persons affiliated with such residents.	
2.	Impossibility to identify and verify the client.	To avoid the impossibility of identification, apply procedures: 1. to request a list of persons who are final beneficiaries, persons who make management decisions and copies of their passports; 2. verification - is the establishment that the identification documents (passports) belong to the person who represents the interests of the client, to do this, use the procedure of interviewing a person through the Zoom platform and record such an interview via video.

Source: The State Financial Monitoring Service of Ukraine, 2022.

In case of identification of relations with natural or legal persons subject to sanctions, the audited entity is obliged to freeze the assets of such persons and stop providing any services. Asset freezing is a ban on the transfer, conversion, placement, movement of assets related to terrorism and its financing, proliferation and financing of weapons of mass destruction, based on UN Security Council resolutions, foreign decisions, courts [8]. In addition, it should be noted that the Law of Ukraine of March 3, 2022 № 2116–IX "On Basic Principles of Compulsory Seizure of Property in the Russian Federation and Its Residents" provides for the forcible seizure of property of the Russian Federation and objects of property rights of its residents on the grounds of public necessity (including cases in which it is urgently required by military necessity) in favor of the state of Ukraine on the basis and in the manner prescribed by this Law.

At the client acceptance stage, an ethical threat assessment is required. The auditors have already dealt with the specifics of assessing the threat of ethics during the pandemic. Recommendations of the International Federation of Accountants (IFAC) "The use of specialists in COVID-19, including considerations for involving specialists in the audit of financial statements. October 2020" were the basis for the development of ethical threat assessment procedures (The International Federation of Accountants (IFAC), 2020). Martial law and the events surrounding the aggressor's invasion also affect such procedures, so the recommendations in a slightly modified form can be applied in 2022 (Table 6).

Table 6. Excerpt from the working document "Assessment of ethical threats related to martial law in an audit firm"

Clientis LLC "XXX"
Date of financial reporting: 31.12.2021
The purpose of the working document: assessment of threats to independence during martial law

№	Issues for consideration	Comments (audit procedures for identification, assessment and counteraction to threats)
1	Are there individual and organization-wide gaps in the skills, knowledge or experience required to fulfill new reporting responsibilities or compliance with martial law requirements	To assess in the working paper whether the audit firm can work remotely
2	Are there sufficient resources to respond to sudden changes and uncertainties arising from martial law, as well as additional tasks and responsibilities that may be required by clients Is it possible in practice to meet immediate needs (eg in resources), given the current travel constraints and personal interactions	To assess in the working document whether the audit firm's employees, for example through the Zoom platform, can be present at client meetings related to the audit, be present during the selective inventory of inventory (the possibility of creating so-called virtual groups)

Table 6. Excerpt from the working document "Assessment of ethical threats related to martial law in an audit firm" (cont.)

№	Issues for consideration	Comments (audit procedures for identification, assessment and counteraction to threats)
3	Is the IT infrastructure able to meet the technological needs of remote groups Is there an increased risk of, for example, fraud, cyber threats	To develop rules for access to electronic documents, rules for their correction, rules for using a remote server (transfer of sensitive data to a remote server or "cloud" storage) To purchase software that prevents threats to sensitive information and prevents unauthorized access to audit firm databases
4	Can the nature or urgency of a customer's request affect a firm's ability to support its existing customers or address martial law issues that affect the firm's own business	To develop procedures for assessing the continuity of the audit firm in martial law and uncertainty
5	Awareness, knowledge and access of the firm to the specific legal and regulatory requirements related to martial law and related recommendations and the ability of the firm to timely synthesize and disseminate knowledge in customerser viceteams	To develop procedures for interaction of the audit firm's staff (daily meetings, reporting to management)
6	Availability of infrastructure, tools and people with sufficient and relevant experience for training and supervision of virtual groups	Ability to create virtual groups and control them
7	Conclusion on ethical threats: 1) threat to self-interest; 2) the threat of self-assessment; 3) threat to protection; 4) the threat of personal relationships; 5) the threat of pressure identified (or not identified), measures to minimize (eliminate).	Conclusion on threats (indicate threats), identified (or not identified), measures to minimize (eliminate).

The document was drafted: name, date, signature
The document was checked: name, date, signature

Source: The International Federation of Accountants (IFAC), 2020.

The proposed working paper makes it possible to assess ethical threats taking into account the factors caused by martial law.

At the planning stage, special attention should be paid to reviewing the risks in the light of military aggression. Martial law is a factor that significantly affects the continuity of any business entity. International Standards on Auditing 570 "Continuity of Activity" defines operational, financial and other factors that affect continuity, but the standard does not provide for the assessment of continuity during hostilities (Information letterof the Boardof the Audit Chamber of Ukraine Consideration by the auditor of the continuity of activities during the audit of financial statements, 2022). All these factors must be considered through the prism of military events. condition (table 7).

Table 7. Operational factors affecting the continuity of the business entity taking into account martial law

Factors	The essence of the factors	Clarification of the essence of the factor (the impact of martial law)
Operating	1. Intentions of management to liquidate an entity or to cease operations. 2. Loss of key management personnel without its replacement. 3. Loss of the main market, key customer (s), franchise, license or main supplier (s). 4. Difficulties with the workforce. 5. Lack of important resources. 6. The emergence of a very successful competitor.	1. The impact of martial law on the company's personnel (mobilization, joining the territorial defense) and the costs associated with the payment of personnel; 2. Suspension or interruption of activities due to disruption of the supply chain, termination of operations, loss of production capacity or commercial facilities, restriction of movement and disruption of logistics; 3. Damage or destruction of property; 4. Failure to comply with the terms of contracts due to force majeure, adverse changes in the terms of contracts, breach of credit agreements, inability to repay accounts payable and delays in repayment of receivables; 5. Significant reduction in sales, profits, cash flows from operating activities.

Source: International standards of quality control, audit, inspection, other assurance and related services, 2018.

In table 8 there are discussed the financial factors that affect the business continuity of the entity, taking into account martial law.

Table 8. Financial factors affecting the continuity of the business entity taking into account martial law

Factors	The essence of the factors	Clarification of the essence of the factor (the impact of martial law)
Financial	1. Excess of liabilities over assets or excess of current liabilities over current assets. 2. Loans with a fixed term, the repayment of which is approaching, in the absence of real prospects for prolongation or repayment; or excessive use of short-term loans to finance long-term assets. 3. Signs of cancellation of financial support by creditors. 4. Negative cash flows from operating activities, as evidenced by financial statements for previous periods or projected financial statements. 5. Negative key financial ratios. Significant operating losses or significant reductions in the value of assets used to generate cash flows. 6. Debts or termination of dividends. 7. Inability to pay creditors on time. Inability to fulfill the terms of loan agreements. 8. Transition from the system of payment on credit for the delivered goods, received service to the payment system at the time of delivery of goods, receipt of services. 9. Inability to obtain financing for the development of new basic products or to finance other significant investments.	1. Arrest or expropriation of assets for the needs of the state after December 31, 2021; 2. Restrictions on access to cash and cash equivalents or restrictions on cash transactions; 3. Impairment of financial and non-financial assets (including events and information after the reporting date); 4. Instability and significant changes in prices for capital instruments, debt securities, commodity prices, foreign exchange rates and/or interest rates after December 31, 2021, which will significantly affect the assessment of assets and liabilities, income and expenses in the following 12 months.

Source: International standards of quality control, audit, inspection, other assurance and related services, 2018.

In table 9 there are analyzed other factors that affect the business continuity of the entity, taking into account martial law.

Table 9. Other factors affecting the continuity of the business entity, taking into account martial law

Factors	The essence of the factors	Clarification of the essence of the factor (the impact of martial law)
Other factors	1. Failure to comply with capital requirements or other legal or regulatory requirements, such as solvency or liquidity requirements for financial institutions. 2. Incomplete legal or regulatory proceedings against an entity that, if satisfied, may give rise to claims that the entity is unlikely to be able to satisfy. 3. Changes in laws or regulations or government policies that are expected to adversely affect the entity. 4. Lack of insurance or insufficient insurance against disasters in case of their occurrence.	1. Announcement of plans to terminate or dispose of fixed assets; 2. Other circumstances that significantly affect the activity.

Source: International standards of quality control, audit, inspection, other assurance and related services, 2018.

Clarifying termination risk factors allows auditor to plan appropriate audit procedures to gather audit evidence and form the auditor's opinion on continuity (table 10).

Table 10. The impact of the assessment of audit evidence on the continuity of the client's activity in the auditor's opinion

Evaluation of audit evidence	The impact of the assessment on the auditor's opinion
There is significant uncertainty about the company's ability to continue as a going concern	1. If continuity is used as a basis for accounting, but there is significant uncertainty when significant uncertainty is properly disclosed in the financial statements, the auditor's report should have a separate section entitled "Significant uncertainty regarding continuity" (paragraphs A21–A22, A28–A31, A34 [5]) 2. If the use of the going concern assumption is acceptable, but there is significant uncertainty when the information about material uncertainty is not properly disclosed in the financial statements, the auditor should express a qualified or negative opinion (paragraphs A24–A25, A32). –A34 [5])

Table 10. The impact of the assessment of audit evidence on the continuity of the client's activity in the auditor's opinion (cont.)

Evaluation of audit evidence	The impact of the assessment on the auditor's opinion
Events and conditions are expected to have some impact, but are not expected to be significant. There is considerable uncertainty, but such uncertainty is not significant	The issue of continuity is a key audit issue.
Deterioration of financial condition and results of operations indicates that the assumption of business continuity is no longer acceptable.	1. If the financial statements have been prepared using a basis of accounting other than that required by the conceptual basis, when the information that the financial statements have not been prepared on a going concern basis is properly disclosed in the financial statements, the auditor's report on the financial statements should include an explanatory paragraph users of the auditor's report that the financial statements have been prepared in accordance with a fundamentally changed basis of accounting, for example on the basis of termination or liquidation. 2. If the use of the going concern assumption is not acceptable, but the financial statements have been prepared using the going concern assumption as the basis of accounting, the auditor should express an adverse opinion. (paragraphs A26–A27 [5])

Source: International standards of quality control, audit, inspection, other assurance and related services, 2018.

4. Inventory as one of the effective procedures for collecting audit evidence

One of the effective procedures for collecting audit evidence is inventory. During hostilities there is destruction, loss of material assets, reduction of their usefulness. Regulations on inventory of assets and liabilities, approved by the order of the Ministry of Finance of Ukraine dated 02.09.2014 № 879 provides for an inventory, in particular, in case of establishing the facts of damage to property (on the date of such facts) the head of the enterprise; in case of man-made accidents, fire or natural disaster (on the day after the end of the phenomena) in the amount determined by the head of the enterprise. conduct an anti-terrorist operation (or their structural units (separate property) are located in these areas), conduct an inventory in cases required for its conduct, when it becomes possible to ensure safe and unimpeded access of authorized persons to assets, primary documents and registers of accounting, which reflects the liabilities and equity of these under enterprises (Ministry of Finance of Ukraine, 2020).

According to the international standard of auditing “Audit evidence – special provisions”, if the inventory is carried out on a date other than the date of financial reporting, the auditor must perform additional audit procedures to obtain audit evidence of proper reflection of changes in inventories between the date of calculation and financial reporting date. Such evidence can be obtained through the procedure of inventory supervision, inventory inspection (presence of the auditor during the inventory). In conditions of limited access to assets due to martial law, it is possible to organize the presence of the auditor during the inventory by using IT technologies: Zoom conference programs, Microsoft Teams, video cameras and recording the inventory procedure conducted by client staff for further audit. This will help the auditor to indirectly confirm the presence or absence of assets, to record the facts of damage and damage. Thus, establishing at the level of the audit firm the possibility of video recording of the inventory process and the use of video conferencing with management of the customer will help the auditor to indirectly confirm the presence or absence of assets, record damage and injury, form the auditor's opinion on the financial statements.

5. Conclusions

The war in Ukraine has forced auditors to tackle new challenges in complying with the latest legal requirements for identifying those involved in military aggression against Ukraine, on the one hand, and requiring careful compliance with International Standards on Auditing. All stages of the audit require the staff of audit firms to take into account such a factor as martial law. Systematized risks associated with identifying the client during the task at the preparatory stage will allow the auditor to identify persons registered in the Russian Federation and the Republic of Belarus, identify their assets in Ukraine, timely provide information to the State Financial Monitoring Service. The working document “Assessment of threats to the ethics of martial law in an audit firm” allows auditor to assess such threats and take measures to minimize or eliminate them. New risk factors for business continuity, which are summarized on the basis of International Standards on Auditing and the clarifications of the Audit Chamber of Ukraine, should also be taken into account when planning an audit. The auditor evaluates the impact of certain factors on the client's financial statements evaluates the audit evidence and, depending on the evidence collected, forms an opinion on the financial statements at the final stage. The auditor uses asset inventory or inspection procedures to perform the engagement. The use of conference software Zoom, Microsoft Teams, inventory videos without physical presence allows auditors to indirectly participate in the inventory and obtain evidence of the availability and condition of assets. Further research will focus on improving the auditor's working papers in order to provide auditors with practical tools.

References

Audit Chamber of Ukraine. Shchodo vplyvu viiskovoi ahresii Rosiiskoi Federatsii na finansovu zvitnist ta audytorski zvity [Concerning the influence of the military aggression of the Russian Federation of financial reporting and auditor's reports]. Retrieved from: https://www.apu.com.ua/wp-content/uploads/2022/03/%D0%86%D0%9D%D0%A4%D0%9E%D0%A0%D0%9C%D0%90%D0%A6%D0%86%D0%99%D0%9D%D0%98%D0%99-%D0%9B%D0%98%D0%A1%D0%A2-%D0%90%D0%9F%D0%A3.pdf [in Ukraine].

Audit Public Oversight Body of Ukraine, (2022). Updated information on reporting by audit entities. Retrieved from: https://www.apob.org.ua/?p=3743&lang=en

Council of the Chamber of Auditors, (2022). Shchodo vplyvu viiskovoi ahresii Rosiiskoi Federatsii na okremi aspekty diialnosti subiektiv audytorskoi diialnosti: Informatsiinyi lyst [Regarding the influence of the military aggression of the Russian Federation on certain aspects of the activities of the subjects of audit activity: Information sheet]. Retrieved from: https://ips.ligazakon.net/document/FN073557 [in Ukraine].

Kyiv School of Economics. Priami zbytky, naneseni infrastrukturi Ukrainy v khodi viiny. [Direct damage caused to the infrastructure of Ukraine during the war]. Retrieved from: https://kse.ua/ua/about-the-school/news/zbitki-naneseni-infrastrukturi-ukrayini-v-hodi-viyni-skladayut-mayzhe-63-mlrd/ [in Ukraine].

Ministry of Finance of Ukraine, (2020). Pro zatverdzhennia Polozhennia pro inventaryzatsiiu aktyviv ta zoboviazan [On the approval of the Regulation on the inventory of assets and liabilities]. Retrieved from: https://zakon.rada.gov.ua/laws/show/z1365-14?lang=en#Text [in Ukraine].

International standards of quality control, audit, inspection, other assurance and related services, (2018). Phoenix Publishing House, 1142 p.

The International Federation of Accountants (IFAC), (2020). Vykorystannia spetsialistiv v umovakh COVID-19, u tomu chysli mirkuvannia shchodo zaluchennia spetsialistiv do audytu z finansovoi zvitnosti [The use of specialists in the context of COVID-19, including considerations regarding the involvement of specialists in the audit of financial statements]. Retrieved from: https://www.ethicsboard.org/publications/using-specialists-covid-19-environment

The State Financial Monitoring Service of Ukraine. (2022). Shchodo mizhvidomchoi koordynatsii systemy finansovoho monitorynhu [On interagency coordination of the financial monitoring system]. Retrieved from: https://fiu.gov.ua/pages/dijalnist/funkcional/news/shhodo-mizhvidomchoji-koordinacziji-sistemi-finansovogo-monitoringu.html [in Ukraine].

Verkhovna Rada of Ukraine, (2001). The Criminal Code of Ukraine [Kryminalnyi kodeks Ukrainy]. Retrieved from: https: //zakon.rada.gov.ua/laws/show/2341-14#Text [in Ukraine].

Verkhovna Rada of Ukraine, (2017). Pro audyt finansovoi zvitnosti ta audytorsku diialnist: Law of Ukraine [About the audit of financial statements and audit activity: Law of Ukraine] Retrieved from: https://zakon.rada.gov.ua/laws/show/2258-19#Text [in Ukraine].

Verkhovna Rada of Ukraine, (2022). Pro zapobihannia ta protydiiu lehalizatsii (vidmyvanniu) dokhodiv, oderzhanykh zlochynnym shliakhom, finansuvanniu teroryzmu ta finansuvanniu rozpovsiudzhennia zbroi masovoho znyshchennia: Law of Ukraine [On the prevention and countermeasures against the legalization (laundering) of proceeds obtained through crime, the financing of terrorism and the financing of the proliferation of weapons of mass destruction: Law of Ukraine]. Retrieved from: https://zakon.rada.gov.ua/laws/show/361-20#Text [in Ukraine].

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 105–130, DOI 10.59170/stattrans-2023-007
Received – 06.07.2022; accepted – 20.10.2022

Current challenges related to the consumer price index (CPI) in Ukraine

Olga Vasyechko[1]

ABSTRACT

The purpose of this study is to contribute to the maintenance and compilation of the consumer price index (CPI) in the current extreme situation caused by the Russian military aggression against Ukraine. In these extreme conditions, official statistics is faced with the task of maintaining the regularity, completeness and quality of the production of statistical information, including the CPI, which is one of the key economic indicators. The interaction between the ideal and conditional concepts of the index and their practical implementation is considered as a potential source of compilation improvement. The author argues that the main factor of the modern criticism of the CPI is the systematic deviation of the practical form of the index from its theoretical foundations. One way to solve this problem is to use new sources of information, especially big data cash registers. In today's extreme conditions, cash data can extensively address the issue of limited and untimely access to primary data sources needed to compile the CPI, as well as promptly take into account the changes in consumption patterns caused by significant migratory flows from the dangerous areas, and changes in the supply offer due to the rupture of supply chains.

Key words: consumer price index (CPI), Russian military aggression against Ukraine, ideal concept of CPI, conditional concept of CPI, cash registers, Big Data.

1. Introduction

The study focuses on current challenges facing consumer price statistics in terms of the evolution of views on the nature of the price index in accordance with current capabilities of official statistics and, above all, technological provision, and how the latter radically changes opportunities and challenges in compiling the CPI. This issue is relevant not only in terms of the CPI methodology, but also in light of the current limitations in obtaining primary data due to the war unleashed by the Russian

1 University Paris 1 Panthéon-Sorbonne, France. E-mail: o.vasyechko_stat@yahoo.fr.
ORCID: https://orcid.org/0000-0001-7461-3086.

© Olga Vasyechko. Article available under the CC BY-SA 4.0 licence

Federation against Ukraine. In conditions where the sources of primary information are partially or completely inaccessible, official statistics requires a revision of the entire paradigm of index compilation starting with the collection of primary data to the method of their aggregation.

The existing challenges for the Consumer Price Index (CPI) are due not only to the current crisis situation, but have formed against the background of the inherent problems of this statistical indicator, which has existed since its inception. The CPI is one of the oldest statistical indicators, which is always in high demand and scrupulous attention from society. The CPI has gone through various stages of its development, both in terms of theoretical justification and methodological provision, and in terms of its practical implementation (Stoevska, 2018). But despite numerous efforts to improve the quality of the index, the CPI continues to be exposed to numerous critical remarks from experts and the general public. The question rightfully follows as to why, for more than a hundred years, the CPI has continued to provoke criticism, and what is the source of the so-called "imperfection" of the CPI, as well as what may be the ways to solve the issues in modern conditions.

For Ukraine, the current period of the CPI development is especially important, because along with the constant search for ways to improve the formula of the index it has set the task of producing official statistics in extreme conditions of military aggression. There is a need to revise the classical approaches to determining information sources and methods of information acquisition, processing and aggregation. This is certainly facilitated by the rapid development of e-commerce in goods and services, web resources and information technology, production and accumulation of big data. Access to fundamentally new sources of primary information creates the opportunity to review the traditional practice of calculating the CPI.

The paper comprises seven sections. Sections 2 to 4 describe the development of ideal concept of the CPI and its conditional concept and argue the reasons for the deviation of the latter from the index number theory. Section 5 deals shortly with issues arising in practical implementation of the conditional CPI concept, while Section 6 considers new data sources and respective prospects for the CPI. Section 7 concludes the essential findings of the study.

2. General issues

The officially accepted formula for calculating and aggregating the CPI is the Laspeyres Index. And the central issue of the modern methodological approach to the CPI is the fact that according to current practice, both at the highest level of aggregation and at the level of elementary indices, the CPI is not a true Laspeyres index (Guide complémentaire, 2004). In this regard, two questions arise:

1) Why is the Laspeyres formula recommended for calculating the aggregate CPI?
2) What is the reason that in practice there has been such a significant deviation from the chosen theoretical formula?

Finding out the answers to these questions is important for being able to go beyond the established limits of the index compilation in order to adapt it to modern extreme conditions. The choice of the theoretical formula of the index and the further departure from its classical form in practical statistical activities are due to the simultaneous coexistence of theoretical, methodological and practical principles of the CPI. The source of problems in the implementation of the CPI is in the plane of constant confrontation of theory and practice. In the latest, 2020 edition, the CPI manual states that in modern conditions, it is possible to use different conceptual frameworks to address fundamental issues related to the nature of the index. At the same time, it is the conceptual basis of the index and the scope of use that should determine the method of its compilation (Consumer price index manual, 2020). The path from index theorizing to its practical implementation can be outlined using the following three levels (Sillard, 2017):

- *Defining the ideal concept of the index*, which is not directly observed, but the existence of which is a postulate of current classical theories.
- *Defining a conditional concept* that is generally acceptable and complements the ideal concept with a number of formulated conditions, thus bringing it closer to reality.
- *Practical implementation of the conditional concept* through its empirical validation through a series of experimental observations and estimates followed by full-scale statistical observations and index calculations using real data.

It should be understood that the results obtained during its implementation of the last stage serve as a source of improvement not only for the conditional concept, but also for the ideal concept of the index. In this regard, W. Erwin Diewert and Robert J. Hill proposed concepts such as *a true index* and an *exact index*. The true index corresponds to the concept of an ideal index. Such an index represents a goal, namely a phenomenon that needs to be measured. The exact index, in turn, is a practical implementation of the conditional concept of the index, while the latter is a transitional link between true and empirical indices. That is, the exact index is considered acceptable if it corresponds to the true index under the existing conditions with a sufficiently high level of approximation (Diewert and Hill, 2010).

Intervention in order to adapt the CPI formula to the extreme conditions of war in Ukraine is possible at all three levels. The most radical changes obviously occur when reviewing the theoretical foundations, namely the ideal concept of the index. But this is a very complex process, and below we will consider the conditions under which it becomes not only possible but also appropriate.

3. The ideal CPI concept

The ideal concept of the CPI is based on both the theory of consumer behaviour and the theory of indices, while its conceptual framework is the System of National Accounts (SNA).

The theory of consumer behaviour suggests that the individual with the existing budget constraints behaves rationally, and his/her preferences can be transitive, and the choice is always in favour of cheaper goods.

In the context of the CPI as an aggregate measure of price levels evolution, it is the behaviour of the *aggregate representative consumer* that is an abstract being that represents all households together with their consumption in a given economic area and during elementary periods of fixed duration.

Ideally, household consumption means *the actual consumption of households within market (monetary) transactions* freely agreed between the supplier and the buyer (so-called "market" consumption). *Actual household consumption covers all goods and services* purchased by households to meet their own needs, whether or not these purchases were subject to expenditure by households. Actual consumption is thus broader in scope than final consumption, which corresponds only to household spending on consumer goods.

Actual consumption is considered to occur in a certain economic area. That is, it is a *place of purchase of goods or services by a representative consumer to meet his consumer needs*, not the place of his residence. The ideal index assumes that *time of purchase and time of payment for goods or services coincide in time, and the period of consumption is as close as possible to the time of purchase and payment*. Consumer demand always takes the form of goods or services, such as certain products that are sold in a commercial entity in a certain economic jurisdiction during the elementary period.

There are two types of elementary periods: those that are well comparable and those that are poorly comparable. Between two *well-comparable periods, consumption needs are identical in nature*. Instead, the number of units consumed may vary. *Periods that are poorly compared are combined according to the index number theory*.

It is believed that the utility function, which characterizes the consumer's choice, is constant between two well-comparable periods. The budget for consumption is considered exogenous (wages received and savings available).

The ideal concept of a consumer price index reflects a change in the budget between two well-comparable periods acceptable to a representative consumer to maintain a constant level of utility.

On the other hand, between two points belonging to two well-comparable periods, the same consumer need can be met by two different products, due to either a change in product characteristics or the disappearance of a previous product from a particular market and, consequently, the emergence and choosing to consume its substitute. Two products that meet the same consumer needs at the two time periods do not have the same characteristics and may not provide the same level of usefulness for a representative consumer. In this regard, adjustments are provided to eliminate the impact of changes in the quality on the level of consumer utility. *The ideal index should reflect the change in unit price with a constant level of quality.*

Thus, the ideal consumer price index should measure *the overall price evolution of all products (goods and services)* consumed by a representative consumer, i.e. be the average value of the evolution of these products. But since the weight of each individual product for the aggregate representative consumer is different, the simple average for calculating the overall index does not take into account the economic weight of each product consumed. In 1812, Young was the first to suggest an approximate weighting of relative prices according to their relative price for the period selected for the study, but did not specify an exact formula for determining weights (Consumer price index manual, 2004).

Hence, *the ideal consumer price index should be a weighted average value, the weight of which reflects the consumption structure of the aggregate representative consumer and, at the same time, does not distort the real dynamics of prices.*

In turn, as the average value of changes in the level of value consumed by the aggregate consumer, the ideal price index must demonstrate such an important property as *associativity*, therefore the ability to aggregate (disaggregate) from any level - price level, elementary indices, sub-indices to the highest level of aggregation.

In order to be able to compare poorly comparable periods, the ideal index should also be compiled into time series, in which any two levels can be compared both in the direction of past periods and in the direction of future periods. The main properties required to compile time series are *transitivity (or circularity) and reversibility.* Due to transitivity, it is also possible to find unknown values of the time series having, for example, the final value of the price level and the chain of indices. Reversibility, in turn, allows moving the base year in a dynamic series of indices, as well as in the direction of past and future periods. There are two paradigms for compiling time series (Balk, 2004): by directly comparing the levels of a data set and by constructing chain indices. The choice of paradigm depends on the chosen formula of the index.

The formula of the ideal price index is generally accepted to be a positive function, hence to demonstrate such property as *monotony* (von Auer, 2004).

Different approaches have gradually been developed to find an index formula that can best meet the requirements formulated for the ideal CPI concept. Today, there are two major schools for measuring inflation in the consumer sector of the economy: the Consumer Price Index (CPI) and the Cost of Living Index (CPI).

Most economic indices, including price indices, are based on the Divisia index, named after the French economist François Divisia, who in 1926 proposed a theoretical construct for building index number series for continuous market data on prices and quantities (Hulten, 2008). But in reality, economic data metrics are discreet, not continuous. Therefore, the reference to Divisia dynamic index numbers are usually the time series that exploit the Divisia's idea but are built for discrete time periods, which can be done using various formulas discussed below.

The general argument for choosing an index formula corresponds to three key approaches, such as the economic approach, axiomatic approach and stochastic approach (Diewert, 1995).

Consumer price index

The economic approach assumes that the CPI aims to detect changes in the level of prices for consumer goods, regardless of the dynamics of consumed quantities of these goods between the two comparative periods, which is achieved by neutralizing the impact of quantitative changes on prices. To solve this problem, the general proposition is to fix the vector of quantities between these two periods.

The question naturally arises: the quantity of which consumed goods should be recorded? It is clear that the number of products circulating in the consumer market is immeasurable, and it is not possible to observe absolutely all goods sold and bought by the population. In this regard, Scottish economist Joseph Lowe in 1823 proposed for the first time the use of the so-called fixed set of quantities (products), in his own words – "reference standard" or "reference table". This means that a certain list is deliberately formed within certain groups of products, the amount of which should be recorded between the two time periods selected for comparison. And it is for products from this list that regular price registration and the CPI calculation are carried out. Therefore, *the economic approach is based on the hypothesis of the existence of a representative consumer basket.*

However, in Lowe's price index, the fixation of quantities is not limited to a specific period. It is assumed that you can use any fixed set of quantities, even hypothetical, that do not apply to any actual period of time. Due to this, the Lowe index can be used not only for comparisons of price dynamics over time, but also for spatial comparisons. And it is due to this that the Lowe price index, as well as elementary (unweighted) indices, *well demonstrate the properties of transitivity and reversibility.* It should be noted that the Lowe indices are very widely used in economic statistics, and price indices are only part of the Lowe index family (Hill, 2010).

It should be understood, however, that the abstract nature of the quantitative vector in the Lowe index is not only its strength, but also its vulnerability. Weighting individual indices to values that do not reflect the real structure of consumption does not allow measuring the real dynamics of consumer prices. To solve this problem, German statisticians Etienne Laspeyres and Hermann Paasche in 1870 proposed indices named after them. The Laspeyres and Paasche indices are special cases of the Lowe indices and, like the Lowe indices, are not limited to price indices, but are widely used in various fields of statistics to measure evolution of price, production, turnover, GDP, and so on. Their difference from Lowe's indices is to specify the period of fixation of the vector, the impact of which must be neutralized (quantitative or price). Namely, in the Laspeyres index, the vector that is neutralized is fixed at the level of the previous period, and in the Paasche index - at the level of the current period.

The study of these indices allowed us to formulate a number of properties that are expected to demonstrate an ideal price index, namely: associativity, proportionality; invariance of the index in the case of proportional change of quantities; increase in the index for rising current prices; decrease of the index with the growth of the base price; the limits of the aggregate index, and the deviation between the Laspeyres and Paasche indices.

The deviation between the Laspeyres and Paasche indices is due to the negative covariance between them, which is known to indicate a systematic excess of the value of one variable over the value of another. That is, in the case of anticorrelation between price vectors and quantity vectors, the Laspeyres index somewhat overestimates inflation, while the Paasche index somewhat underestimates inflation. That is, for the same elementary indices, the value of the Laspeyres index will always exceed the value of the Paasche index while the Paasche index, on the other hand, does not systematically take into account the structure of consumption that existed in the previous (base) period.

The search for an ideal formula, free from the shortcomings of the Laspeyres and Paasche indices, gave impetus to the development of an axiomatic approach.

The axiomatic approach was formulated as independent by the British economist Alfred Marshall (1887), the American economist Korea Moylan Walsh (1921), the Irish economist Francis Edgeworth (1925) and the American economist Irwin Fischer (1930). Marshall proposed a modification of the Lowe index formula with a simple arithmetic mean of the two comparable periods, developed then in Edgeworth's work, which is why the corresponding index is named after two economists. A variant of this index, but with the geometric mean of the two periods compared, was proposed by Walsh.

The Marshall-Edgeworth and Walsh indices demonstrate the property of reversibility. However, they do not have the ability to circularity, although their deviation from the target value for this property is smaller than in the case of the Laspeyres and Paasche indices. The values of the Marshall-Edgeworth and Walsh indices are located between the values of the Laspeyres and Paasche indices, i.e. with the help of these indices it is possible to reduce the systematic errors in measuring inflation, characteristic of the Laspeyres and Paasche indices. In the development of the ideas of Marshall-Edgeworth and Walsh, Fischer proposed an index that is the geometric mean of the Laspeyres and Paasche indices.

To select the formula of the index-candidate an axiomatic approach offers twenty tests to check its ability to demonstrate both the properties formulated in the economic approach and additional ones. All tests are grouped into six groups: 1) general; 2) for homogeneity; 3) monotony; 4) symmetry; 5) the limits of the aggregate index; 6) tests of the index of dual quantities.

Among the proposed indices, the only one that demonstrates all twenty formulated properties is the Fisher's index, which is why it is called the superlative index.

Within the axiomatic approach, the concept of the index as a weighted geometric mean of price level ratios was also developed, in which it is proposed to use a symmetric mean cost share (relative weights) for the two periods as weights. In this approach, fourteen properties that the candidate indexes must satisfy are formulated, and two formulas are proposed. Finnish statistician and economist Leo Waldemar Törnqvist (1936) proposed to use the *arithmetic* mean of the relative weights of two periods, and Walsh (in the development of his previous index) - the *geometric* mean of the relative weights of two periods. These indices are based on two important hypotheses. The first hypothesis assumes the separation of the weighting system in the calculation of the index. That is, if, for example, the price of only one commodity has changed between two comparative periods, the costs incurred by the consumer for other commodities will not affect the aggregate index. The second hypothesis assumes the invariance of the index in relation to the dynamics of prices for goods that for some reason did not gain weight. That is, if the price of only one commodity has changed between two comparison periods, but the weight of this commodity is zero, the aggregate index will be equal to one (or 100 percent).

Of the two presented indices, the Törnqvist index is quite symmetrical and fully meets the stated requirements, which is why, like the Fisher index, it was included in the group of superlative indices. The Törnqvist index is a further significant contribution to the formation of the ideal CPI concept, namely in the separation of the weighing system, which significantly increased the possibility of various operations *in case of* *missing or incomplete primary data, as well as embedding other formulas for sub-indexes within the aggregate index.*

The results of calculations obtained using the Fisher and Törnqvist indices are quite similar in value. In some classifications, the Walsh index is also included in the group of superlative indices. The property of superlative indices is their ability to take into account both prices and quantities during defined periods. *Superlative indices are symmetric and provide a close approximation for theoretical price indices. The results these indices give are close in value.*

In general, the superlative index is defined in index number theory as equal to its own theoretical prototype (ideal index), defined for a particular functional form of its reference quantitative vector. Hence, it is an index that is "... a flexible functional form that can provide a second-order approximation to other twice-differentiable functions around the same point" (Diewert, 1976; Export and import price index manual, 2009).

The **stochastic approach** is based on the hypothesis that individual indices oscillate around a trend of the distribution of these indices. This idea was first proposed by the Italian economist Gian Rinaldo Carli in 1764, was deeply studied in the works of British economist William Stanley Jevons (1863), and finally substantiated by Edgeworth.

Within this concept, there are two approaches: 1) unweighted stochastic approach, and 2) weighted stochastic approach.

Within the *unweighted stochastic approach*, there are two types of indices - the Carly Index and the Jevons Index. The following hypothesis was formulated for the unweighted stochastic model:

- The lower limit of the price ratios is zero, and their distribution relative to zero is not symmetric.
- Model residuals (white noise) are independent, equally distributed, and centered random variables, hence, they are usually distributed around zero.

The aim of the model is to minimize the squared error of the residuals, which is achieved using the least squares method.

However, this approach was criticized by Keynes (1930), who questioned the existence of a general level of inflation, around which all price ratios randomly fluctuate, and thus the hypothesis of white noise, that is the independence of deviations between price ratios and overall inflation. Keynes noted that deviations between price ratios and overall inflation correlate with each other, and this correlation is largely due to the relative level of spending on consumed products (Diewert, 1995).

The weighted stochastic approach is also represented by two indices: the weighted version of the Carly index and the weighted version of the Jevons index.

For the weighted stochastic model, the hypothesis of heteroscedasticity (autocorrelation of residuals) was formulated, according to which price fluctuations of one commodity affect changes in the price of other commodities, as well as the assumption that weighted individual effects compensate each other, therefore there is a collinearity effect. This means that, unlike the previous model, the weighted model contains an individual trend, and inflation specific to a particular product is the sum of total inflation and individual effect.

In general, with the help of both stochastic approaches, an attempt was made to justify the calculation of the aggregate price index as a certain form of the average. The advantage of a *weighted linear estimate* is that it generally coincides with the Laspeyres index (unbiased estimate of the overall inflation rate), while the *weighted logarithmic estimate* coincides with the formula of the Törnqvist index. Both are convenient to use and to calculate confidence intervals.

However, the stochastic approach, like any other, has its weaknesses. For example, a stochastic weighted estimate assumes that the larger the share of the budget spent on a given product is, the less its relative prices are dispersed around the mean, although no empirical evidence has been obtained for it.

Promising for this method is the possibility not to track prices directly, because it is not about a set of affordable prices, it is about the simple two-stage random sample (period and costs).

The cost-of-living index (COLI) is directly derived from the theory of consumer behaviour and is a kind of economic approach. Within the concept of the cost of living, such indices are considered as:

- The true cost index or the index of constant utility, proposed by the Russian economist Alexander Konüs in 1924;
- The superlative cost-of-living index proposed by the American economist Walter Erwin Diewert in 1976;
- Dynamic cost-of-living index proposed by the Portuguese economist Ricardo A. M. R. Reis in 2009.

The initial hypothesis of this group of indices assumes that in the event of an increase in the price of a commodity, the consumer can replace it with another commodity: minimizing the value provided that a certain level of utility is achieved. That is, the cost-of-living index is based on consumer preferences in terms of iso-utility choice and reflects the evolution of costs in different price contexts for the same utility.

The basic formula ***of the true Konüs cost-of-living index*** correlates the function of consumer expenditure at one-year prices with the function of consumer expenditure at another year's prices, therefore it is about the cost of achieving a certain level of utility (or standard of living) in one year relative to the cost of achieving the same level in another year chosen for comparison.

The achieved level of utility reflects the quantitative needs of the consumer, so a variation of Konüs index is a formula in which the function of quantities is used instead of the utility function. This transformation brings Konüs cost-of-living index closer to the previously considered Laspeyres and Paasche indices (depending on the period of fixation of the quantity function).

However, the result of the calculation of the Konüs-Laspeyres cost-of-living index *is less* than that obtained using the classical Laspeyres index, while the result of the calculation of the Konüs-Paasche cost-of-living index *exceeds* the result of the classical Paasche index. Hence, it confirms the view that the Laspeyres index overestimates to some extent the cost of living and, accordingly, inflation, while the Paasche index does not overestimate the cost of living and inflation. The reason for this systematic error lies in the imperfection of the assessment of the substitution effect compared to the Konüs index (Manser and McDonald, 1988).

Dievert's perfect indices are an approximation of the indices of constant utility. These indices give good results for cases of homothetic consumer preferences. This is the case when the utility function has a constant elasticity of substitution. Ideal cost-of-living indices include modified Törnqvist indices and all indices derived from the quadratic utility approximation (primarily the Taylor and Fisher indices).

The advantage of ideal cost-of-living indices is their ease of use in the sense that they do not require an assessment of the parameters of the demand function, while the shortcomings of these indices are common to the shortcomings of the corresponding classical price indices, which complicates their practical use and will be reviewed further.

The problem of this group of indices is also ignoring the fact of the existence of different categories of households and different markets for the same product (points of sale and sales network). In order to eliminate this shortcoming, the cost-of-living index was further developed by introducing the so-called "democratic" and "plutocratic" weights into its formula.

"Democratic" weights are obtained by averaging the weighted percentage of expenditures for the whole population. "Plutocratic" weights are calculated by dividing total aggregate expenditures by population number and, accordingly, give more weight to high expenditure households.

Studies show that the democratic index better reflects the impact of inflation on the average household, although in practice, because in many countries the plutocratic index is used to deflate national accounts and to measure inflation, it is rarely used (Guide complémentaire, 2004).

The dynamic cost-of-living index measures the so-called dynamic inflation by extending the concept of the *constant* utility index to the idea *of a time-varying* utility function. In this model, uncertainty and expectations play a leading role, in contrast to the model with constant utility. This method is based, in particular, on the a priori stochastic evolution of product prices (for example, price logarithms correspond to first-order autoregressive processes, VAR models, etc.).

This approach has the advantage of comprehensively modelling forms of intertemporal substitutions in consumer preferences, but longer than monthly substitutions. Interesting results are provided by long-term panel surveys, which show significant changes in preferences depending on age, social status and living conditions, as well as the transfer of budget savings between generations: from older generations to younger (help of middle-aged parents to student children) and, conversely, from adult children to elderly parents (when children who have reached a certain social status and income level, in turn, provide financial assistance to parents). According to Reyes, this dynamic inflation rate takes into account the fact that the consumer lives a large number of elementary periods and throughout his life will optimally (from his point of view) distribute their own consumption (Reis, 2009).

In summary, we can say that the ideal concept is a scientific hypothesis about the possibility of measuring inflation in the consumer sector of the economy using the chosen formula, which must meet a number of established requirements. The implementation of the presented theoretical concept is possible under the condition of its certain approximation. Such approximation is the expression of declared but difficult to observe phenomena through other, more accessible (close in value) or simpler ones, which is common practice in economic research. The process of approximation of the ideal CPI occurs through finding the conditional concept of the CPI.

4. Conditional CPI concept

The conditional concept of the CPI is the transitional link that allows further empirical testing of the theoretical (ideal) concept, and determines the methodological principles of the index. Expression of the ideal concept through the conditional concept reflects the real conditions for the economic agents of each country and the primary information accessibility. Adjustments are made to the basic components of the CPI, such as:

- aggregate representative consumer;
- actual final consumption of households;
- geographical coverage of the index;
- consumer needs units;
- comparison period;
- where, when and how the price of consumer products (goods and services) should be recorded;
- the aggregate index formula and the sub-indices formula.

Precision of the concept of aggregate representative consumer is taking into account different geographical and consumer categories of the population: residents-nonresidents and urban-rural population. According to the ideal concept, the price index should cover the entire population of the country, which consists of such categories as residents and non-residents, as well as urban and rural population. But in practice, there are a number of reasons that complicate this task, both due to the limited budget of statistical observations and due to limited access to certain data sources, thus leading to a significant deviation from the ideal concept.

Given the existing budgetary and organizational constraints, the factors of choice between urban and rural populations are:

- the purpose for which the price index is expected to be used;
- population of each category;
- contribution of consumption (transactions) of each category in the total amount of consumption;
- geographical accessibility for statistical observations;
- evolution of consumption structure and the ability to obtain data to build the weight structure of the index for each category.

The first parameter that outlines *the conditional limits of the* aggregate representative consumer is the choice between:

1) a resident, regardless of whether he carries out his consumption within the country or also abroad;

and, at the same time,

2) a resident and a non-resident, but under the condition of their consumption exclusively within the territorial boundaries of the country.

The need for such a choice is due to two different approaches to determining the national product in national accounting. The first approach is designed to calculate gross national product (GNP), and the second - to calculate gross domestic product (GDP). However, it should be noted that both approaches in their unchanged form are quite difficult to implement in consumer price statistics. In the first approach (GNP method) it is almost impossible to obtain information on consumption and prices of products of residents living outside the country. In addition, the consumption of residents abroad reflects the inflationary processes of the economy of their current stay. The second approach (GDP method) also makes it difficult to monitor on an ongoing basis the consumption patterns of non-residents living within the country: such households are volatile and statistical monitoring often requires appropriate legislation. Non-residents may also have an income level that deviates significantly from the average income level of residents, which introduces a systematic error in the weight structure. In addition, the share of non-resident households in the total number of households in the country is usually very small, so the consumption inherent to them may not be considered as typical for a given economic area.

The choice between these methods is a very important issue in today's conditions when a large number of people temporarily left Ukraine because of the war. Classically, residents who consume within the territorial boundaries of the country are the only appropriate conditional definition of the aggregate representative consumer. But should the CPI of Poland, for example, take into account the consumption of millions displaced Ukrainians? And does the current CPI of Ukraine take into account the real structure of household consumption, when households are largely truncated because of population migration?

The second parameter of the formation of the *conditional boundaries of the* aggregate representative consumer is the representation of urban and/or rural population in its structure. In favour of the rural population there is the fact that in the case of using the index to analyze poverty, the registration of prices in rural areas is more informative than in cities. In addition, in the vast majority of countries, the rural population significantly outnumbers the urban population. At the same time, the volume of money transactions for the purchase of goods and services in rural areas is smaller compared to urban agglomerations, hence for the price index the total weight of urban consumption is more significant than in rural areas, while the cost of registering prices in rural areas due to territorial dispersion, and often limited availability, is significantly higher.

The decisive factor in choosing the coverage of the population for the CPI is the ability to obtain and timely update the weights to build an aggregate index. Changes in the structure of urban consumption happen faster than in rural areas, both due to greater availability of new goods and services in the city and due to a more dynamic and less traditional way of life. And because scales for rural households are rarely available for measurement, the ultimate preference is often given to the urban population. The last factor is especially important in times of war when rural zone accessibility is restricted for statisticians.

Thus, in the conditional CPI concept, the definition of the aggregate representative consumer is much narrower than in the ideal concept, and in the vast majority of cases it covers *urban households that consume within a specific economic jurisdiction (country).*

The definition *of the actual final consumption of households in the conditional index* directly follows from the definition of the aggregate representative consumer as consumption within a certain territory. In addition, it is also determined by the time factor and the chosen method of estimating consumption costs, namely – the choice between methods: purchase, use, and payment.

In the case *of purchase*, it means the *moment of purchase of* goods or services, regardless of when they are actually used or consumed. The moment of purchase of goods is the moment of transfer from the seller to the buyer of ownership of the goods or services. In the case *of use*, it is the *period during which the product is consumed or used many times.* Payment means *the period of time of actual payment for a product,* which may differ from the period of purchase or use.

The choice of method depends, on the one hand, on the structure of consumption, and, on the other hand, on the development of banking services and payment system in trade and services. The first means that the share in the structure of consumption of fast moving, semi-durable and durable consumer goods and services varies from one country to another. In poorer countries, the share of fast moving consumer goods is much higher than in developed countries, while the latter are characterized by significant volumes of durable consumer goods and services. On the other hand, in developed countries there is a widespread practice of seasonal sales and discounts, when goods are sold cheaper than usual, and the consumer tends to buy on this occasion bigger quantity of goods than she currently needs to consume or use them future periods. Economically advanced countries also have a developed and extensive banking system and offer financial services in the form of bank checks, credit purchases, online payments and remittances, which are often prevalent in trade as opposed to cash. E-commerce became especially widespread during the COVID-19 pandemic. The preference for non-cash payments was also given by the banking system of Ukraine in the first months of the war, which allowed reducing the outflow of large money from the bank accounts of consumers. E-commerce in Ukraine is facilitated also by the well-developed shipping business, which is flexible according to the need of the customers, even during wartime.

A new trend in the behaviour of Western consumers, which is getting more widespread, is also the purchase of goods on individual orders in accordance with the proposed samples. This is a deliberate deferred consumption (sometimes up to 2 months) of the ordered goods, which are not mass-produced, but only on individual orders. The goods bought in this way are of course more expensive, but are more relevant in terms of compliance with modern fashion trends and personal preferences. This approach also suits both the producer and the seller, because it allows not to invest in mass production of products that often do not find their consumer, settle in warehouses and eventually gets destroyed or deposited in landfills of poor countries, which entails losses for the owner, and environmental pollution.

This means that in countries with a less modern system of financial services, less e-commerce and a more traditional style of consumption, the time lag between the moment of payment, purchase and consumption can be either insignificant or non-existent, and accordingly, the three methods of determining the actual household final consumption considered to produce similar results. For this reason, according to the CPI Manual, the vast majority of countries use the purchasing method to determine consumer spending (Consumer price index manual, 2020). However, the rapid evolution of trade methods requires a revision of this approach, and many statistical services make some adjustments to determine the actual final consumption. This adjustment allows the direct use of big data on real-time transactions made by the population to purchase and pay for consumer goods and services, which reduces the restrictions on obtaining primary data imposed by the war conditions.

By general agreement, there are *units of homogeneous needs and units of heterogeneous needs*. Units of homogeneous needs cannot be replaced. Units of heterogeneous needs are interchangeable with a substitution elasticity equal to one. Units of needs are defined at the level of the elementary market, which corresponds to the market of a particular type of product or service for a particular area. Outside the specified elementary market, the units of needs cannot be replaced. Together with the "purchase" method, this satisfies the requirement of an ideal concept such as tracking prices at the place of purchase of goods and services, not at the place of residence of the consumer. The type of product or service corresponds to a variety of products. In the case of goods or services purchased over the Internet, only their parameters are taken into account to determine their basic market.

The specification of the elementary comparison periods depends on such factors as *cyclical consumption* and *price volatility* for a particular product. It is necessary to distinguish between the period of comparison and the frequency of registration of prices.

The production and corresponding renewal of stocks of goods in sales outlets largely depend on how often the consumer needs to purchase them. For example, fresh food can be renewed in stores on daily basis, two or three times a week, weekly or decadal, which, in turn, may be accompanied by changes in price and range.

Price volatility also depends on the type of product and the system of its commercialization. For example, ticket prices for the same flight of the same airline can fluctuate significantly each time a consumer makes an online request. The greater the volatility in the price level of a particular product, the more frequent is the need for its registration, which is especially relevant in wartime, when the established periodicity of price registration can be radically revised.

The months of the year were chosen as the conditional CPI concept as the comparison periods. The month corresponds to the smallest standard business cycle, which covers production-sales-consumption of most products. During the year, each current month (month of observation) is compared directly with December of the previous year. Every two of such months *are considered to be well comparable.* The argument in favour of December is the existence of a pronounced annual (calendar) periodicity of changes in price levels. December is the final month of the year and at the same time a transition between two consecutive years. January marks not only the beginning of the year, but also the beginning of economic recovery, which is systematically accompanied by new orders and rising consumer prices. All other periods *are considered hardly comparable.* In the event of an emergency situation not at the beginning of the calendar year, but later, the question arises of the comparability of the periods that actually characterize such a situation and the periods before its occurrence.

The properties of transitivity and reversibility are well demonstrated by elementary, that is unweighted, indices. However, as we know, the CPI should take into account not only the changes in the level of prices of consumed products, but also the structure of final consumption, so the aggregate CPI cannot be calculated as a simple average value from elementary indices, but only as a weighted average. For each well-comparable period, the system of consumer expenditure weights of units of needs is based on the consumer expenditures of households registered for the previous year as a whole. At the same time, in the conditions of high instability of the structure of household consumption during the war, caused primarily by the migration of many households from dangerous regions and the formation of the so-called "truncated" households, the system of weights needs to be significantly updated.

Thus, according to the conditional concept, the CPI is an indicator of the monthly average change in the level of prices paid by urban consumers for the market basket of goods and services (Konny et al., 2019), which, however, must be substantially clarified in the light of the extreme conditions of war.

The choice of formula for the conditional CPI concept depends on a number of factors and is related to the chosen approach and practicality. Conceptualization of a common approach to a possible assessment of consumption inflation, namely *a choice between CPI and COLI*, depends on statistical tradition: American (USA and Canada) and European. The American school considers COLI as an ideal target metrics, while the European school prefers the CPI. Accordingly, in the American school the conditional concept of the index is substantiated within the limits of COLI, and in the European school - within the CPI. Hence, this is the justification of a particular formula of the aggregate index and, if necessary, sub-indices.

Among those considered within the ideal concept, the basic formulas for the aggregate index are: Laspeyres, Paasche, Marshall-Edgeworth and Walsh, Fischer, Törnqvist. In terms of the ideal concept and in accordance with the requirements formulated within the economic, axiomatic and stochastic approaches, the best properties are shown by the Törnqvist index (Table 1) (Chauvet-Peyrard, 2014).

Table 1. Estimation of theoretical properties of indices for aggregate CPI estimation
(3 points - the best score, 2 points - the average score, 1 point - the worst score, x *- is not considered within the specified method)*

Approaches	Aggregate index formulas					
	Laspeyres	Paashe	Walsh	Jevons	Fisher	Törnqvist
Economic	1	1	2	1	2	3
Axiomatic	1	1	2	2	3	3
Stochastic	1	1	1	2	x	3

Fisher's index is also effective from the point of view of the axiomatic approach and somewhat worse from the point of view of the economic approach. The Walsh and Jevons indices also meet the requirements of the axiomatic approach quite well, although the Walsh index has better properties from the economic point of view and the Jevons index from the point of view of the stochastic approach. It is interesting to note that the Laspeyres and Paasche indices have the lowest score for all three approaches. According to the results of the generalization presented above, if we proceed solely from the theoretical properties of all these indices, preference should certainly be given to the Törnqvist and Fisher indices. However, numerous empirical studies indicate the existence of a number of limitations in their practical application. In fact, only the Laspeyres index and to some extent Jevons are possible for practical application (Table 2).

Table 2. Evaluation of practical properties of indices for aggregate CPI evaluation
(3 points - the best score, 2 points - the average score, 1 point - the worst score, x *- not possible for practical use)*

Aggregate index formulas					
Laspeyres	Paashe	Walsh	Jevons	Fisher	Törnqvist
3 points	x	x	1 point	x	x

It should be considered, however, that even the Laspeyres index in its theoretical form cannot be directly used in statistical practice, but requires some modifications. Practical use of index formulas is influenced by such factor as a*ccessibility and timeliness of inflow of primary data necessary for CPI compilation on a regular basis*. Since, according to all the above-mentioned formulas, the aggregate CPI is an averaged value of various kind, its calculation requires primarily two types of primary data: the structure of consumption of goods and services and registration of price levels. The availability of these data is a decisive factor in developing a conditional CPI concept. Unlike all the considered indices, the formula of the Laspeyres index assumes the use of data that can (with some conditions) be obtained at the time of the operational calculation of the index by official statistics.

Thus, in practice, the processing of data on household expenditures surveys takes a long time, so there is a lag between the period of the survey of unit costs and their first use to compile the index. This, of course, makes it impossible to use the Laspeyres formula for real-time CPI compilation. Therefore, the share of costs obtained for the previous period can be used to weight price changes provided by the transformation of the classic Laspeyres formula into the formula proposed by Jung in 1812 [Consumer price index manual, 2004], in which the period of the weight structure of the index is independent of the comparison periods. The Laspeyres index advantage is also in the fact that the formula allows for an easy solution to the task of inclusion to the common system of calculation of other formulas for the lowest levels of aggregation. It is not possible to obtain a weight structure at these levels, so the formulas of Carly, Jevons and Dutot are the basis for elementary sub-indexes. The Carly index is the arithmetic average of individual indices, while the Jevons index is the ratio of the geometric average of individual price levels, and the Dutot index is the ratio of the arithmetic average of individual price levels.

Practical implementation of the conditional CPI concept

Practical implementation is the application of a random variable arising from previous assumptions, by organizing statistical observation. The practical implementation of the conditional concept includes numerous stages that generally correspond to those given in international recommendations and methodological materials of the State Statistics Service of Ukraine (Methodological provisions. State Statistics Service of Ukraine. 2020):

In the context of the transformation of the ideal concept (through the stage of the conditional concept) to its practical implementation, it is important to take into account the impossibility of direct use in statistical practice not only the ideal but also the conditional concept. That is, even the conditional CPI concept needs some adaptation in order to be able to collect data and calculate the index on a regular basis, while ensuring the continuity of the time series. This is mainly due to the fact that consumer needs, on the one hand, and the market supply of goods and services, on the other hand, are not fixed in time. Changes in the economic conditions of life, the impact of the technical progress and fashion push the consumers to seek new degrees of usefulness, which naturally entails the emergence of new and disappearance of obsolete, morally and technologically, products. In turn, the manufacturers are in constant pursuit of models and developments of new competitive goods and services that they could offer to the consumers and thus encourage them to regularly review their preferences in favour of more expensive goods and services. This circle is thus an upward spiral, in which the renewal of supply is accompanied by changes in product quality and rising consumer prices.

To this is added the phenomenon of seasonality of goods, which periodically partially or completely disappear from the market, and appearing next season, have, as a rule, new characteristics and higher prices. All this requires a corresponding adjustment of the conditional index formula.

There are numerous techniques for taking into account the disappearance of old and the emergence of new products and, accordingly, taking into account changes in product quality. An important point in terms of the adaptation of the conditional concept is that their application requires the separation of calculations for the respective product groups using a whole arsenal of models and formulas. As an example, we can cite the approach of *option pricing and production cost*, as well as *hedonic regression*, which are widely used in the practice by leading national services [Chauvet-Peyrard, 2014]. In the case of the option pricing and production cost approach, the difference in price can be taken into account due to the difference in the total price of options that complete the finished product, i.e. the price is formed as the sum of base cost and added options cost. A striking example of the option method is the price of cars, which can vary significantly within one model, depending on the options chosen by the buyer.

In turn, the hedonic approach assumes the existence of an implicit market for all possible product characteristics, and the product itself can be likened to a certain vector of characteristics, i.e. it is believed that there is a real product for any possible combination of characteristics under conditions of perfect competition. It is assumed that the consumer maximizes his own utility in the face of budget constraints, so that the marginal utility provided by each characteristic should be equal to its marginal cost. Hedonic regression of quality adjustment is often used in the case of monitoring the prices of high-tech goods and services. The products for which this approach is considered appropriate are those that are sold in highly competitive and dynamic markets and, accordingly, change rapidly, but at the same time, they are not difficult to track on a regular basis. These include products such as computers, computer software and IT tools, as well as Internet services and more.

The attempt to improve the quality of the aggregate index through the use of numerous additional adjustments based on inclusions in the basic formula gives a compilation of the practical concept of the index. This, in turn, significantly distances practical concept from the conditional concept, which is a potential source of systematic criticism. Is it possible to resolve this contradiction? And how can this be taken into account when constructing a consumer price index in the current extreme situation?

As the experience of recent years shows, great prospects in this sense exist in the development of fundamentally new sources of information and in new technologies for primary data registration, processing and storage.

New sources of information and prospects

Today, in statistical practice, such sources of primary data are traditionally used to build the CPI, such as:

1. Household budget survey;
2. Price collection survey: price registration is carried out by price collectors from different outlets and often with use of electronic devices.
3. Data of websites where relevant goods or services are sold: data are registered either by statisticians manually or with the help of robots (web scraping) and special mobile applications. Some institutions provide statistical services with their own Application Programming Interface (API) for direct access to pricing information. Collecting data with the API is often simpler and clearer than long-term support for web scraping code.
4. Administrative data: files of administrative records on prices and levels of consumption of certain, primarily social services and goods.

All of these sources have both strengths and weaknesses. The main problem is the growing distance of the conditional concept from its ideal prototype. First, because the CPI aims to measure the constant quality of price change, over time, when a particular product disappears from sale, it is necessary to choose a substitute product, and any change in quality between the original and the substitute product must be evaluated and eliminated to display the so-called “net” price change. Second, new products entering the market must be accounted for in a timely manner with an appropriate economic weight for which data are not available. Third, the CPI is based on samples that may contain systematic error. Finally, traditional data sources often provide only supply prices and not actual consumption prices, as they do not reflect the practice of discounts, which is very common and takes various forms in the retail system.

Alternative sources that different countries have been developing recently are:

1. Third-party data are data collected by a third party that contains prices for goods or services obtained from several sources. The third party is the so-called aggregator, which standardizes the elements and structure of data obtained from various commercial institutions. Such datasets are usually obtained by official statistics on a paid basis, or, in some cases, free of charge.
2. Corporate data (that is that owned by corporations or companies) is a set of data obtained directly from the headquarters of a company that collects data in its own outlets or collects it on sales websites. Because such datasets are usually created for internal use, it is the owner who determines the architecture of this data. The information provided to statistical agencies is that the company is willing to provide, i.e. may vary in structure, so statisticians should negotiate with data owners on the optimal level of detail to ensure the confidentiality of personal records. Adapting the obtained corporate data to statistical needs is one of the key challenges when using such sources.

The most promising source of information on the evolution of consumer prices in corporate data is direct *cash data of sales outlets (data of cash registers)*. Cash register data refer to transactions that reflect both the price paid for each unit of goods and the number of units of goods purchased. These data are automatically registered and collected by retailers when the buyer goes through the cash register and pays for the goods purchased by him.

Cash register data have long been used in marketing research. But in recent years, leading statistical agencies have also initiated first pilot and later regular collection of cash data for statistical purposes, namely to build the CPI. Among the pioneers in this field are the national statistical agencies of the Netherlands (the first CPI publication calculated using cash data was published in 2002), Norway (2005), Switzerland (2008), Sweden (2012) and Belgium (2015), Denmark (2013), Iceland (2016), Luxembourg, Italy (2018), France (2020) and the United States (2021). A significant methodological contribution to the cash data adaptation for statistical purposes was made by Eurostat (2017).

Due to certain difficulties, the use of cash register data to build national CPI is currently limited to certain items of goods sold in supermarket and hypermarket chains. The use of cash data implies the presence on each product of a global item number (Global Trade Item Number - GTIN), which is a 14-digit international product number. Thanks to the GTIN, it is possible to scan and computer data on each product throughout the commercialization chain. These codes are an internationally recognized system and are supported by the international organization GS1.

The main motive that prompted the development of cash data as a new resource of statistical information was the need to restructure official statistics in accordance with the requirements of the big data era (because cash data are a kind of big data). In addition, it is also obvious that it is expedient to use this type of corporate data for statistical purposes in conditions of limited access to traditional sources of information, which is especially relevant in the context of Russian military aggression in Ukraine.

From a financial point of view, this resource is to some extent "free" for statistics, as the production of cash data is outside the statistical process, even if their processing for statistical use requires significant investment.

Also, the obvious advantage for statisticians is the comprehensive nature of cash data, real-time data flow and the availability of previously unavailable information, namely the number of units consumed and the corresponding household expenditures, which certainly opens up many opportunities for price statistics (Leclair, 2019).

A fundamental component of the price index number theory and, accordingly, the ideal CPI concept is the availability of real-time information on household expenditures. Until recently, such information was not available, so theoretically perfect formulas, such as superlative indices, could not be used in statistical practice.

Access to cash data allows not only to improve the CPI estimates, but also to bring the aggregate index formula closer to its ideal concept (Konny et al., 2019).

The comprehensive nature of cash data allows for more accurate statistics and the possibility of creating more detailed versions of consumer price indices, for example, by specific segments of consumption. Experimental studies have shown that cash data can also be used successfully for spatial price comparisons (Léonard et al., 2019).

We should also take into account the current trend towards branching out and consolidating the network of sales outlets of various sizes in both urban and rural areas. In many cities, the so-called "neighbourhood" trade is becoming more widespread, offering goods in small supermarkets in urban areas, but not on the outskirts of cities. On the other hand, modern residents of small settlements, traditionally not covered by the CPI survey, are increasingly buying consumer goods in large supermarkets located in their region, which are becoming the norm in rural areas. The possibility of directly obtaining cash data from supermarkets allows at the same time to better cover the urban population and include data on rural consumption, thereby significantly increasing the level of representativeness of the CPI, both nationally and regionally. This, in turn, brings the practical concept of the aggregate consumer closer to its ideal vision.

Cash data allow us also to quickly identify new products that need to be added to the CPI basket, or products that are obsolete and that should be removed in order to update and match the basket to the actual structure of household consumption. The same applies to the timely recording of changes in the structure of household consumption caused by mass migratory flows of the population during the war.

Thanks to detailed cash data, it is also possible to choose index formulas that can take into account the lowest level of substitution aggregation carried out by the consumer due to rising prices during the two observation periods (Leclair, 2019).

In general, *cash data better control the concept of price to be measured by the CPI than traditional methods,* as the prices reflected on price tags may differ from the prices actually paid by the consumer through various promotions.

It should also be kept in mind that the detection and replacement of products that appear and disappear during the year is an operation that requires significant human resources. The continuous nature of cash data allows one to automatically search for the previous price of the required product in the data accumulated over previous periods, which is also highly relevant in extreme conditions, when supply and price levels are very volatile.

Today, there are two common approaches to the use of cash data: sampling data collection and exhaustive data collection.

The sampling data collection is the most common among countries that already use cash data to build the CPI. Cash data provide an opportunity to form a qualitative basis for the sample, namely - the list and parameters of all goods sold in sales outlets, with the weight of each of these items in the turnover of sales outlets, which allows random selection of products, on the one hand, and on the other hand, to control the sample bias.

The exhaustive data collection has been developed and implemented in Danish statistical practice since 2013. The use of the exhaustive data collection was primarily aimed at reducing differences in the approaches used by statisticians for different sales outlets and consumer goods within the Danish CPI. That is, it is possible to apply a generalized approach and an index formula to a significant number of goods and services, which significantly reduces the compilation nature of the practical CPI formula mentioned above. Today, this approach is still used for uniformly defined groups of goods. According to the proposed methodology, the price index is calculated as the ratio of the turnover index and the weighted quantity index, i.e. it is the Geary-Khamis method applied to time series. At the same time, the quantitative scales for homogeneous goods are updated monthly during the current year, based on the prices and quantities of sales registered by sales outlets, which is highly relevant in wartime. The advantage of this method is that it does not cause bias in chain indices, because at the end of each year such price indices coincide with transitive indices (Chessa, 2016).

The challenge of working with large cash data is their incomparably greater volume than traditional data. For example, in France it is 1.7 billion monthly records received in real time [Leclair, 2019]. The procedures for their transmission through communication channels, processing, storage and protection differ significantly from those provided for traditional statistical data sets. According to the international experience gained in recent years, a prerequisite for the successful involvement of cash data in the CPI calculation covers a number of issues (Leclair, 2019), namely:

- evaluation and analysis of new sources of information;
- legal support of the right of statistical access to them;
- form of data access (free / paid), budget and frequency;
- data reliability;
- methodological support;
- and technological capability of national statistical services to gather, treat and storage cash data.

5. Conclusion

The current challenges in the field of price statistics are both new and ancient, i.e. related to the conceptual framework of the CPI. The natural transition from the theoretical concept of the index to its conditional concept and then to its practical implementation has led to a significant deviation of the aggregate formula from the theoretical foundations. This contradiction can be resolved by revising the paradigm of sources of statistical information for the compilation of the CPI. The use of big data, especially direct cash data, has the potential to reduce the cost of statistical surveys, expand the sample size and improve its design while reducing the burden on respondents, obtaining more reliable data on transaction prices and comprehensively improving the price index by including real-time information on household expenditures (Konny et al., 2019).

The revision of the paradigm of primary data sources allows for a significant reduction in the methodological and organizational limitations imposed by the extreme conditions of Russia's military aggression against Ukraine. In the conditions caused by the war, this kind of information allows regular estimates of the consumer price index for a large number of goods without the loss of quality, and control the structure of consumption both in general and by region, and opens prospects for reducing discrepancies between conventional concept of the CPI, its ideal concepts and their practical application.

References

Balk, B. M., (2004). Direct and chained indices: a review of two paradigms. URL: https://www.researchgate.net/publication/228443023_Direct_and_chained_indices_a_review_of_two_paradigms.

Balk, B. M., (2018). Mixed-form indices: a study of their properties. URL: https://ec.europa.eu/eurostat/cros/system/files/euronaissue1-2018-article3.pdf.

Chauvet-Peyrard, A., (2013). Les indices de prix. De la théorie à la pratique. URL: https://dokumen.tips/documents/cours-les-indices-de-prix-de-la-theorie-a-la-pratique-acp-version-publique-2013.html.

Chessa, A. G., (2016). A new methodology for processing scanner data in the Dutch CPI. URL: https://ec.europa.eu/eurostat/cros/system/files/euronaissue1-2016-art2.pdf.

Consumer price index manual. Concepts and Methods – 2020. International Labour Office. International Monetary Fund. Organisation for Economic Co-operation and Development. European Union. United Nations. World Bank.

Diewert, W. E., (1995). On the Stochastic Approach to Index Numbers. Document of discussion, no DP 95-31, University of British Columbia.

Diewert, W. E., Hill, R. J., (2010). Chapter 12 Alternative approaches to index number theory. Price and productivity measurement: Volume 6 – Index Number Theory. URL: http://www.indexmeasures.ca/Vol6_10,09,26.pdf.

Diewert, W. E., (1995). Axiomatic and economic approaches to elementary price indexes. *NBER Working Paper*, No. w5104.

Diewert, W. E., (1976). Exact and superlative index numbers. *Journal of Econometrics,* 1976, vol. 4, issue 2, pp. 115–145.

Export and import price index manual: theory and practice, (2009). International Monetary Fund, Washington, D.C.

Guide complémentaire du Manuel de l'indice des prix à la consummation, OIT, (2004). URL: http://www.imf.org/external/data.htm#guide.

Hill, T. P., (2010). Lowe indices. *Price and productivity measurement*, Vol. 6, pp. 197–216.

Hulten, Ch. R., (2008). Divisia index. The New Palgrave Dictionary of Economics, 2nd ed.

Konny, C. G., Williams, B. K., Friedman, D. M., (2019). Big Data in the U.S. Consumer Price Index: Experiences & Plans. URL: https://www.nber.org/books-and-chapters/big-data-twenty-first-century-economic-statistics/big-data-us-consumer-price-index-experiences-and-plans.

Leclair, M. et al., (2019). Utiliser les données de caisses pour le calcul de l'indice des prix à la consummation. *Courrier des statistiques*, 3, pp. 61–75.

Léonard, I., Sillard, P. & Varlet, G., (2019). Écarts spatiaux de prix dans l'alimentaire avec les données de caisse, Economie et Statistique/*Economics and Statitistics*, 509, pp. 73–87.

Manser, M. E., McDonald, R. J, (1988). An Analysis of Substitution Bias in Measuring Inflation, 1959-85, *Econometrica*, Vol. 56(4), pp. 909–930.

Reis, R., (2009). A dynamic measure of inflation. NBER. URL: https://www.nber.org/papers/w11746.

Sillard, P., (2017). Indices de prix à la consommation. Insee, Document de travail, N°F1706.

State Statistics Service of Ukraine, (2016). Metodolohichni polozhennia shchodo orhanizatsii statystychnoho sposterezhennia za zminamy tsin (taryfiv) na spozhyvchi tovary (posluhy) I rozrakhunkiv indeksiv spozhyvchykh tsin [Methodological guidance on organization of statistical observation of change in prices (tariffs) on consumer goods (services) and calculation of consumer prices indices]. URL: http://ukrstat.gov.ua/norm_doc/2021/310/310.pdf [in Ukrainian].

Stoevska, V., (2018). Official consumer price indices – historical perspective, ILO Department of Statistics. URL: https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/ILO.pdf.

The World Bank Consumer price index manual, (2004). Theory and practice, Geneva, International Labour Office.

Von Auer, L., (2004). Economic monotonicity of price index formulas. Otto-*von*-Guericke-Universität Magdeburg. URL: http://www.ipeer.ca/papers/Auer,Nov.04,RevisedCommentSSHRC46.pdf.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 131–144, DOI 10.59170/stattrans-2023-008
Received – 06.07.2022; accepted – 28.11.2022

Approach to population estimation in Ukraine using mobile operators' data

Volodymyr Sarioglo[1], **Maryna Ogay**[2]

ABSTRACT

Estimating the size and places of residence of the population of Ukraine has been this country's problem for the past decade, and is related to the lack of census data for the 2010 round, large-scale processes of external and internal labour migration, and Russia's armed aggression against Ukraine that started in 2014. This already disadvantageous situation has been significantly exacerbated by Russia's full-scale war against Ukraine which began on 24th February 2022. Conducting statistical surveys, especially surveys regarding the population, turned out to impossible under war circumstances. Therefore, the task of developing effective approaches to estimating the population size using data from existing sources, in particular the data of mobile operators regarding the number, location and mobility of subscribers, has become even more pressing. The article highlights the results of a study on the use of data from mobile operators, data from administrative registers, and the results of a special population sample survey on the use of mobile communication for the purpose of estimating the population. It also provides the results of experimental calculations of the population size in Ukraine as a whole and in particular regions. The study moreover showed that the size of Ukraine's population in November 2019, unlike the official estimate of 41,940.7 thousand people, was probably about 37,289.4 thousand people.
The developed approaches can be used to estimate the number and location of the population of Ukraine during the intercensal period or significant population movements due to environmental disasters or military conflicts.

Key words: population estimation, mobile operators, mobile subscribers, administrative data, sample survey.

JEL Classification: C82, C83, J10

---

[1] Ptoukha Institute for Demography and Social Studies of the National Academy of Science of Ukraine, Ukraine. E-mail: sarioglo@ukr.net.net. ORCID: https://orcid.org/0000-0003-4381-9633.

[2] Ptoukha Institute for Demography and Social Studies of the National Academy of Science of Ukraine, Ukraine. E-mail: ogay@ukr.net. ORCID: https://orcid.org/0000-0001-9292-5710.

© Volodymyr Sarioglo, Maryna Ogay. Article available under the CC BY-SA 4.0 licence

Problem statement

In independent Ukraine, the census was conducted only once, in 2001. Accordingly, in recent years, in the implementation of the census of large-scale processes of virtually uncontrolled labour migration, lack of reliable data on migration within the country, official estimates of the population as a whole and by individual territories and settlements have long been criticized by specialists. The situation was significantly complicated by the impact of Russian aggression in the East of Ukraine in 2014 and became one of the most relevant for the country in the full-scale war of 2022, when a significant number of people were forced to seek new places to live in other regions of Ukraine or abroad, or they were forcibly moved to the occupying country or found themselves in the temporarily occupied territories. Human losses due to the death of military and civilians should also be taken into account. All this has exacerbated the issue of finding and applying approaches to estimating and monitoring the population based on existing data sources in Ukraine without conducting a traditional census.

In our opinion, for the development of approaches to monitoring the population of Ukraine in the current conditions, the national experience of 2019-2020 is very useful to clarify the population of Ukraine based on available information from various sources, including data from mobile operators (hereinafter – DMO). Relevant surveys were conducted by the State Statistics Service of Ukraine (hereinafter – the State Statistics), Institute of Demography and Social Research named after M. Ptukha of the National Academy of Sciences of Ukraine, members of a working group from other organizations, at the initiative of Minister of Ukraine D. Dubilet.

The aim of the study is to highlight approaches to the operational assessment of the current population of Ukraine and its regions based on data from administrative registers and mobile operators.

In preparing the article, methods of statistical data processing from different sources, methods of data aggregation, methods of evaluation of indicators based on the results of sample surveys, etc., were used.

Overview of existing research and experience. According to the international experience, traditional data collection methods, such as censuses, cannot meet the needs of users for prompt and timely information when it is necessary to estimate the population living in certain areas in the event of natural disasters, armed conflicts or epidemics. In fact, censuses conducted once every ten or five years have not been considered satisfactory to users lately, even under normal conditions, without cataclysms. This is explained by the fact that the availability of reliable and up-to-date information on the population is now an increasingly important prerequisite for the development and implementation of effective and efficient measures of socio-economic policy at all levels of administrative-territorial division. Under such conditions, more and more attention of researchers, primarily statisticians and data scientists, is paid to the development of methodological approaches and applied tools for estimating population, demographic phenomena and processes based on available data from sources other than traditional censuses. An important result of these studies is also the possibility of a radical modernization of the methodology and tools of censuses. A promising process in this regard in many countries is the use of data from administrative registers, which contain information about the population as a whole, its individual groups, households, etc., as well as DMO on the location and/or movement of subscribers.

The article focuses on approaches to estimating the number and location of the population based on data on mobile network subscribers and data from some other sources. Mobile network or DMO data is the data collected by mobile network operators as a product of the operation of a mobile network. DMOs consist of two types of data: communication data – mainly calls and text messages; and positioning data – determination and fixation by mobile operators of the approximate location of the device based on the identification of the antenna used for communication (see, for example, Handbook on the use of Mobile Phone data for Official Statistics (2017)).

Data of mobile operators are increasingly used in statistics to estimate population movements (labour migration, pendulum migration, tourism, number of internally displaced persons, etc.), population density, population of certain territories, economic activity, etc. The main advantages of DMO are their efficiency (data are recorded on a regular basis and can be accessed with a slight delay, approximately a week, due to the need to prepare and verify data by a mobile operator) and completeness (almost all data are recorded and stored for a long time).

As with the problems of estimating the number and density of the population on the basis of DMO, their main disadvantages are: the presence of several mobile operators in the country, which are commercial organizations, have different databases, heterogeneous coverage areas and with whom it is necessary to agree on data format and content; inaccuracy in determining the positioning of mobile devices for small areas, in particular for small settlements; the need for special control and processing of data by mobile operators for the purposes of use in population estimates to exclude cases of double counting, understatement or other errors; shifting population estimates when using DMO due to the fact that one person may own several phones or SIM-cards, which leads to overestimation, while a person may not use a mobile phone at all, which will lead to underestimation for certain social groups; one or more operators may refuse to provide mobile data, then the coverage of the population under study may be incomplete; the possibility of atypical behaviour of mobile phone owners (phone owners may be temporarily in a certain area in connection with a business trip, on vacation, in the hospital, etc.), which leads to systematic errors in assessments.

With reference to these shortcomings of the DMO as a source of data in estimating the population, their use implies a mandatory "link" to the existing estimates of population size or density based on special statistical models at the initial stage of implementation of the relevant tools. Such models should be developed or improved, first of all, based on the results of population censuses. Under this approach, the model is developed according to the census, and then the estimates are carried out with the necessary frequency according to the current DMO and the model. Under other conditions, models for estimating the population can be developed on the basis of data from relevant administrative registers, which contain up-to-date, reliable and complete data on individual populations, such as children of a certain age, persons of retirement age, etc. It is advisable to have information on the use of mobile devices by the population.

A wide range of studies demonstrate the potential of using DMO as a source of data in the following main areas: estimating the spatio-temporal distribution of the population within certain areas, given, for example, in Deville P., Linarde C., Martine S., and others (2014); estimates of the extent of movement of people within individual settlements at different times of the day or different days of the week and holidays covered in particular in the publications Järv O., Ahas R., Saluveer E., and others (2012), Ratti C., Pulselli R. M., Williams S., and Frenchman D. (2006), Reades J., Calabrese F., Ratti C. (2009); measuring the extent of labour migration (see, for example, De Meersman F., Seynaeve G., Debusschere M., and others (2016); real-time estimates of the number of people during large-scale events (such as a concert) or at the scene of an emergency covered, in particular, in the publication Lu H., Frauendorfer D., Rabbi M., and Mast M. (2012), etc.

The study Deville P., Linarde C., Martine S., and others (2014) based on datasets of MO call records from Portugal and France shows how spatially and temporarily estimations of population densities can be produced at national scales. Therefore, the parameters of the created models, which represent the ratio between mobile phone user density and population density, are defined by using the census-derived national population.

In the Handbook on the use of Mobile Phone data for Official Statistics (2017) results of some research on the use of mobile positioning data in population statistics are mentioned. In addition to the study discussed above, references are given to "The Study on Urban Mobility and Dynamic by Using Aggregate Mobile Phone

Sources" (Japan), "Overview of the Sources and Challengers of Mobile Positioning Data for Statistics" (Estonia), "Estimating Population Density Distribution from Network-Based Mobile Phone Data" (Pan-European level), "Use of Mobile Phone Data to Estimate Mobility Flows. Measuring Urban Population and Intercity Mobility Using Big Data in an Integrated Approach" (Italy), and a research on dynamic population monitoring platform based on mobile phone data and methods for estimating population numbers in China. Unfortunately, some study materials were not available through the links provided. In general, given the lack of up-to-date population census data in Ukraine, as well as data of population registers, the use of approaches presented in publications is impossible. Although the presented study unquestionably is based on the ideas underlying existing research.

The study data sources. In Ukraine, at the end of the second decade of the 21st century, the population counting situation was as follows: as previously noted, the last census was conducted only in 2001, the demographic register and other population-related registers were not introduced, although regulations on some of them were adopted. Experts and scientists, individual politicians, and international organizations have increasingly raised the question of the need to conduct a census based on traditional statistical methods or using modern technologies, in particular, by filling out census materials online.

An attempt to quickly estimate the number and location of the population of Ukraine was made using data from mobile operators and some existing data from administrative registers, which were partially covered in such publications as the Population of Ukraine. Demographic dimension of living standards: joint monograph (2019), Sarioglo V., Dubilet D., Werner I. (2020). To carry out this operational assessment of the population of Ukraine, data were used from the following main sources:

- data from a special survey of the population on the use of mobile devices, the number of devices, mobile operators, etc.;
- data of three main Ukrainian mobile operators – Kyivstar, Vodafone Ukraine, lifecell on the number of subscribers by administrative-territorial units;
- data of the Pension Fund of Ukraine on the number of persons aged 60 and older by administrative-territorial units;
- administrative data on the number of children under 14 inclusive by administrative-territorial units.

The main purpose of the survey was to estimate the use of mobile devices by the population of Ukraine, taking into account the mobile operator that provides network services. The survey was conducted by the State Statistics in October-November 2019 on the basis of a representative (probabilistic) stratified multi-stage sample of households. About 32,000 households were selected for the survey, and the participation rate was about 84%. When developing the survey tools, it was taken into account that in order to estimate the population by the number of subscribers, it should be considered that the activity of mobile communication is different among different population groups (by age, region, type, etc.) together with the duplication of data of mobile operators in the part of persons who have several SIM-cards of different mobile operators (including several SIM-cards of one operator). Therefore, DMOs should be adjusted to accept these effects. In this context, mobile operators, as a rule, do not have their own reliable data for such an adjustment and it should be implemented based on the results of the survey.

The results of the survey were extended to the entire population of Ukraine, in view of the following basic assumptions. First, it was assumed that the vast majority of the institutional population that is not covered by state statistical sample surveys (individuals in prisons, on compulsory military service, in care houses, etc.) use mobile communications similarly to individuals, relevant sex and age groups living in private households. Second, the population (individuals living in settlements located in the territory affected by radioactive contamination as a result of the Chernobyl disaster, homeless individuals, etc.) is a small population, which is statistically insignificant.

The survey results provided an opportunity to reliably estimate the main characteristics of mobile users in Ukraine by gender and age groups, type of area – large city (with a population of 100 thousand people and more), small town, rural area – and by consolidated geographical regions. In total, five geographical regions were formed: Northern region (Kyiv, Kyiv region, Zhytomyr region, Sumy region, Chernihiv region); Central region (Cherkasy region, Poltava region, Kirovohrad region, Vinnytsia region); Eastern region (Dnipropetrovsk region, Donetsk region, Zaporizhia region, Luhansk region, Kharkiv region); Southern region (Odesa region, Mykolaiv region, Kherson region); Western region (Ivano-Frankivsk region, Khmelnytsky region, Chernivtsi region, Lviv region, Rivne region, Ternopil region, Volyn region, Zakarpattia region).

The survey found, in particular, that men and women use mobile communication approximately with the same intensity, and, accordingly, it can be assumed that the use of mobile communication by the population does not depend on gender. Mobile usage varies slightly by location and age group. As can be expected, the number of respondents using mobile communication decreases slightly in the older age groups for all three types of residence and is most noticeable for rural areas (Figure 1).

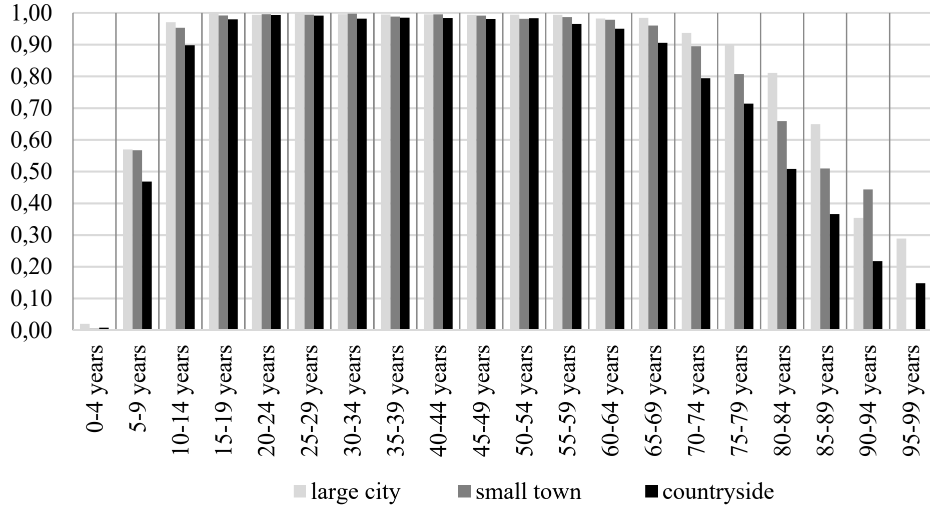


Figure 1. Proportion of mobile users by age group

Source: built by authors using data of the State Statistics survey.

In this regard, the use of mobile communications by type of residence differs markedly only for the population in the older age groups, in fact for people aged 65 years and older. It should be noted that for individuals aged 85 years and older, the survey results are not statistically reliable. As for children, at the age of 10–14 in rural areas the share of children who use mobile communication is about 6 pp (percentage points) smaller than in cities, and this difference increases for younger age groups.

Characterizing the peculiarities of the use of mobile communications by region, it should be noted that here the differences are observed only for older people. At the same time, the level of mobile communication use is the lowest in the western regions of Ukraine, and the highest – in its eastern regions. For example, for the age group 75–79 years the difference in the share of the population that uses mobile communication reaches 20 pp, in the western regions in this age group mobile communication is used by about 70% of the population, and in the eastern – about 90% of the population. The differences of the respective shares of the population are similar for individuals aged 80 and older. This is primarily due to the fact that the share of urban population in Ukraine is significantly higher for the eastern regions than for the western ones. Other regions of Ukraine – northern, central and southern – have an intermediate position on the use of mobile communications. It should be noted that the results of the survey of the use of mobile communication by the population of Ukraine provided an opportunity to reliably quantify the relevant indicators.

The research has shown that the degree of coverage of mobile services in the regions for the three main operators was also assessed (Figure 2). As can be seen from the above data, Kyivstar had the largest number of subscribers in the network in 2019, which was the most significant ahead of other companies in the western, central and northern regions of Ukraine. Vodafone Ukraine had the main share of subscribers in the eastern and southern regions of Ukraine. The mobile operator lifecell had the largest share of subscribers in the northern regions, but, in general, was significantly behind the number of subscribers to two other companies. Information on the coverage of the territory of Ukraine by mobile communication services provided by various operators is useful in estimating the population at the level of regions and administrative-territorial units.

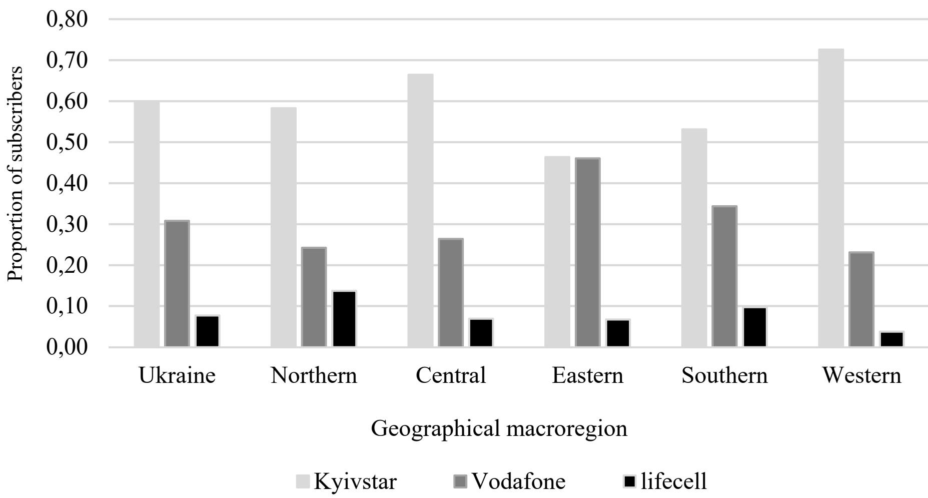


Figure 2. Proportion of subscribers of the main mobile operators of Ukraine by geographical macroregions in 2019

Source: built by authors using data of the State Statistics survey.

The information received from mobile operators had the form of a table with aggregated data by administrative-territorial units, the form of which is presented in Table 1. In the formation of the data set, operators have carried out significant work on their preparation. First of all, telephone numbers that belong to people but not to devices were set. Then the place of predominant stay of individuals (user telephone numbers) was established during a certain period and time (determination of home location). The basic number of the subscriber was also determined provided that many subscribers can use the services of different operators, as well as have several SIM-cards of one operator. It should be noted that mobile operators in the preparation of data tried to determine the sex of subscribers, but per their attribution, such estimates were very approximate and therefore were not used in estimating the population.

Table 1. Example of aggregated data for Zhytomyr region, obtained from major mobile operators, units

Territory	Mobile operator			
	number of phone numbers (total)	number of phone numbers		
		basic phone number	nonbasic phone number	phone number type not defined
Zhytomyr Region	960533	710757	125792	123984
Zhytomyr	244265	164923	29693	49649
Berdychiv	68642	49681	8441	10520
Korosten	52018	37358	6324	8336
Malyn	19720	14620	3088	2012
Novograd-Volynsky	33805	22124	6038	5643
Andrushiv District	24854	19324	3195	2335
Andrushivka	9895	7695	1134	1066
Baraniv District	31780	24842	4256	2682
…				

Source: data of mobile operators.

Estimation of population

The results of the population survey and the data of mobile operators make it possible to make estimates of the population. For example, based on the data of three mobile operators, the total number of subscribers (numbers) for the study period was 41,916 thousand. According to the results of a sample survey, the average number of SIM-cards per mobile user was 1.2133, and the average share of the population using mobile connection was 0.8908. Thus, the estimate of the total number of mobile users is about 41916 / 1.2133 = 34547.1 thousand people. The estimate of the total population of Ukraine, made on the basis of the above data, is: 34547.1/0.8908 = 38782.1 thousand people. It should be noted that according to the State Statistics, the current population of Ukraine as of November 1, 2019 was 41,940.7 thousand people. The calculations at the national level took into account PFC data on the number of individuals aged 60 and over and 65 years and older, and estimates of the proportion of these individuals according to demographic statistics and the results of state sample surveys of the population (households). It was taken into account that some retirees live in the temporarily occupied territories of Donetsk and Luhansk regions, the relevant data were obtained from the PFC. The calculation resulted in an average estimate of the current population of Ukraine, which amounted to 37289.4 thousand people.

The calculations of population estimates by region also took into account administrative data on the number of children under 14 and the number of individuals aged 60 and older, and estimates of the number of people aged 15-59 years obtained from mobile operators. It should be noted that the level of mobile use by the population of this age group is the highest and, accordingly, estimates of the number of subscribers are the most accurate. Regarding the estimate of the number of SIM-cards per person for this age group and the proportion of people who use mobile communications, the relevant characteristics were calculated following the survey by geographical regions, taking into account the type of place of residence. For mobile use, the corresponding estimates are 0.995 for all urban regions and 0.982 for rural areas. The number of SIM-cards for people aged 15-59 by geographical regions is presented in Table 2.

Table 2. Estimation of the number of SIM-cards per mobile user aged 15–59 by geographical macroregion and type of settlement

Geographical macroregion	Large city	Small town	Countryside
Northern	1.2781	1.2722	1.2680
Central	1.4376	1.2888	1.2057
Eastern	1.3462	1.2674	1.2855
Southern	1.3775	1.3733	1.3110
Western	1.2737	1.3016	1.2617

Source: built by authors using data of the State Statistics survey.

At the regional level as a whole, the estimated population was calculated as the sum of the number of children under 15 according to the administrative register, the number of individuals aged 60 and older as shown on the PFC records and the estimated number of individuals aged 15-59 determined by mobile operators, based on the proportion of such users among all users, the number of SIM-cards per user and the proportion of mobile users in the entire population of the region. After the calculations for all regions, the estimates were adjusted so that the total number of available population by region corresponds to the accepted estimate of the population of Ukraine. The results of population estimates by region are presented in Table 3.

Table 3. Estimation of the current population by regions of Ukraine

Region	The number of the available population according to the state statistics dated 1.11.2019	Estimation of the available population
Ukraine	**41940726**	**37289395**
Vinnytsia Region	1547575	1331403
Volyn Region	1032178	903669
Dnipropetrovsk Region	3181481	3229959
Donetsk Region	4137246	1981154
Zhytomyr Region	1210242	1062657
Transcarpathian Region	1254327	924677
Zaporizhzhia Region	1690423	1656736
Ivano-Frankivsk Region	1368761	1125696
Kyiv Region	1778132	2286396
Kirovohrad Region	935158	826817
Luhansk Region	2138533	1127460
Lviv Region	2513820	2290052
Mykolaiv Region	1121778	1053218
Odesa Region	2376607	2347885
Poltava	1389289	1337002
Rivne Region	1153702	943623
Sumy Region	1069941	932520
Ternopil Region	1039702	763582
Kharkiv Region	2661509	2794956
Kherson Region	1029615	913743
Khmelnytsky Region	1256356	1024697
Cherkasy Region	1194584	1088149
Chernivtsi Region	901908	727676
Chernihiv Region	993758	912547
Kyiv	2964101	3703121

Source: built by authors.

The population of the regions was distributed by administrative-territorial units (cities and districts) in proportion to the number of mobile subscribers. The admissibility of such a procedure was confirmed by the very high correlation between the distributions of mobile subscribers, individuals of retirement age according to the PFC and the population according to the State Statistics by administrative-territorial units for all oblasts. At the same time, the reliability of population estimates by administrative-territorial units, especially for small settlements, was expertly assessed as insufficient for use.

It should be noted that the obtained population estimates using the data of mobile operators fully confirm the existence of more noticeable effects of urbanization of Ukraine than shown by state statistics. For example, as can be seen from Table 3, according to estimates, the current population of Kyiv is almost 25% larger than according to the State Statistics. This is also confirmed by estimates from other studies conducted over the last decade, which estimated the population of Kyiv at even 4 million people. Similar trends are typical for some other large cities in Ukraine, although the effects are not so significant.

Conclusions

The development and application of modern approaches to population estimation and distribution based on data from such sources as administrative population records, data of mobile operators, results of special sample surveys, etc., are very relevant for Ukraine, especially in conditions of significant population movements as a result of the military aggression by Russia in 2014 and 2022, large-scale processes of external and internal labour migration, as well as the lack of opportunities for traditional censuses.

The results of the research showed that the data on number of subscribers by the main mobile operators – Kyivstar, Vodafone Ukraine, lifecell – are quite suitable for use in estimating and monitoring the population of Ukraine in case of availability of reliable and detailed data on the use of mobile devices by population. The potential efficiency of the use of data from mobile operators can increase significantly if census data are available, as this would provide an opportunity to build adequate statistical models to estimate the population at different levels of administrative-territorial division and control the reliability of such estimates.

The main advantages of these mobile operators are their efficiency and completeness, which is due to the fact that the results of mobile use are recorded in real time, stored for a long time period and mobile devices are used by the vast majority of the population. At the same time, it should be borne in mind that mobile data belong to commercial organizations, and this limits their availability, requires special processing and unification, which is due in particular to the fact that mobile operators have their own database architectures, record and store slightly different characteristics of subscribers and transactions, use different classifiers of territories, etc.

The experimental estimates of the population in Ukraine as a whole, by regions and administrative-territorial units, were significantly different from those calculated by the State Statistics on the basis of the 2001 census and administrative data on births, deaths and migration. Thus, the total population of Ukraine was estimated at about 37,289.4 thousand people, in contrast to the State Statistics estimate, which is 41,940.7 thousand people. In general, the population estimates obtained have not raised serious remarks from experts in Ukraine.

The developed approaches can be used to assess and monitor the number and location of the population of Ukraine, provided the availability and proper preparation of data of mobile operators, the availability of administrative records containing information about the population, the availability of sample surveys, in particular on the peculiar use of mobile communications by the population. At the same time, the presence of operational estimates of the population does not eliminate the need to conduct, if possible, a census of the population of Ukraine, including with the use of modern technologies for collecting information about the population.

References

Handbook on the use of Mobile Phone data for Official Statistics, (2019). UN Global Working Group on Big Data for Official Statistics, September 2019. Retrieved from: https://unstats.un.org/bigdata/taskteams/mobilephone/MPD%20Handbook%20220191004.pdf.

Lyuds′kyy rozvytok v Ukrayini. Otsinka ta prohnoz rivnya zhyttya naselennya: kol. monohr/NAN Ukrayiny, In-t demohrafiyi ta sotsial′nykh doslidzhen′ imeni M. V., (2019). Ptukhy — Kyiv, 270 p [in Ukrainian].

Sarioglo, V., Dubilet, D., Verner, I., (2020). Otsinka chysel′nosti nayavnoho naselennya Ukrayiny. Prezentatsiya, Kyiv,. Retrieved from: https://buhgalter.com.ua/upload/news/2020/Оцінка_наявного_населення_України.pdf [in Ukrainian].

Deville P., Linarde, C., Martine, S., Gilbert, M., Stevens, F. R., & Gaughan, A. E., (2014). Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences,* 111(45), pp. 15888–15893. DOI: 10.1073/pnas.1408439111.

Järv, O., Ahas, R., Saluveer, E., Derudder, B. and Witlox, F., (2012). Mobile phones in a traffic flow: A geographical perspective to evening rush hour traffic analysis using call detail records. *PLoS One,* 7 (11), e49171. Retrieved from: https://doi.org/10.1371/journal.pone.0049171.

Ratti, C, Pulselli, R. M., Williams, S. and Frenchman, D., (2006). Mobile landscapes: Using location data from cell phones for urban analysis. *Environment and Planning B: Planning and Design,* 33(5), pp. 727–748. DOI:10.1068/b32047.

Reades, J., Calabrese, F. and Ratti, C., (2009). Eigenplaces: Analysing cities using the space-time structure of the mobile phone network. *Environment and Planning B: Planning and Design*, 36(5), pp. 824–836. Retrieved from: https://doi.org/10.1068/b34133t.

De Meersman, F., Seynaeve, G., Debusschere, M., Lusyne, P., Dewitte, P., Baeyens, Y., (2016). Assessing the Quality of Mobile Phone Data as a Source of Statistics. *European Conference on Quality in Official Statistics (Q2016).* Madrid, 31 May – 3 June 2016. Retrieved from: https://ec.europa.eu/eurostat/cros/system/files/assessing_the_quality_of_mobile_phone_data_as_a_source_of_statistics_q2016.pdf.

Lu, H., Frauendorfer, D., Rabbi, M., Mast, M., (2012). StressSense: Detecting stress in unconstrained acoustic environments using smartphones. *UbiComp'12.ACM Conference on Ubiquitous Computing,* Pittsburgh, 5–8 September, pp. 351–360. DOI:10.1145/2370216.2370270.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 145–168, DOI 10.59170/stattrans-2023-009
Received – 01.07.2022; accepted – 03.10.2022

Budgetary policy of Ukraine in time of challenges and its impact on financial security

Taisiia Bondaruk[1], Liudmyla Momotiuk[2], Iryna Zaichko[3]

ABSTRACT

Ukraine has recently experienced a significant economic downturn as a result of the COVID-19 pandemic and the war caused by a large-scale military aggression of the Russian Federation. In conditions of the constant fluctuations of the national economy, the stimulating effect of the budgetary policy aimed at minimizing the consequences of such fluctuations and guaranteeing a sufficient level of financial security of the state becomes especially important.

The aim of the study is to deepen the theoretical and methodological foundations of the creation and implementation of budgetary policy in Ukraine, evaluation of its impact on the financial security in time of challenges.

The study uses methods of comparative analysis, grouping in the process of evaluating the current state of budgetary policy indicators, methods of normalization and standardization of data, modelling, and graphical analysis of data for normalizing the financial security indicators and determining the dynamics of financial security components. The materials and reports containing statistical data from the Ministry of Finance of Ukraine and the State Statistics Service of Ukraine served as the basis of the study.

We found out that the components of the financial security of the state in the face of the challenges posed by martial law and the pandemic do not take into account the impact of budgetary policy. We substantiated the thesis that the creation of Ukraine's budgetary policy under martial law requires adjustments to the financial security assessment system. The most statistically significant and reliable models of interrelation were selected for further use in multifactor modelling and forecasting the financial security of the state (on the basis of ranking the linear, polynominal, exponential, logarithmic and power dependencies within one-factor equations). It was experimentally proved that out of 122 statistically significant indicators, budgetary policy indicators such as the coefficient of financing the national

[1] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: bondaruk23@ukr.net. ORCID: https://orcid.org/0000-0001-9410-6428.

[2] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: momotyuk_le@ukr.net. ORCID: https://orcid.org/0000-0002-0445-5948.

[3] National Academy of Statistics, Accounting and Audit, Kyiv, Ukraine. E-mail: zkiv26@gmail.com. ORCID: https://orcid.org/0000-0002-6080-2959.

© Taisiia Bondaruk, Liudmyla Momotiuk, Iryna Zaichko. Article available under the CC BY-SA 4.0 licence

functions, the coefficient of public debt service and redemption, and the coefficient of the proportionality of financing the national security agencies had the greatest impact on the financial security of Ukraine.
We also substantiated the scientific provisions behind the modelling of the level of financial security of Ukraine taking into account the impact of budgetary policy in the period of challenges. In the process of modelling, the indicators of budgetary policy were identified, while regression analysis revealed the factors influencing the budgetary policy.

Key words: budgetary policy, financial security, multifactor modelling, linear dependences, polynomial dependences, logarithmic dependences, power dependences.

1. Introduction

The challenges posed by the damage to the domestic economy caused directly by the coronavirus pandemic COVID-19, as well as by Russia's large-scale military aggression, have led to a significant economic downturn. In conditions of constant fluctuations of the national economy, the stimulating effect of the budgetary policy aimed at minimizing the consequences of such fluctuations, ensuring the stability of the budget, creating the necessary conditions for its fulfilment and ensuring the appropriate level of socio-economic development becomes especially important. The trend of recent years indicates a predominant focus of budgetary policy to cover current budget expenditures, rather than ensure the implementation of strategic tasks of state development and working out the measures to improve its financial security.

First of all, it is critical for Ukraine to restore the country's economic potential, which has suffered losses caused by the destruction or shutdown of the enterprises, as well as the destruction of the infrastructure. According to the expert estimates of the Ministry of Economy, as of April 1, 2022, the total (both direct and indirect) losses of Ukraine's economy due to the war, such as: declining GDP, cessation of investment, outflow of labour, additional spending on defence and social support, etc., reached almost 600 billion USD. At the same time, the International Monetary Fund (IMF) predicts a 35% drop in Ukraine's GDP according to the results of 2022, and the World Bank forecasts a 45% drop in GDP and 15% inflation.

In these conditions, the budgetary policy in the system of financial security should be aimed at minimizing risks in the budget sphere, determining the forms of interaction of its components, conducting economic transformations to improve financial security. That is why the search for new methods, forms and means of implementing budgetary policy for the purpose of ensuring the financial security of the state in modern conditions needs special attention.

The aim of the study is to deepen the theoretical and methodological foundations of the formation and implementation of Ukraine's budgetary policy in times of challenges and its impact on the state of financial security.

2. The second section

2.1. Problems

The theoretical basis of modern public finance theory, on which the budget policy is based, is the theory of pure public goods, which was developed in the works of P. Samuelson (1955), R. Musgrave (1994) and J. Buchanan (1997). The theory is based on the model of general equilibrium, which analyzes the activities of both the state itself and private economic entities, and taxes serve as prices of public goods. However, in a market economy, the state can provide not only public but also private goods. The methodological basis for solving such problems was laid by K. J. Arrow (1971).

Examining the influence of government on the formation of budgetary policy, C. Tiebout (1956), R. Musgrave (1959, 1986) and W. Oates (2008) assert that neither large-scale centralization of government nor fully decentralized power, consisting of many small and local jurisdictions are likely to be effective. The central government should focus on the provision of national public services, the benefits of which are distributed throughout the whole country and the provision of which has significant economies of scale. Typical examples are defence, international relations, national infrastructure, monetary policy, macroeconomic stabilization, income redistribution and poverty reduction policies.

W. Dziemianowicz (2004) paid considerable attention to the study of Central European countries that had successfully carried out economic and institutional transformation in the field of public finance, namely, in the policy of attracting foreign capital and its impact on economic development in Poland.

Most of the scientific papers of these economists are focused on identifying general trends in the functioning of public finance. Paying tribute to the development of the theoretical basis of budgetary policy and financial security, it should be noted that the impact of budgetary policy on the financial security of the state has not acquired a holistic scientific vision, and this necessitates further research. Particular attention needs to be paid to determining the performance indicators that should be used to assess the level of financial security. A strategic guideline for budgetary policy with regard to financial security requires scientific rethinking. The search for effective budgetary tools to increase the level of financial security remains relevant.

2.2. Materials and methods

To evaluate the economic and financial security of the state, the following methods are proposed: observation of key macroeconomic indicators and comparing them with the threshold values for which the world averages are taken; assessment of the country's economic growth rates according to the main macroeconomic indicators, as well as the dynamics of their change; methods of expert evaluation used to describe the quantitative and qualitative characteristics of the studied processes. The first of these methods includes setting threshold values or indicators, i.e. in this case an indicator (indicative) approach is used, and it can be applied for the financial security. There are many papers dealing with indicative approach to the economic security of the country. They all differ mainly in the sets of economic indicators used. Besides, all these methods do not allow us to unequivocally assert the quantitative level of financial security.

The most important problem of characterizing financial security at the macro level is the definition of its main criteria and indicators. The basis for the formation of these indicators is the close relationship between the concept of "security" and the category of "risk".

The concept of risk in the economic security strategy contains two most important elements: risk assessment and risk management. Risk assessment is expert, probabilistic, in nature. Risk management involves potentially critical socio-economic situations in order to prevent, weaken and mitigate their effects. Thus, the assessment of the level of economic security allows, together with the analysis of risk factors, the use of categories of losses (damages) - actual, expected, potential, those that are compensated and not compensated.

In the process of research the following methods were also applied: the method of comparative analysis to specify the methodological foundations of the studying the impact of the budgetary policy; methods of statistical processing of information, comparative analysis, grouping in the process of assessing the current state of budgetary policy indicators; the method of normalization and standardization of data, graphical analysis of data when normalizing financial security indicators and determining the dynamics of financial security components.

The information base of the study consists of laws and regulations of the Verkhovna Rada of Ukraine and the Cabinet of Ministers of Ukraine, materials and reported statistical data of the Ministry of Finance of Ukraine, the State Statistics Service of Ukraine, materials of information and analytical bulletins, works of researchers, scientists and practitioners, personal analytical papers of the authors.

3. Results. Analysis of the current state of budgetary policy in the system of financial security

On February 24, 2022, the Verkhovna Rada of Ukraine imposed martial law because of Russia's direct full-scale invasion of Ukraine. According to Article 1 of the Law of Ukraine "On the legal regime of martial law" (2015), martial law is defined as a special legal regime that is introduced in Ukraine or in some of its localities in the event of armed aggression or a threat of attack and provides for the granting to the relevant state authorities of the powers necessary to prevent the threat and ensure national security.

The problems of the countries affected by the war are quite acute as they face the task, firstly, of creating a dynamically developing economy and, secondly, of creating the conditions for increasing economic and social integration. Martial law crises in Ukraine has challenged traditional governance mechanisms. The events that take place in a state of martial law require a change in approaches to the analysis of economic dynamics and justification of budgetary policy measures.

Budgetary policy has a decisive impact on budget security, and the latter is one of the most important components of both financial and national security. Most scholars point out that all aspects of national security are interconnected and interdependent, with budget security being of particular importance, as there is no aspect of national security that does not directly depend on the level of budget security. At the same time, the level of budget security itself is to a great extent determined by the level of other aspects of national security. This reveals the dual nature of the category of budget security. Consideration of these relationships and interdependences between budget security and other aspects of national security is the basis for developing measures to avoid and overcome threats to the national interests of the state in the budget sphere.

The dependence of all elements of national security on its financial and budgetary component is extremely simple: lack of financial resources leads to underfunding of the most urgent needs in various spheres of life, such as: in economy, social security, military and law enforcement activities, in the field of health care and education, causing serious threats to these areas. The main disadvantage of most state programs in Ukraine in recent years has been that one or another path of development was proposed either as an ideological dogma (regardless of whether it was a program of the former state plan or liberal programs), or as a set of projects and expenses. During the transformation period, there was a traditionally high activity of proposing the theories and schemes which have not been confirmed by world science (2020).

The main purpose of the budget security system is the material (financial) support of the process of strengthening all spheres of national security without exception: defence, environmental, informational, demographic, economic, political, social and energy. In other words, budget security is a basic component of the entire state security system, the failure to ensure which will inevitably lead to the deprivation of funds for the organization and functioning of the state security system as a whole.

The main goal of the budget security is to constantly maintain the state of the budget system, which should be characterized by equilibrium, resistance to internal and external negative influences, the ability to ensure effective socio-economic development of the country in the period of challenges caused by both pandemics and martial law. The parameters of the state's budget security must ensure its internal and external equilibrium, and their values must be sufficient to ensure proper resilience of the system to the action of threats in the period of challenges.

Thus, the central element of the study of budget security issues is the system of its ensuring. An important aspect is to assess the effectiveness of its construction and performance.

We will continue the study of the impact of the budgetary policy on the level of financial security by analyzing four indicators of Ukraine's budget security, which have reflected its state in dynamics over the past 13 years. The data given in Table 1 show that during 2009–2021 in the budgetary sphere of Ukraine there were changes in different directions. Thus, in 2018 compared to 2009, the state budget deficit in relation to GDP decreased by 2.1 p.p. (percentage points). The deficit of budgetary and extra-budgetary funds of the general government sector in relation to GDP in 2018 compared to 2009 decreased by 3.4 p.p., reaching a maximum (of 0.13%) in 2016 and a minimum (of 3.6%) in 2009. The volume of consolidated budget revenues at the end of 2018 amounted to one third of GDP, with an increase of 4.4 p.p. compared to 2009 and reached 34.1% (maximum) in 2017 and 28.1% (minimum) in 2010. Total payments for servicing and redemption of public debt against state budget revenues in 2009 were at the lowest level (of 4.7%), increasing in 2018 by 7.8 p.p. and reaching the highest level (of 16.2%) in 2015.

Table 1. Indicators of budget security of Ukraine in 2009–2021

Indicator name	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
1 The ratio of state budget deficit/ surplus to GDP, %	–3.75	–5.73	–1.75	–3.66	–4.25	–4.92	–2.27	–2.94	–1.60	–1.66	–1.96	–5.18	–3.63
2 Deficit / surplus of budget and off-budget funds of the general government sector,% of GDP	–3.60	–0.67	–0.84	–0.51	0.05	–0.23	–0.11	0.13	0.03	–0.21	0.07	–0.30	–1.70
3 The level of redistribution of GDP through the consolidated budget,%	28.82	28.07	29.54	30.53	29.08	28.74	32.79	32.81	34.08	33.26	32.45	32.82	30.45
4. The ratio of total payments for service and redemption of public debt to state budget revenues, %	4.72	7.30	7.82	7.44	10.14	14.29	16.24	15.80	14.05	12.53	12.03	11.27	11.80

Source: authors' calculations based on the Ministry of Finance of Ukraine and the National Bank of Ukraine.

From the main provisions of the methodical guidelines (Guidelines for calculating the level of economic security of Ukraine 2013) it is clear that financial security indicators can acquire features of 3 main types: indicators-stimulants (type C), the growth of which clearly leads to an increase in the level of financial security; disincentive indicators (type B), the growth of which clearly leads to a decrease in the level of financial security; mixed action indicators (type A), which, growing to a certain optimal level, behave as stimulants, but their further increase has a negative impact on the level of financial security, which corresponds to the behaviour of the disincentive.

Given the characteristic values, the first three indicators of the budget security were indicators of type A, which is taken into account when normalizing them (Table 2).

Table 2. Normalized indicators and sub-index of budget security of Ukraine

Indicator name	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
1.Normalized ratio of the state budget deficit/surplus to GDP	0.550	0.147	1.000	0.533	0.250	0.384	0.654	0.788	1.000	1.000	1.000	0.202	0.526
2.Normalized deficit/surplus of budget and off-budget funds of the general government sector	0.199	0.734	0.768	0.703	1.000	0.645	0.623	1.000	1.000	0.641	1.000	0.940	0.460
3.Normalized level of redistribution of GDP through consolidated budget	0.918	0.993	0.846	0.764	0.892	0.926	0.614	0.612	0.528	0.583	0.637	0.612	0.770
4.Normalized ratio of total payments for service and redemption of public debt to state budget revenues	1.000	0.780	0.746	0.771	0.590	0.314	0.184	0.213	0.330	0.431	0.465	0.516	0.480
5 Subindex of budget security	0.663	0.646	0.841	0.689	0.668	0.550	0.513	0.650	0.719	0.669	0.778	0.558	0.551

Source: authors' calculations based on the Ministry of Finance of Ukraine and the National Bank of Ukraine.

The data in Table 2 show that in 2014 and 2015, as well as in 2020 and 2021, the value of the sub-index of budget security corresponded to an unsatisfactory level of 0.550-0.513 and 0.550-0.551, respectively. Such values can be explained by the fact that in 2014 Ukraine faced the largest challenges of the 21st century, including the economic crisis, the military conflict in the East, the annexation of Crimea by Russia. The decline in domestic demand and weak external demand led to a decline in real GDP in 2014 by 6.8%. In 2020–2021, the challenges were caused by the COVID-19 pandemic and the development of the global economic and social crisis, the largest in decades (Heyets, V., Lunina, I., 2021). Ukraine was in a state of intensive circulation of an infectious disease caused by the coronavirus SARS-CoV-19 (Stiglitz, 2020) and needed sufficient financial resources, which was a key factor in increasing budget expenditures and, in turn, reflected on budget security indicators. Recently, the government has faced the need to make complex budgetary policy decisions related to rising defence and security spending amid limited budget revenues, in conditions of significant economic and political uncertainty.

The study of the place and role of budgetary policy in the system of ensuring financial security involves, in addition to determining the indicators of assessment of budgetary policy, the implementation of analytical procedures to assess its impact on the state of financial security. The main purpose of this assessment is to determine the reliability, statistical significance and adequacy of the impact of key factors of budgetary policy, which most influenced the state of financial security during 2009–2021, and which can be used to work out the prospects for developing the budgetary policy.

The authors propose a number of additional indicators of financial security and substantiate their expediency and importance in considering the objective impact of budgetary policy on banking security, non-banking financial market security, debt security, budget security, currency security and monetary security. Among these indicators are: coefficient of participation of the National Bank of Ukraine (NBU) in the state budget: the share of revenues from the NBU in state budget revenues; the coefficient of influence of non-state pension funds (NPF): the ratio of the amount of pension contributions of NPF to the income of the Pension Fund of Ukraine itself; the coefficient of efficiency of Domestic Government Bonds (DGB): the ratio of DGB revenues to consolidated budget expenditures; the coefficient of public debt service: the share of costs for servicing and redemption of public debt in the consolidated budget expenditures; the coefficient of proportionality of financing of the national security agencies: the ratio of expenditures on public order, security and the judiciary to expenditures on defence; the coefficient of debt dependence: the volume of expenditures for servicing and redemption of public debt to gross domestic product (GDP); the coefficient of devaluation stability of the budget revenue base: the ratio of consolidated budget revenues to hryvnia UA – dollar USA exchange rate;

the coefficient of budget dependence on crediting: the share of crediting in the consolidated budget expenditures; the coefficient of household income stability: the ratio of consumer loans, given to households, to households income. In order to ensure the objectivity of the study, calculations of the proposed indicators were made for their further use in the assessment system (Table 3).

Table 3. Dynamics of the proposed additional indicators of budgetary policy in the system of financial security during 2009–2021

Indicators	The value of the indicator by year												
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Coefficient of participation of the NBU in the state budget	2.41	6.46	3.78	6.82	8.35	6.39	11.56	6.19	5.59	5.45	5.03	3.10	1.47
Coefficient of influence of NPF	0.76	0.78	0.79	0.83	0.95	1.09	1.11	1.70	1.19	0.99	0.89	0.84	0.76
Coefficient of efficiency of DGB	6.09	10.66	6.87	8.00	9.98	12.71	1.46	4.42	3.09	5.20	16.57	16.22	15.61
Coefficient of debt service	3.43	4.63	6.08	6.01	7.09	9.94	12.96	11.66	10.55	10.31	8.77	7.60	8.30
Coefficient of proportionality of financing the national security agencies	2.52	2.55	2.47	2.53	2.66	1.64	1.06	1.21	1.19	1.22	1.33	1.34	1.44
Coefficient of debt dependence	1.12	1.57	1.90	1.85	2.36	3.31	4.45	4.10	3.74	3.62	3.02	2.89	2.80
Coefficient of devaluation stability of the budget revenue base	26.92	30.32	39.49	43.31	42.44	30.04	24.48	24.12	29.83	34.12	49.90	51.01	60.92
Coefficient of budget dependence on crediting	0.92	0.36	1.14	0.87	0.11	0.95	0.45	0.22	0.20	0.15	0.29	0.33	0.26
Coefficient of household income stability	15.33	11.17	9.96	8.58	8.87	8.91	5.92	4.95	4.60	4.67	4.64	4.22	4.56

Source: authors' calculations based on the Ministry of Finance of Ukraine and the National Bank of Ukraine.

The data given in Table 3 show that the studied indicators had different vector dynamics. Thus, the coefficient of proportionality of financing national security agencies in 2021 compared to 2009 decreased more than twice, reaching the highest value of 2.7 in 2013, and the lowest one of 1.1 in 2015. Such a significant reduction in this indicator tells about a significant increase in defence spending caused by hostilities.

The coefficient of participation of the NBU in the state budget is declining from its highest value of 11.6 in 2015 to its lowest one of 1.47 in 2021. Over the past three years, there has been a negative trend towards a gradual decline in the share, i.e. from 5.45% in 2018 to 1.47% in 2021, which indicates a significant reduction in revenues from the NBU activity in the state budget revenues. The ratio of NPF pension contributions to PFU own revenues increased by 1.3 times, reaching the highest value of 1.7 in 2016 and the lowest value of 0.76 in 2021. This indicator tells about the underdevelopment of the NPF system, as their pension contributions are less than 1% of the PFU's own income in 2009-2013 and 2018-2021. Over the last five years, this figure has been declining, indicating a further decrease in the share of NPFs in pension insurance in Ukraine.

The coefficient of efficiency of DGB during 2019-2021 is growing compared to the previous period. The rapid, abrupt increase in the coefficient took place in 2019 and then it decreased slightly to 15.61 in 2021. This indicates the active implementation of DGB as one of the methods of solving debt problems caused by challenges, and increasing of the budget deficit. It should be noted that a special feature of government domestic borrowing in the last two years is also the fact that a significant part of DGB was purchased by non-residents. The high level of profitability attracts foreign investors. At the beginning of 2020, out of the total DGB package, non-residents owned bonds worth UAH 118.8 billion, i.e., almost 15% of their total volume.

The coefficient of proportionality of financing of the national security agencies indicates an increase in defence expenditures relative to expenditures on public order, security and the judiciary since 2014 after the start of hostilities in eastern Ukraine and the annexation of Crimea. Consolidated budget revenues, expressed in million USD at the end of 2018, amounted to 34.1 billion USD, having increased by 12.5% compared to 2009. At the same time, in 2009-2013 there was an increase in consolidated budget revenues in USD by 40%, and in 2014 a reduction by 41% and a further reduction to 76% in 2016 compared to 2013 took place. The consolidated budget revenues in USD in 2018 accounted for only 80% of 2013. Further growth took place in the period from 2019 to 2021. Such changes occurred due to the growth of the US dollar against hryvnia by 3.4 times over the past decade. The coefficient of budget dependence on crediting at the end of 2018 was only 0.15, having decreased by 0.8 in 2009, reaching the highest value of 1.14 in 2010 and the lowest one of 0.11 in 2013.

To assess the impact of budgetary policy on the state of financial security of the country, we consider the methods of factor analysis, the most acceptable for the purposes of our study. Assessing the impact of each indicator of budgetary policy on the integral index of financial security allowed us to conclude that it is necessary to use standard tools for data graphical analysis. This will help to determine the factor dependence between the selected indicators of budgetary policy and to objectively assess the state of financial security in the public sector. The following dependence equations are considered:

1) linear dependence , as

$$y = a_0 + a_1 x, \tag{1}$$

where y is the resulting feature or dependent variable (in our case - the integral index of financial security; x - factorial feature or independent variable (in our case - the studied indicator of budget policy);

a1 is the regression coefficient, showing by how many unities the resulting feature changes with the growth of the factor feature by 1;

a0 is a constant showing the value of the resulting feature at x = 0;

2) polynomial dependence, as

$$y = a_0 + a_1 x + a_2 x^2; \tag{2}$$

3) exponential dependence, as

$$y = a_0 \cdot \exp(a_1 x); \tag{3}$$

4) logarithmic dependence, as

$$y = a_0 + a_1 \cdot \ln(x); \tag{4}$$

5) power dependence, as

$$y = a_0 \cdot x^{a1}. \tag{5}$$

We find it expedient to consider the level of approximation R^2 (coefficient of determination), which shows the statistical significance and reliability of the dependence equation, as a criterion for selecting one or another dependence. The initial condition for the statistical significance and reliability of the dependence equation is the value of $R^2 > 0.5$.

Figure 1 presents five models of the dependence of the integral index of financial security on state budget expenditures as a percentage of GDP. It should be noted that among the indicators of budgetary policy, the coefficient of redistribution of expenditures in relation to GDP is one of the few whose impact on the level of financial security of Ukraine is described by five statistically significant dependences. According to the results of the trend analysis, five equations of approximation were obtained. For factor analysis in terms of our study, we consider it appropriate to use the equations coefficient of determination R^2 of which is larger than 0.5. The conducted study indicates the significance and reliability of all obtained equations, but the quadratic equation has the highest level of reliability, which is 69%. Besides, all graphs of the dependence equations illustrate the negative slope, i.e. the inverse relationship – the growth of the factorial feature leads to a decrease of the effective feature.

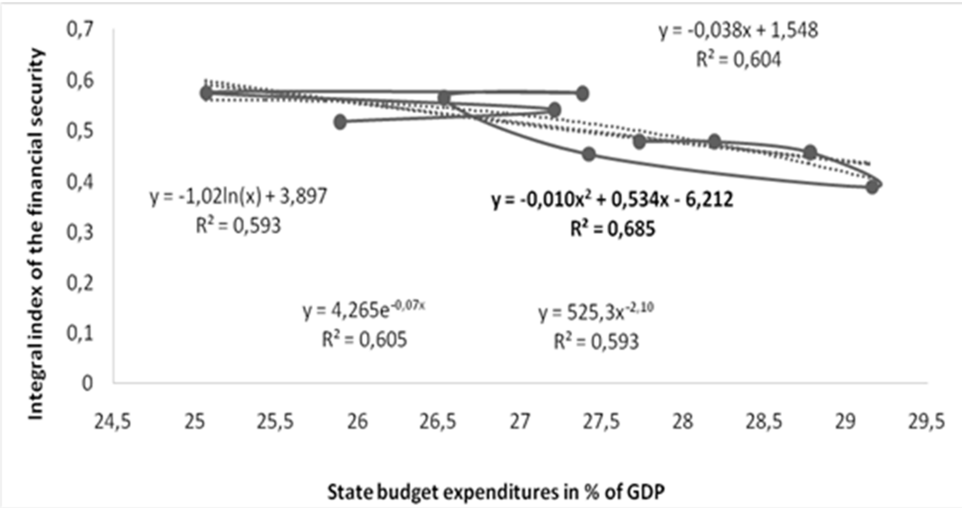


Figure 1. Results of modelling interrelation between the coefficient of redistribution of expenditures in relation to GDP and the integral index of financial security of the state

Source: developed according to Table 1, Table 2.

Similarly to Figure 1, the models of influence of all indicators of budgetary policy have been developed, the equations of linear dependences of an integral index of financial security of the state for 122 indicators of budgetary policy have been obtained as well as the corresponding values of coefficients of determination. Of 122 equations obtained, only 19 were statistically significant and reliable (R^2> 0.5), that is only 15.6% (Table 4). According to the linear model, the greatest influence on the level of financial security of Ukraine is exerted by such budget policy indicators as the coefficient of financing national functions (reliability 79.5%), the coefficient of proportionality of financing of the national security agencies (reliability 78.3%), the share of defence expenditures in the consolidated budget (reliability 73.2%) and the coefficient of budget dependence on crediting (reliability 70.3%).

Given the large number of results obtained on the impact of budget policy on financial security, in the future we will consider only models of interrelation with high reliability (R^2> 0.5) based on the linear function (see Table 4), quadratic function (second order polynomial, Table 5), exponential function (Table 6), logarithmic function (natural logarithm, Table 7) and power function (Table 8).

Table 4. Results of modelling the impact of budget policy indicators on the level of financial security of Ukraine using linear dependence

Indicator name	Equation of linear regression	R^2
1. Coefficient of redistribution of state budget expenditures relative to GDP	$y = -0.0382\,x + 1.5482$	0.6048
2. Coefficient of efficiency of local budget revenues by tax sources	$y = -0.0051\,x + 0.93$	0.5090
3. The share of revenues from indirect taxes in the consolidated budget	$y = -0.0314\,x + 2.2128$	0.6070
4. Coefficient of direct taxes in the tax revenues of the consolidated budget	$y = 0.0314\,x - 0.9281$	0.6070
5. Coefficient of financing national functions	$y = -0.0234x + 0.8242$	0.7951
6. The share of expenditures on national functions in the state budget	$y = -0.0113x + 0.7292$	0.6786
7. Coefficient of budget dependence on crediting	$y = -0.0158x + 0.6273$	0.7028
8. The share of expenditures for public debt service in the state budget	$y = -0.0086x + 0.6188$	0.6531
9. The share of defence expenditures in the consolidated budget	$y = -0.0241x + 0.6237$	0.7323
10. The share of defence expenditures in the state budget	$y = -0.013x + 0.6172$	0.6688
11. The share of expenditures on economic activities in the consolidated budget	$y = 0.0228x + 0.2613$	0.6021
12. The share of expenditures on economic activities in the state budget	$y = 0.0105x + 0.3665$	0.6351
13. The share of expenditures on environmental protection in the consolidated budget	$y = 0.2781x + 0.2727$	0.4984
14. The share of expenditures on health care in the consolidated budget	$y = 0.0362x + 0.1101$	0.4956
15. The share of expenditures on education in the consolidated budget	$y = 0.0207x + 0.1099$	0.5952
16. Expenditures on servicing and redemption of public debt in % to GDP	$y = -0.0437\,x + 0.626$	0.6956
17. The share of spending for servicing and redemption of public debt in expenditures	$y = -0.0158\,x + 0.6338$	0.678
18. Coefficient of proportionality of financing the national security agencies	$y = 0.0788\,x + 0.3534$	0.7829
19. Coefficient of devaluation stability of the budget revenue base	$y = 7 \cdot 10^{-6}\,x + 0.2758$	0.645

Source: compiled and calculated by the authors

Table 5. Results of modelling of the impact of budget policy indicators on the level of financial security of Ukraine using polynomial dependence

Indicator name	Equation of quadratic regression	R^2
1. Coefficient of redistribution of state budget expenditures relative to GDP	$y = -0.0106\,x^2 + 0.5347\,x - 6.2126$	0.6857
2. Deficit of local budgets	$y = -5 \cdot 10^{-10}x^2 - 2 \cdot 10^{-8}x + 0.5353$	0.5874
3. Coefficient of budget deficit (surplus) of local budgets relative to GDP	$y = -0.0479\,x^2 - 3 \cdot 10^{-5}x + 0.5354$	0.5900
4. The share of revenues from indirect taxes in the consolidated budget	$y = -0.008\,x^2 + 0.833\,x - 21.226$	0.7167
5. The share of revenues from indirect taxes in local budgets	$y = -0.0079\,x^2 + 0.0521\,x + 0.4862$	0.7571
6. Coefficient of direct taxes in the tax revenues of the consolidated budget	$y = -0.008\,x^2 + 0.7597\,x - 17.563$	0.7167
7. The share of revenues from direct taxes in local budgets	$y = -0.0079\,x^2 + 1.5239\,x - 73.104$	0.7571
8. Expenditures on national functions in the consolidated budget	$y = 1 \cdot 10^{-5}x^2 - 0.0034x + 0.6794$	0.5574
9. Expenditures on national functions in the state budget	$y = 2 \cdot 10^{-5}x^2 - 0.0037x + 0.6626$	0.5678
10. Coefficient of budget dependence on crediting	$y = -0.0029x^2 + 0.0307x + 0.4718$	0.8456
11. The share of expenditures for public debt service in the state budget	$y = -0.0011x^2 + 0.0213x + 0.455$	0.8186
12. The share of defence expenditures in the consolidated budget	$y = 0.0059x^2 - 0.0859x + 0.7595$	0.7649
13. The share of defence expenditures in the state budget	$y = 0.0033x^2 - 0.0743x + 0.8543$	0.7805
14. The share of expenditures on public order, security and the judiciary in the consolidated budget	$y = 0.075x^2 - 1.3173x + 6.2371$	0.5722
15. The share of expenditures on economic activities in the consolidated budget	$y = -0.0025x^2 + 0.0758x - 0.012$	0.6186
16. The share of expenditures on economic activities in the state budget	$y = 0.0002x^2 + 0.0044x + 0.4016$	0.6378
17. The share of expenditures on environmental protection in the consolidated budget	$y = -0.3897x^2 + 0.9794x - 0.0323$	0.5099
18. The share of expenditures on health care in the consolidated budget	$y = 0.0341x^2 - 0.6895x + 3.9192$	0.7286
19. The share of expenditures on health care in the state budget	$y = 0.0263x^2 - 0.1092x + 0.5428$	0.7715
20. The share of expenditures on education in the consolidated budget	$y = 0.0022x^2 - 0.0607x + 0.8603$	0.6057

Table 5. Results of modelling of the impact of budget policy indicators on the level of financial security of Ukraine using polynomial dependence (cont.)

Indicator name	Equation of quadratic regression	R^2
21. The share of expenditures on education in the state budget	$y = -0.0025x^2 + 0.0653x + 0.1142$	0.5219
22. The share of expenditures on education in local budgets	$y = 0.0045x^2 - 0.263x + 4.2718$	0.5121
23. Coefficient of stability of revenue base of local budgets	$y = -72.165\,x^2 + 145.15\,x - 72.453$	0.6778
24. Coefficient of deficit (surplus) of local budgets	$y = -72.165\,x^2 + 0.825\,x + 0.5366$	0.6778
25. Coefficient of debt dependence relative to GDP	$y = -0.0268\,x^2 + 0.1059\,x + 0.4501$	0.8582
26. Coefficient of public debt service and redemption	$y = -0.0036\,x^2 + 0.043\,x + 0.4271$	0.8907
27. Coefficient of budget dependence	$y = -61.452\,x^2 + 40.856\,x - 6.2477$	0.5720
28. Coefficient of proportionality of financing of the national security agencies	$y = -0.0004\,x^2 + 0.0804\,x + 0.3522$	0.7829
29. Coefficient of influence of NPF	$y = 0.2978\,x^2 - 0.8482\,x + 1.037$	0.5559
30. Coefficient of devaluation stability of the budget revenue base	$y = -1 \cdot 10^{-10} x^2 + 1 \cdot 10^{-5} x + 0.1524$	0.6494

Source: compiled and calculated by the authors.

As a result of modelling the impact of budget policy indicators on the level of financial security of Ukraine using polynomial dependence, 122 quadratic dependences were obtained, of which 30 equations were statistically significant and reliable ($R^2 > 0.5$), which is 24.6% of their total, with the greatest impact on the level of financial security of Ukraine by quadratic function exerted by such indicators of budget policy as: the coefficient of public debt service and redemption (reliability 89.1%); the coefficient of debt dependence relative to GDP (reliability 85.8%); the coefficient of budget dependence on crediting (reliability 84.6%), and the share of public debt service expenditures in the state budget (reliability 81.9%). At the same time, the reliability of quadratic functions was higher than that of the linear ones, and the coefficient of proportionality of financing of the national security agencies was ranked only fifth with the same reliability.

Table 6. The results of modelling the impact of budget policy indicators on the level of financial security of Ukraine using exponential dependence

Indicator name	Exponential regression equation	R^2
1. Coefficient of redistribution of state budget expenditures relative to GDP	$y = 4.2659 \exp(-0.078\ x)$	0.6058
2. The share of revenues from indirect taxes in the consolidated budget	$y = 17.446 \exp(-0.065x)$	0.6236
3. Coefficient of direct taxes in the tax revenues of the consolidated budget	$y = 0.0255 \exp(0.0653x)$	0.6236
4. Coefficient of financing national functions	$y = 0.964\exp(-0.048\ x)$	0.7928
5. The share of expenditures on national functions in the state budget	$y = 0.7921\exp(-0.023x)$	0.6710
6. Coefficient of budget dependence on crediting	$y = 0.6451\exp(-0.033x)$	0.7085
7. The share of expenditures for public debt service in the state budget	$y = 0.6332\exp(-0.018\ x)$	0.6522
8. The share of defence expenditures in the consolidated budget	$y = 0.6378\exp(-0.049x)$	0.7146
9. The share of defence expenditures in the state budget	$y = 0.6288\exp(-0.026x)$	0.6467
10. The share of expenditures on economic activities in the consolidated budget	$y = 0.305\exp(0.0464x)$	0.5960
11. The share of expenditures on economic activities in the state budget	$y = 0.3798\exp(0.0211x)$	0.6087
12. The share of expenditures on health care in the state budget	$y = 0.286\exp(0.1543x)$	0.7290
13. The share of expenditures on education in the consolidated budget	$y = 0.2255\exp(0.0419x)$	0.5789
14. Coefficient of debt dependence relative to GDP	$y = 0.6429 \exp(-0.09x)$	0.6964
15. Coefficient of public debt service and redemption	$y = 0.6538 \exp(-0.032x)$	0.6842
16. Coefficient of proportionality of financing of the national security agencies	$y = 0.3689 \exp(0.1596x)$	0.7635
17. Coefficient of devaluation stability of the budget revenue base	$y = 0.3165 \exp(0.00005x)$	0.6183

Source: compiled and calculated by the authors.

When modelling the impact of budget policy indicators on the level of financial security of Ukraine using exponential dependence, 122 equations were obtained, with only 17, i.e. 13.9% of the total, being statistically significant and reliable (R^2> 0.5). It was established that the greatest impact on the level of financial security of Ukraine according to the exponential model is exerted by such budget policy indicators as: the coefficient of proportionality of financing of the national security agencies (reliability 76.4%); the share of expenditures on health care in the state budget (reliability 72.9%); the share of defence expenditures in the consolidated budget (reliability 71.5%) and the share of expenditures on public debt service in the consolidated budget (reliability 70.9%). The equations of logarithmic dependence are summarized in Table 7.

Table 7. Results of modelling the impact of budgetary policy indicators on the level of financial security of Ukraine using logarithmic dependence

Indicator name	Logarithmic regression equation	R^2
1. The share of expenditures on national functions in the state budget	$y = -0.213\ln(x) + 1.1377$	0.6373
2. Coefficient of financing national functions	$y = -0.315\ln(x) + 1.3229$	0.7688
3. Coefficient of direct taxes in the tax revenues of the consolidated budget	$y = 1.4459 \ln(x) - 5.0182$	0.6164
4. The share of revenues from indirect taxes in the consolidated budget	$y = -1.692 \ln(x) + 7.2659$	0.5990
5. Coefficient of redistribution of state budget expenditures relative to GDP	$y = -1.026 \ln(x) + 3.8972$	0.5938
6. The share of expenditures on economic activities in the consolidated budget	$y = -0.1\ln(x) + 0.7004$	0.5616
7. The share of defence expenditures in the state budget	$y = -0.089\ln(x) + 0.7261$	0.5093
8. The share of defence expenditures in the consolidated budget	$y = -0.119\ln(x) + 0.6845$	0.7551
9. The share of expenditures for public debt service in the state budget	$y = -0.114\ln(x) + 0.7415$	0.7077
10. Coefficient of budget dependence on crediting	$y = 0.2403\ln(x) - 0.0604$	0.6141
11. The share of expenditures on economic activities in the state budget	$y = 0.1239\ln(x) + 0.193$	0.6012
12. The share of expenditures on environmental protection in the consolidated budget	$y = 0.2472\ln(x) + 0.5532$	0.5048
13. The share of expenditures on health care in the state budget	$y = 0.2583\ln(x) + 0.1759$	0.7228

Table 7. Results of modelling the impact of budgetary policy indicators on the level of financial security of Ukraine using logarithmic dependence (cont.)

Indicator name	Logarithmic regression equation	R^2
14. The share of expenditures on education in the consolidated budget	$y = 0.3821\ln(x) - 0.6189$	0.5878
15. The share of expenditures on education in the state budget	$y = 0.1649\ln(x) + 0.1332$	0.5017
16. Coefficient of debt dependence relative to GDP	$y = -0.098 \ln(x) + 0.5958$	0.5587
17. Coefficient of public debt service and redemption	$y = -0.102 \ln(\mathrm{x}) + 0.7108$	0.5244
18. Coefficient of proportionality of financing of the national security agencies	$y = 0.1398 \ln(x) + 0.4226$	0.7830
19. Coefficient of devaluation stability of the budget revenue base	$y = 0.2325 \ln(x) - 1.907$	0.6509

Source: compiled and calculated by the authors.

According to the data in Table 7, when modelling using a logarithmic equation, 19 statistically significant and reliable equations were obtained ($R^2 > 0.5$), which is 15.6% of the total. According to the results of calculations, the greatest impact on the level of financial security of Ukraine according to the logarithmic model is exerted by such budgetary policy indicators as: the coefficient of proportionality of financing of the national security agencies (reliability 78.3%), the coefficient of financing national functions (reliability 76.7%), the share of defence expenses in the consolidated budget (reliability 75.5%,) and the share of expenditures on health care in the state budget (reliability 72.3%). It should be noted that the obtained factors are also the factors of greatest influence in linear and exponential models.

Table 8. Results of modelling the impact of budgetary policy indicators on the level of financial security of Ukraine using the power dependence

Indicator name	Power regression equation	R^2
1. Coefficient of redistribution of state budget expenditures relative to GDP	$y = 525.38\, x^{-2.104}$	0.5936
2. The share of revenues from indirect taxes in the consolidated budget	$y = 629568\, x^{-3.515}$	0.6146
3. Coefficient of direct taxes in the tax revenues of the consolidated budget	$y = 5 \cdot 10^{-6} x^{3.0071}$	0.6341
4. Coefficient of financing national functions	$y = 2.6684x^{-0.643}$	0.7638

Table 8. Results of modelling the impact of budgetary policy indicators on the level of financial security of Ukraine using the power dependence (cont.)

Indicator name	Power regression equation	R^2
5. The share of expenditures on national functions in the state budget	$y = 1.8223x^{-0.435}$	0.6305
6. Coefficient of budget dependence on crediting	$y = 0.7499x^{-0.206}$	0.5662
7. The share of expenditures on public debt service in the state budget	$y = 0.7899\ x^{-0.183}$	0.5112
8. The share of defence expenditures in the consolidated budget	$y = 0.7208x^{-0.241}$	0.7338
9. The share of defence expenditures in the state budget	$y = 0.808x^{-0.231}$	0.6844
10. The share of expenditures on economic activities in the consolidated budget	$y = 0.158x^{0.491}$	0.6098
11. The share of expenditures on economic activities in the state budget	$y = 0.2683x^{0.2483}$	0.5744
12. The share of expenditures on health care in the state budget	$\mathrm{y} = 0.2574\mathrm{x}^{0.5233}$	0.7057
13. The share of expenditures on education in the consolidated budget	$y = 0.0517x^{0.7721}$	0.5710
14. Coefficient of debt dependence relative to GDP	$y = 0.6043\ x^{-0.202}$	0.5601
15. Coefficient of public debt service and redemption	$y = 0.7666\ x^{-0.21}$	0.5303
16. Coefficient of proportionality of financing of the national security agencies	$\mathrm{y} = 0.4242\ x^{0.284}$	0.7682
17. Coefficient of devaluation stability of the budget revenue base	$y = 0.0039\ x^{0.4685}$	0.6287

Source: compiled and calculated by the authors.

The results of the calculations show that the greatest impact on the level of financial security of Ukraine according to the power model is exerted by budgetary policy indicators such as: the coefficient of financing national functions (reliability 76.4%); the coefficient of proportionality of financing of the national security agencies (reliability 76.8%); the share of defence expenditures in the consolidated budget (reliability 73.4%) and the share of expenditures on health care in the state budget (reliability 70.6%).

Thus, in the process of assessment of the impact of budgetary policy indicators on the level of financial security of the state the most statistically significant and reliable equations of interrelation were determined in each group of statistical dependencies, as well as key factors of influence. The results of the assessment of the impact of budgetary policy on the financial security of the state are summarized in Table 9.

Table 9. Summarized results of modelling the impact of budgetary policy indicators on the level of financial security of Ukraine

Model name	Number of equations	Average determination coefficient R^2
1. Linear	19	0.6575
2. Polynomial (quadratic)	30	0.6956
3. Exponential	17	0.6629
4. Logarithmic	19	0.6070
5. Power	17	0.6151
Total	102	0.6549

Source: compiled and calculated by the authors according to Tables 4–8.

The results of the analysis of the impact of budget policy indicators on the level of financial security of Ukraine are characterized by a tendency to recur in different models of interrelation. These indicators interact with each other, responding accordingly to changes in budgetary policy. Given the above, it can be concluded that the following indicators have a significant impact on the level of financial security:

I. *Indicators of efficiency of budgetary policy of revenues and expenses*:

1. Expenditures on national functions in the state budget.
2. Expenditures on national functions in the consolidated budget.
3. Deficit of local budgets.
4. The share of expenditures on national functions in the state budget.
5. Coefficient of financing national functions.
6. The share of expenditures on public debt service in the state budget.
7. Coefficient of budget dependence on crediting.
8. The share of expenditures on public order, security and the judiciary in the consolidated budget.
9. The share of defence expenditures in the state budget.
10. The share of defence expenditures in the consolidated budget.
11. The share of expenditures on economic activities in the state budget.
12. The share of expenditures on economic activities in the consolidated budget.
13. The share of expenditures on education in the state budget.
14. The share of expenditures on education in the consolidated budget.
15. The share of expenditures on education in local budgets.
16. The share of expenditures on health care in the state budget.
17. The share of expenditures on health care in the consolidated budget.

18. The share of expenditures on environmental protection in the consolidated budget.
19. Percentage share of revenues from indirect taxes in the consolidated budget.
20. Percentage share of revenues from indirect taxes in local budgets.
21. Coefficient of direct taxes in the tax revenues of the consolidated budget.
22. Percentage share of revenues from direct taxes in local budgets.
23. Coefficient of efficiency of local budget revenues by tax sources.
24. Coefficient of redistribution of state budget expenditures relative to GDP.

II. *Indicators of dependence of budget expenditures on revenues:*

25. Coefficient of stability of the revenue part of local budgets.
26. Coefficient of deficit (surplus) of the local budgets relative to GDP.
27. Coefficient of deficit (surplus) of the local budgets.
28. Coefficient of debt dependence relative to GDP.
29. Coefficient of public debt service and redemption.

III. *Indicator of effectiveness of inter-budget policy:*

30. Coefficient of budget dependence.

IV. *Additional indicators of budgetary policy in assessing the level of financial security:*

31. Coefficient of proportionality of financing of the national security agencies.
32. Coefficient of debt dependency.
33. Coefficient of debt service.
34. Coefficient of influence of NPF.
35. Coefficient of devaluation stability of the budget revenue base, million USD.

Thus, in assessing the impact of budgetary policy on the state of financial security of the country, the authors performed the following tasks:

1) to ensure the objectivity, breadth and comprehensiveness of the factor analysis of the level of financial security of Ukraine, calculations were made for 122 absolute and relative indicators, which represent most aspects of modern budget policy, 35 of which have a significant impact on financial policy;
2) by the criterion of statistical significance and reliability of the regression equation (coefficient of determination $R^2 > 0.5$) the ranking of one-factor equations of linear, polynomial, exponential, logarithmic and power dependences was carried out, which allowed to select the most statistically significant and reliable models of interrelation for use in multifactor modelling and forecasting of financial security of the state;
3) it has been experimentally proved that of 122 statistically significant indicators the following indicators of budget policy have the greatest impact on the level of financial security of Ukraine: the coefficient of financing national functions (linear dependence), coefficient of debt service (polynomial dependence), coefficient of

proportionality of financing the national security agencies (exponential, logarithmic and power dependences).

4) it was found out that the most reliable groups of equations relative to the influence on the state of financial security are the group of polynomial (quadratic) dependences with the highest average reliability of 69.6%, the group of exponential equations with average reliability of 66.3% and the group of linear dependencies with average group reliability of 65.8%.

4. Conclusions

It was determined that the components of the state's financial security in the face of martial law and pandemic do not take into account the impact of budgetary policy. Therefore, in the course of comprehensive integrated assessment of the financial security of the state, additional indicators were proposed, such as: the coefficient of the NBU's participation in the state budget; the coefficient of influence of NPF; the coefficient of DGB efficiency and the coefficient of debt service; the coefficient of debt dependence, and the coefficient of proportionality of financing of the national security agencies; the coefficient of devaluation stability of the budget revenue base; the coefficient of the household income stability and the coefficient of budget dependence on crediting.

In the course of assessing the impact of budgetary policy on the state of financial security of the country, the expediency has been justified to consider 122 absolute relative indicators which represent most aspects of modern budgetary policy as factorial features of the level of financial security of Ukraine. The most statistically significant and reliable models of interrelation have been selected for their further use in multifactor modelling and forecasting of financial security of the state based on ranking of one-factor equations of linear, polynomial, exponential, logarithmic and power dependences. It has been experimentally proved that out of 122 statistically significant indicators, the greatest impact on the level of financial security of Ukraine is exerted by such budgetary policy indicators as the coefficient of financing national functions, the coefficient of public debt servicing and redemption, the coefficient of proportionality of financing of the national security agencies.

It is proved that the formation of Ukraine's budgetary policy in the period of challenges, under martial law in particular, requires adjustments to the indicators of the financial security assessment system. As a result of the analysis, the preconditions for developing a comprehensive multifactor model of financial security of Ukraine as a basis for forecasting and developing strategic guidelines for improving the level of financial security, taking into account the factors influencing budgetary policy, have been formulated.

References

Arrow, K. J., (1971). A Utilitarian Approach to the Concept of Equality in Public Expenditures. *Quarterly Journal of Economics,* vol.85, pp. 409–415.

Bjjukenen, D., (1997). *Sochynenyja* [Essays]. Moscow: Taurus Aljfa (in Russian).

Covid-19 Explorer. Ukraine, (2020). World Health Organization. URL: https://covid19.who.int/explorer.

Heyets, V., Lunina, I., Stepanova, O., (2021). The Formation of Fiscal Space While Overcoming Vaccine Nationalism and Ensuring Sustainability of Development. *Science andInnovation*, vol. 17, no. 1. pp. 29–41.

Jalowiecki, B., Dziemianowicz, W., (2004). *Polityka miejskaa inwestycje zagraniczne w polskich metropoliach [Urban politics and foreign investments in Polish metropolies]*, Warszawa: Scholar, pp. 134 (in Polish).

Ministry of Finance of Ukraine, (2019). “Bjudzhet Ukrajiny 2018: Statystychnyj Zbirnyk” [Budget of Ukraine 2018. Statistical yearbook] (in Ukrainian), Minfin. Kyiv, Ukraine, pp. 11–12.

Ministry of Finance of Ukraine, (2020). “Bjudzhet Ukrajiny 2019: Statystychnyj Zbirnyk” [Budget of Ukraine 2019. Statistical yearbook] (in Ukrainian), Minfin. Kyiv, Ukraine, pp.11–13.

Musgrave, R., Musgrave, P., Kullmer, L., (1994). Die Öffentliche Finanzen in Theorie und Praxis. Germany: Mohr (in German).

Musgrave, R. A., (1986). Economics of Fiscal Federalism. *Public Finance in a Democratic Society: Coll. Papers,* vol. 2, pp. 33–41.

Musgrave, R. A., (1959). The Theory of Public Finance: A Study in Public Economy. New York: McGraw-Hill. p. 84.

National Bank of Ukraine. FX market. URL: https://bank.gov.ua/ua/markets/currency-market.

National Bank of Ukraine. External Sector Statistics. URL: https://bank.gov.ua/statistic/sector-external/data-sector-external#4.

Oates, W. E., (2008) On the evolution of fiscal federalism: theory and institutions. *National Tax Journal*, vol. 61, pp. 313–334.

Osaulenko, O., Bondaruk, T., Momotiuk, L., (2020). Ukraine’s State Regulation of the Economic Development of Territories in the Context of Budgetary Decentralisation. *Statistics in transition new series*, vol. 21, no. 3, pp. 129–148.

Samuelson, P., (1955). Diagrammatic exposition of a theory of public expenditure. *Review of Economics and Statistics,* vol. 37, pp. 350–356.

Shhetynin, A., Uljjanova, L., Chajka, J., (2022). Deficyt derzhavnogho bjudzhetu I derzhavnyj borgh Ukrajiny: suchasnyj stan ta pidkhody do jogho pokrashhennja [State budget deficit and public debt of Ukraine: current state and approaches to its improvement]. *Economic space,* vol. 179, pp. 37–43.

Stiglitz, J., (2020). Conquering the Great Divide. *Finance & Development*. URL: https://www.imf.org/external/pubs/ft/fandd/2020/09/COVID19-and-global-inequality-joseph-stiglitz.htm.

Tiebout, C., (1956). A pure theory of local expenditures. *Journal of Political Economy*, vol. 64, pp. 416–424.

Ministry of Economic Development and Trade of Ukraine (2013) Methodological recommendations for calculating the level of economic security of Ukraine. URL:https://zakon.rada.gov.ua/rada/show/v1277731-13.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 169–197, DOI 10.59170/stattrans-2023-010
Received – 01.07.2022; accepted – 21.10.2022

Statistical modelling and forecasting of wheat and meslin export from Ukraine using the singular spectral analysis

Tetyana Chala[1], Oleksiy Korepanov[2], Iuliia Lazebnyk[3], Daryna Chernenko[4], Georgii Korepanov[5]

ABSTRACT

The article addresses the problems related to the functioning of the worldwide market of wheat and meslin. The authors identify the countries that over the past 17 years have been among the top 10 world leaders in terms of the value of export and import of wheat and meslin. The structure of wheat export by Ukrainian regions is analysed in comparison with the total export. The localisation coefficient is applied to measure the regional unevenness of the distribution of wheat export volumes and the total export by regions of the country. The modelling and forecasting of the volumes and prices of export of wheat and meslin from Ukraine are based on Singular Spectrum Analysis. The study particularly focuses on the individual components of time series, such as trend, annual, semi-annual, four-month, three-month seasonal components. The reliability of the forecast is confirmed by the calculation of the MAPE forecast error and Henry Theil's inequality coefficient. The article proposes an algorithm for calculating the relative indicators of the structure for the individual components of the reconstructed time series, identified through the singular spectral analysis.

Key words: export, forecasting, singular spectrum analysis, Ukraine, wheat and meslin.

[1] V. N. Karazin Kharkiv National University. Department of Statistics, Accounting and Auditing, Ukraine. E-mail: t.g.chala@karazin.ua. ORCID: http://orcid.org/0000-0001-7499-0308.
[2] V. N. Karazin Kharkiv National University. Department of Statistics, Accounting and Auditing, Ukraine. E-mail: o.s.korepanov@karazin.ua. ORCID: http://orcid.org/0000-0002-8499-0819.
[3] V. N. Karazin Kharkiv National University. Department of Statistics, Accounting and Auditing, Ukraine. E-mail: y.a.lazebnyk@karazin.ua. ORCID: http://orcid.org/0000-0002-2567-9764.
[4] V. N. Karazin Kharkiv National University. Department of Statistics, Accounting and Auditing, Ukraine. E-mail: d.i.chernenko@karazin.ua. ORCID: http://orcid.org/0000-0001-8655-0019.
[5] V. N. Karazin Kharkiv National University. Department of Statistics, Accounting and Auditing, Ukraine. E-mail: g.s.korepanov@karazin.ua. ORCID: http://orcid.org/0000-0001-7724-9339.
© T. Chala, O. Korepanov, I. Lazebnyk, D. Chernenko, G. Korepanov. Article available under the CC BY-SA 4.0 licence

1. Introduction

In February 2022, the world was shocked by the Russian invasion of Ukraine. The war has a devastating effect on human life and destroys the economies of both countries. It also causes significant economic losses in many parts of the world. This is the second major shock since the COVID-19 pandemic, which has caused an economic downturn worldwide.

With the beginning of the Russian invasion of Ukraine, martial law was imposed, which led to difficulties in conducting operational activities of the entire business environment of Ukraine. The change of plans and work also affected agribusiness.

The study of agricultural market conditions, aggregate supply and demand, in particular the country's export potential is one of the most pressing tasks at the present stage of human development at both the macro and meso levels (Agricultural markets, 2022; Europe, 2022). After all, information about the state of the agricultural market, the specific situation that has developed on it at the moment, as well as its analysis and evaluation is essential at the state level for effective management of agriculture and the economy as a whole and at the region and individual enterprise (Kozera-Kowalska, 2021).

Many scientific communities (ECA, EU Science Hub), foreign and domestic scientists have been and are engaged in research of agrarian economy in different regions of the world. Among them there are economists such as Maeda E. E., Abera T. A., Siljander M., Heiskanen J. (Maeda, 2021), Kozera-Kowalska M., Uglis J. (Kozera-Kowalska, 2021), Cameron A. (Cameron, 2022) and others. The issues connected with carrying out of researches of the agro-market, development of agrarian policy, etc., are considered in their works. As a result of studying these works, we can make conclusions about the complexity of internal processes in the agricultural market and the need for constant analysis of the main indicators of development of agro-industrial complex. This determines the relevance of the study of the features of the agricultural market especially during the COVID-19 pandemic.

This article was planned before the war, focusing on modelling and forecasting the monthly export value and average monthly wheat export prices from Ukraine using the Singular Spectrum Analysis (SSA) method during the COVID-19 pandemic. Special attention is attracted to the approach proposed by the authors to calculate the percentage structure of time series for the selected additive components.

The general issues of wheat export from Ukraine in modern conditions cannot be avoided. The main global problem during the war is the lack of stable logistics chains that can provide the necessary export. Due to the fact that grain sellers need to work from scratch, in the most efficient mode, it is not always possible to take into account all the nuances. Currently, grain traders are hampered by a number of issues that hinder the normalization of export processes. Among them are: blocked seaports of Ukraine; low capacity of railway stations on the border with European countries; lack of sufficient European wagons to transport grain; limited capacity of European logistics centres; limited freight by land; lack of drivers for international road transport (Suvoryj eksport, 2022).

Thus, the purpose of this article is to model and forecast the value of wheat and meslin export in Ukraine. To achieve the goal, the following tasks were set and completed: to determine the features of export of wheat and meslin from Ukraine in conditions of martial law, to analyse the value of export of wheat and meslin from Ukraine, to analyse the structure of wheat export by types of transport and regions of Ukraine, to identify its main partners in the world, to model and forecast the value and prices of wheat and meslin export in Ukraine using the Singular Spectrum Analysis.

2. The worldwide market of wheat and meslin under the influence of Russia's military aggression in Ukraine

The problems caused by Russia's war are estimated at about US$ 120 billion in world grain trade. Deliveries from Ukraine and Russia, which account for about 25 percent of world grain trade, are making it difficult to raise food security concerns (see Durisin, 2022).

The war in Ukraine is changing the world flows of agricultural export. For more than 15 years, Ukraine has been one of the ten largest wheat exporting countries in the world. This thesis is confirmed by the results of ranking the world by the value of export in 2005-2021 (see Table 1, 2). To perform correct calculations, the study was limited to the product group 1001 "Wheat and meslin" (from chapter 10 – "Cereals", section II – "Vegetable Products") according to The Harmonized Commodity Description and Coding System (HS) (RAMON). The content of product group 1001 "Wheat and meslin" did not change in different versions of the HS classification. Therefore, it is correct to compare data on international trade in wheat and meslin for the entire period under study.

The Harmonized Commodity Description and Coding System (HS) of the Customs Cooperation Council (CCC) is used worldwide as a reference for classifications of external trade statistics and for customs tariffs. It has the legal status of an international convention and has been in effect since 01.01.1988. The HS classification is further disaggregated at European Union level into a classification called Combined Nomenclature (CN) (RAMON).

The information source of this part of the study is annual data from the UN Comtrade Database. UN Comtrade is a repository of official international trade statistics (UN Comtrade Database).

Table 1. Top 10 countries exporting wheat and meslin (HS 2002 commodity code – 1001) in the world in 2005, 2010, 2015

№	Export of wheat and meslin in 2005			Export of wheat and meslin in 2010			Export of wheat and meslin in 2015		
	Country	bil US$	%	Country	bil US$	%	Country	bil US$	%
	World	**17.75**	**100.00**	**World**	**32.67**	**100.00**	**World**	**38.55**	**100.00**
1	USA	4.41	24.85	USA	6.78	20.74	Canada	6.22	16.13
2	France	2.40	13.55	France	4.66	14.25	USA	5.63	14.61
3	Australia	2.28	12.83	Canada	4.54	13.89	Australia	4.37	11.34
4	Canada	2.24	12.60	Australia	3.84	11.76	France	4.27	11.07
5	Argentina	1.28	7.22	RF	2.07	6.33	RF	3.95	10.24
6	RF	1.13	6.39	Germany	2.02	6.19	Germany	2.43	6.30
7	Germany	0.77	4.32	Kazakhstan	0.91	2.79	Ukraine	2.24	5.81
8	Ukraine	0.65	3.67	Ukraine	0.91	2.77	Argentina	1.03	2.68
9	UK	0.37	2.07	Argentina	0.90	2.76	Poland	0.85	2.21
10	Hungary	0.23	1.27	UK	0.71	2.18	Romania	0.77	2.00
	others	1.99	11.23	others	5.34	16.34	others	6.79	17.61

Source: authors' calculation based on UN Comtrade Database.

Analysing the global export of wheat and meslin in 2005–2015, we can note its significant growth (2.17 times), from US$ 17.75 billion in 2005 to US$ 38.55 billion in 2015. Export of wheat and meslin in 2005, 2010 and 2015 was mainly carried out in the USA, Canada, France, Australia.

In 2005 and 2010, Ukraine ranked 8th among wheat and meslin exporters in the world, and in 2015 – 7th among 107 exporting countries.

Global export of wheat and meslin continued to increase in 2019–2021, reaching US$ 46.33 billion in 2021. Export of wheat and meslin in 2019–2021 was mainly carried out by the Russian Federation, the USA, Canada, France and Ukraine. Ukraine in 2019–2021 ranked 5th among wheat and meslin exporters in the world.

Table 2. Top 10 countries exporting wheat and meslin (HS 2017 commodity code – 1001) in the world in 2019, 2020, 2021

№	Export of wheat and meslin in 2019			Export of wheat and meslin in 2020			Export of wheat and meslin in 2021		
	Country	bil US$	%	Country	bil US$	%	Country	bil US$	%
	World	**40.33**	**100.00**	**World**	**44.80**	**100.00**	**World**	**46.33**	**100.00**
1	RF	6.40	15.88	RF	7.92	17.68	RF	8.88	19.17
2	USA	6.27	15.54	USA	6.32	14.10	USA	7.29	15.73
3	Canada	5.38	13.34	Canada	6.30	14.06	Australia	7.11	15.35
4	France	4.36	10.80	France	4.54	10.14	Canada	6.64	14.33
5	Ukraine	3.66	9.07	Ukraine	3.59	8.02	Ukraine	4.72	10.19
6	Australia	2.48	6.16	Australia	2.70	6.02	Germany	1.98	4.27

Table 2. Top 10 countries exporting wheat and meslin (HS 2017 commodity code – 1001) in the world in 2019, 2020, 2021 (cont.)

№	Export of wheat and meslin in 2019			Export of wheat and meslin in 2020			Export of wheat and meslin in 2021		
	Country	bil US$	%	Country	bil US$	%	Country	bil US$	%
7	Argentina	2.30	5.69	Germany	2.12	4.73	Romania	1.82	3.93
8	Romania	1.27	3.15	Argentina	2.03	4.53	India	1.72	3.71
9	Germany	1.25	3.11	Kazakhstan	1.14	2.54	Poland	0.99	2.14
10	Kazakhstan	1.00	2.49	Poland	1.05	2.33	Lithuania	0.83	1.79
	others	5.96	14.78	others	7.09	15.83	others	4.35	9.39

Source: authors' calculation based on UN Comtrade Database.

It should be noted that according to the UN Comtrade Database, India, which has historically sold its huge wheat crops in the domestic market due to government regulation, entered the top 10 world exporters in 2021, selling record volumes across Asia. According to Bloomberg, Brazilian wheat export in the first three months of 2022 significantly exceeded export in 2021, and Egypt is considering exchanging fertilizers for Romanian grain and is negotiating the purchase of wheat from Argentina. In Table 3 the world's largest importers of wheat and meslin in 2010, 2015, 2020 are presented.

Table 3. Top 10 importers of wheat and meslin (HS 2007 commodity code – 1001) in the world in 2010, 2015, 2020

№	Import of wheat and meslin in 2010			Import of wheat and meslin in 2015			Import of wheat and meslin in 2020		
	Country	bil US$	%	Country	bil US$	%	Country	bil US$	%
	World	**35.09**	**100.0**	**World**	**40.90**	**100.0**	**World**	**40.90**	**100.0**
1	Egypt	2.18	6.22	Egypt	2.52	6.17	Egypt	2.69	6.59
2	Italy	1.87	5.34	Algeria	2.40	5.87	Indonesia	2.62	6.40
3	Brazil	1.70	4.85	Indonesia	2.08	5.09	Turkey	2.33	5.71
4	Japan	1.67	4.75	Italy	2.05	5.00	China	2.26	5.53
5	Indonesia	1.42	4.06	Japan	1.65	4.04	Nigeria	2.15	5.26
6	Fmr Sudan	1.29	3.68	Brazil	1.33	3.24	Italy	2.04	5.00
7	Algeria	1.25	3.57	Nigeria	1.24	3.04	Philippines	1.63	3.98
8	Netherlands	1.08	3.07	Spain	1.21	2.95	Japan	1.52	3.73
9	Rep. of Korea	1.07	3.04	Netherlands	1.17	2.87	Brazil	1.46	3.57
10	Spain	1.06	3.02	Thailand	1.14	2.78	Morocco	1.42	3.48
	others	20.50	58.41	others	24.12	58.96	others	20.76	50.77

Source: authors' calculation based on UN Comtrade Database.

Egypt is the undisputed leader-importer. Indonesia, Italy, Japan, Brazil are also constantly presented in the top 10 importers of wheat and meslin in 2010, 2015, 2020.

3. Export of wheat and meslin from Ukraine

Ukraine has significantly increased grain export in recent decades. In particular, the volume of wheat and meslin export in 2019 reached the highest values since independence – 20.0 million tons. This is four times higher than in 2010 and almost 100 times higher than in 2000. The trade value of wheat and meslin export has a stable upward trend, and since 2015 – the growth rate has increased significantly (Figure 1).

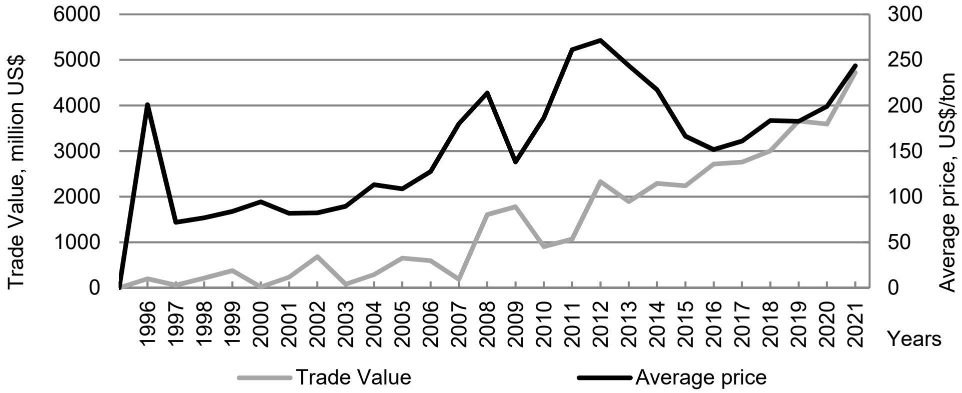


Figure 1. Export of wheat and meslin (HS 1992 commodity code – 1001) from Ukraine to the world in 1996–2021

Source: based on UN Comtrade Database.

In 2021, compared to 2010, the trade value of wheat and meslin export increased more than 5 times. And in comparison with 2000 – almost 250 times. The average price of wheat export in 2021 also became the highest in the last 10 years – 243.5 US$/ton and almost reached the level of 2012 (271.3 US$/ton). This is 22% more than in 2018, and four times higher than in 2010. Table 4 presents data on wheat and meslin export from Ukraine to the world in 2005, 2010, 2015.

Table 4. Top 10 countries in the world to which Ukraine exported wheat and meslin (HS 2002 commodity code – 1001) in 2005, 2010, 2015

№	Export of wheat and meslin in 2005			Export of wheat and meslin in 2010			Export of wheat and meslin in 2015		
	Country	million US$	%	Country	million US$	%	Country	million US$	%
	World	**652.17**	**100.0**	**World**	**906.42**	**100.0**	**World**	**2238.18**	**100.0**
1	Spain	154.15	23.64	Egypt	155.47	17.15	Egypt	310.76	13.88
2	Tunisia	57.88	8.87	Tunisia	85.08	9.39	Thailand	260.15	11.62
3	Israel	54.72	8.39	Israel	83.45	9.21	Indonesia	157.75	7.05
4	Italy	37.14	5.70	Bangladesh	74.86	8.26	Rep. of Korea	156.45	6.99

Table 4. Top 10 countries in the world to which Ukraine exported wheat and meslin (HS 2002 commodity code – 1001) in 2005, 2010, 2015 (cont.)

№	Export of wheat and meslin in 2005			Export of wheat and meslin in 2010			Export of wheat and meslin in 2015		
	Country	million US$	%	Country	million US$	%	Country	million US$	%
5	Algeria	36.55	5.60	Kenya	66.85	7.38	Bangladesh	136.78	6.11
6	Morocco	36.18	5.55	Libya	54.34	5.99	Spain	135.12	6.04
7	Indonesia	35.54	5.45	Rep. of Korea	53.71	5.93	Italy	107.30	4.79
8	Rep. of Korea	32.84	5.04	Turkey	49.78	5.49	Tunisia	105.90	4.73
9	Hungary	32.32	4.96	Viet Nam	32.98	3.64	Philippines	102.90	4.60
10	Libya	16.04	2.46	Syria	31.76	3.50	Israel	93.61	4.18
	others	493.35	75.65	others	218.14	24.07	others	671.46	30.00

Source: authors' calculation based on UN Comtrade Database.

Export of wheat and meslin from Ukraine in 2005, 2010 and 2015 was mainly to Spain, Egypt, Tunisia, Israel. The total number of countries to which Ukraine exported wheat and meslin increased from 53 countries in 2005 to 66 countries in 2015 and in 2021.

Table 5 presents data on wheat and meslin export from Ukraine to the countries of the world in 2019–2021.

Table 5. Top 10 countries in the world to which Ukraine exported wheat and meslin (HS 2017 commodity code – 1001) in 2019–2021

№	Export of wheat and meslin in 2019			Export of wheat and meslin in 2020			Export of wheat and meslin in 2021		
	Country	million US$	%	Country	million US$	%	Country	million US$	%
	World	**3658.40**	**100.0**	**World**	**3594.22**	**100.0**	**World**	**4722.75**	**100.0**
1	Egypt	664.47	18.16	Egypt	610.47	16.98	Egypt	850.72	18.01
2	Indonesia	537.05	14.68	Indonesia	543.19	15.11	Indonesia	638.02	13.51
3	Bangladesh	418.62	11.44	Bangladesh	294.85	8.20	Turkey	384.67	8.14
4	Turkey	207.44	5.67	Pakistan	258.01	7.18	Pakistan	350.28	7.42
5	Tunisia	195.54	5.35	Turkey	207.20	5.76	Morocco	275.01	5.82
6	Morocco	174.74	4.78	Morocco	196.16	5.46	Yemen	198.32	4.20
7	Philippines	169.80	4.64	Tunisia	194.38	5.41	Bangladesh	192.32	4.07
8	Thailand	155.66	4.25	Yemen	144.38	4.02	Saudi Arabia	170.62	3.61
9	Libya	129.64	3.54	Lebanon	133.23	3.71	Lebanon	159.45	3.38
10	Yemen	119.57	3.27	Philippines	118.89	3.31	Tunisia	151.96	3.22
	others	885.87	24.21	others	893.46	24.86	others	1351.37	28.61

Source: authors' calculation based on UN Comtrade Database.

Export of wheat and meslin from Ukraine to the world in 2019-2021 was mainly carried out to Egypt, Indonesia, Bangladesh, Turkey, Tunisia. The total number of countries in the world to which Ukraine in 2019 exported wheat and meslin reached 74 countries.

The total export of wheat and meslin from Ukraine to the world in 2020 amounted to 3594.22 million US dollars. Export of wheat and meslin to Ukraine in 2020 was mainly carried out to Egypt, Indonesia, Bangladesh, Pakistan, Turkey. The total number of countries in the world to which Ukraine exported wheat and meslin in 2020 reached 64 countries.

The total export of wheat and meslin to Ukraine to the world in 2021 amounted to 4722.75 million US dollars. Export of wheat and meslin to Ukraine in 2021 was mainly carried out to Egypt, Indonesia, Turkey, Pakistan, Morocco. The total number of countries in the world to which Ukraine exported wheat and meslin reached 66 countries in 2021.

4. The structure of wheat export by types of transport and regions of Ukraine

The analysis of market proportionality, namely: study of structural differences of certain market parts or segments, assessment of structural changes, analysis of uneven distribution between individual components of the population (assessment of concentration and localization in the market), etc., is of great practical importance in the statistical analysis and modelling of wheat and meslin export in Ukraine in general and in the regional context in particular. (Adán, 2020; Isard, 1960; Obrębalski, 2015; Rubalcaba, 2013; Sayago-Gomez, 2017). One of the manifestations of the transformation processes taking place in the wheat and meslin market in Ukraine is the structural changes in the volume of export by various modes of transport. Therefore, it is important to objectively display information about structural changes. This is possible only with the help of qualitative in-depth statistical analysis of the structure.

Export of grain from Ukraine in the pre-war period was carried out mainly by sea. In Table 6 the structure of wheat and meslin export from Ukraine in 2017–2021 by type of transport is presented.

Table 6. Structure of wheat and meslin (HS 2017 (equal to Ukrainian Classification of Goods for Foreign Economic Activity) commodity code – 1001) export from Ukraine in 2017–2021 by type of transport

Type of transport modes	Shares of wheat and meslin export, %				
	2017	2018	2019	2020	2021
Road	0.15	0.15	0.10	0.12	0.06
Rail	0.11	0.06	0.19	0.02	0.00
Maritime (Sea)	99.74	99.79	99.71	99.85	99.89
Others	0.00	0.00	0.00	0.01	0.05
Total	100.00	100.00	100.00	100.00	100.00

Source: authors' calculation based on the data of State Customs Service of Ukraine.

After the annexation of Crimea in 2014 and before the war, there were 13 seaports in Ukraine, 7 of which are located in Odessa region (Belgorod-Dniester, Odessa, Chernomorsk, Rhine, Izmail, Ust-Danube, South), 2 – in Mykolaiv (Mykolaiv and Olbia), 2 in Kherson (Kherson and Skadovsk), 1 in Zaporizhia (Berdyansk) and 1 in Donetsk region (Mariupol).

As of the end of April, access to 4 Ukrainian ports in Kherson, Donetsk and Zaporizhia oblasts was blocked. It is impossible for ships to enter and leave the Belgorod-Dniester port due to the lack of cargo and fleet. The ports of Mykolaiv, Olbia, Chernomorsk and Odesa are partially operational, but entry and exit of ships is also impossible.

The calculated shares of regions' wheat export in country's wheat export and regions' total export of goods in Ukraine's total export of goods in 2018–2021 are given in Table 7.

Table 7. Structure of wheat and meslin export (HS 2017 (equal to UCGFEA) commodity code – 1001) and total export of goods by regions (Total of all HS (equal to UCGFEA) commodities) of Ukraine in 2018–2021

Regions of Ukraine	2018		2019		2020		2021	
	Wheat and meslin export,%	Total export of goods,%	Wheat and meslin export,%	Total export of goods,%	Wheat and meslin export,%	Total export of goods,%	Wheat and meslin export,%	Total export of goods,%
	d_{ik}	d_{ij}	d_{ik}	d_{ij}	d_{ik}	d_{ij}	d_{ik}	d_{ij}
Vinnytsya	0.002	3.86	0.001	3.88	0.001	3.75	0.000	2.49
Volyn	0.047	1.94	0.017	1.85	0.139	1.75	0.003	1.58
Dnipropetrovsk	0.001	20.85	0.158	21.08	0.026	20.69	0.142	23.13
Donetsk	2.681	13.08	7.510	12.35	4.266	10.72	4.395	13.38
Zhytomyr	0.030	1.79	0.096	1.92	0.039	1.86	0.021	1.47
Zakarpattya	0.002	4.48	0.002	3.97	0.000	3.68	0.002	3.20
Zaporizhzhya	4.851	9.12	0.000	8.21	5.736	7.97	3.680	9.02
Ivano-Frankivsk	0.007	2.35	0.004	2.43	0.003	2.06	0.002	2.22
Kyiv	0.009	5.01	0.002	5.19	0.008	5.36	0.001	4.85
Kirovohrad	0.003	1.41	0.000	1.89	0.002	2.49	0.000	1.94
Luhansk	0.000	0.55	0.000	0.41	0.000	0.35	0.000	0.32
Lviv	0.017	5.12	0.192	5.87	0.016	6.33	0.002	5.58
Mykolayiv	**39.514**	5.70	**25.645**	5.74	**38.148**	6.13	**42.686**	6.57
Odesa	**51.675**	4.50	**65.498**	3.69	**50.618**	3.71	**48.509**	3.23
Poltava	0.031	5.15	0.028	5.62	0.013	6.31	0.001	6.07
Rivne	0.020	1.11	0.027	1.17	0.029	1.28	0.015	1.34
Sumy	0.018	2.17	0.000	2.36	0.000	2.66	0.001	2.02
Ternopil	0.186	1.22	0.127	1.16	0.274	1.22	0.027	1.29
Kharkiv	0.032	3.45	0.123	3.77	0.083	4.01	0.036	3.43

Table 7. Structure of wheat and meslin export (HS 2017 (equal to UCGFEA) commodity code – 1001) and total export of goods by regions (Total of all HS (equal to UCGFEA) commodities) of Ukraine in 2018–2021 (cont.)

Regions of Ukraine	2018		2019		2020		2021	
	Wheat and meslin export,%	Total export of goods,%	Wheat and meslin export,%	Total export of goods,%	Wheat and meslin export,%	Total export of goods,%	Wheat and meslin export,%	Total export of goods,%
	d_{ik}	d_{ij}	d_{ik}	d_{ij}	d_{ik}	d_{ij}	d_{ik}	d_{ij}
Kherson	0.832	0.73	0.540	0.71	0.548	0.76	0.466	0.79
Khmelnytskiy	0.026	1.72	0.000	1.71	0.018	1.81	0.000	1.75
Cherkasy	0.000	2.12	0.000	2.30	0.009	2.21	0.000	1.68
Chernivtsi	0.000	0.53	0.000	0.57	0.000	0.46	0.000	0.39
Chernihiv	0.016	2.04	0.030	2.15	0.024	2.43	0.011	2.26
Ukraine	100.000	100.00	100.000	100.00	100.000	100.00	100.000	100.00

Source: authors' calculations for wheat and meslin export based on State Customs Service of Ukraine and for total export of goods – State Statistics Service of Ukraine.

The analysis of the calculated shares showed that Mykolaiv and Odessa regions in 2018–2021 took the largest parts of all export of wheat. It is explained by big volumes of transportation by sea transport.

An important area of statistical analysis of structures is the characteristics of the uneven distribution of attribute values between the individual components of the population, the assessment of their concentration in individual parts (Flowers, 2020; Isard, 1960; Obrębalski, 2015; Rubalcaba, 2013; Sayago-Gomez, 2017).

This assessment is based on a comparison of the fractions of two distributions – the number of elements in the population d_j and the amount of values of the feature d_k. If the distribution is uniform, then $d_j = d_k$. The deviation of the particles indicates a certain uneven distribution, which is measured by the coefficients: localization (location quotient) and concentration (Flowers, 2020; Isard, 1960; Obrębalski, 2015; Rubalcaba, 2013; Sayago-Gomez, 2017).

The formula for a Wheat Export Location Quotient is given below in (1).

The location quotient is calculated by the ratio of the shares of the two distributions for each component of the population:

$$L_i = \frac{d_{ik}}{d_{ij}}, \tag{1}$$

where:

L_i is the wheat export location quotient for region *i*;

d_{ik} is the share of wheat and meslin export for region *i* in wheat and meslin export for Ukraine;

d_{ij} is the share of total export of goods for region i in total export of goods for Ukraine.

In the case of uniform distribution of all values $L_i = 1$. In the case of concentration of feature values in the i-th region $L_i > 1$, and vice versa.

Location quotients are widely used in regional analysis to assess the uniformity of territorial distribution (Obrębalski, 2015; Sayago-Gomez, 2017).

Based on the obtained data (Table 7), the Wheat Export Location Quotient (ratio of the shares of wheat export and the shares of total export of goods) by regions of Ukraine in 2017–2021 was calculated, and is given in Table 8.

Table 8. Wheat Export Location Quotient by regions of Ukraine in 2017–2021

Regions of Ukraine	Wheat Export Location Quotient				
	2017	2018	2019	2020	2021
Odesa	**10.668**	**11.474**	**13.717**	**17.667**	**15.023**
Mykolayiv	**5.639**	**6.929**	**6.647**	**4.185**	**6.492**
Kherson	**1.845**	**1.141**	0.766	0.706	0.593
Zaporizhzhya	0.592	0.532	0.698	0.000	0.408
Donetsk	0.212	0.205	0.345	0.700	0.328
Ternopil	0.062	0.152	0.237	0.104	0.021
Zhytomyr	0.034	0.017	0.020	0.052	0.015
Rivne	0.011	0.018	0.024	0.021	0.012
Kharkiv	0.007	0.009	0.022	0.031	0.010
Dnipropetrovsk	0.000	0.000	0.001	0.008	0.006
Chernihiv	0.012	0.008	0.011	0.012	0.005
Volyn	0.002	0.024	0.075	0.010	0.002
Zakarpattya	0.001	0.001	0.000	0.001	0.001
Ivano-Frankivsk	0.001	0.003	0.001	0.002	0.001
Vinnytsya	0.005	0.000	0.000	0.000	0.000
Kyiv	0.008	0.002	0.002	0.000	0.000
Kirovohrad	0.009	0.002	0.001	0.000	0.000
Luhansk	0.000	0.000	0.000	0.000	0.000
Lviv	0.002	0.003	0.003	0.030	0.000
Poltava	0.003	0.006	0.002	0.004	0.000
Sumy	0.002	0.009	0.000	0.000	0.000
Khmelnytskiy	0.008	0.015	0.010	0.000	0.000
Cherkasy	0.003	0.000	0.004	0.000	0.000
Chernivtsi	0.000	0.000	0.000	0.000	0.000

Source: authors' calculations for wheat and meslin export based on State Customs Service of Ukraine and for total export of goods – State Statistics Service of Ukraine.

The values of the calculated coefficients indicate a significant localization of wheat export in 2017–2021 in Odessa and Mykolaiv regions ($L_i > 1$). In 2017 and 2018, the localization of wheat export was also observed in Kherson region of Ukraine.

Odessa and Mykolayiv regions have the highest coefficients during the whole period, which indicates the largest localization of wheat export in these regions. This is due to the availability of sea transport for transportation and the largest international ports in Ukraine.

5. SSA analysis of wheat and meslin export from Ukraine

Analysis of the dynamics of the main indicators of wheat and meslin export from Ukraine confirmed the presence of the trend, as well as cyclical and seasonal fluctuations. During the period of sharp changes in the development of the agricultural market as a whole and its individual components to obtain a forecast of the dynamics of indicators studied under such conditions, various forecasting methods were tested, including various models of dynamics, namely: trend, autoregression and moving averages, seasonal decomposition, spectral Fourier analysis, multifactor prediction models (Armstrong, 2018; Dama, 2021; Golyandina, 2018; Hassani, 2007; Kwas, 2022; Petropoulos, 2022; Sukhanova, 2019; Theil, 1996; Wang, 2011; Wickham, 2017; Yerina, 2014).

Analysis of the deviations of empirical and calculated theoretical values of time series characterizing wheat export, as well as calculated forecast errors allowed the authors to conclude that the advantages of proven classical forecasting methods combine Singular Spectrum Analysis (SSA) (Golyandina, 2013; Golyandina, 2018; Hassani, 2007). To justify the chosen method, the initial time series were reduced by six months in order to be able to calculate the forecast error based on these data.

The SSA method combines elements of classical analysis, methods of nonlinear dynamics and signal processing. A distinctive feature of the method is that it allows one to work with a wide range of non-stationary time series and does not require a preliminary task of the series model, but it allows one to decompose the time series into additive components that can be interpreted as trend (in particular, linear or exponential), periodic and seasonal components and noise. It is not necessary to know in advance the parametric type of trend, as well as the presence of oscillating components and their periods.

We consider it expedient to build an additive model of time series of volumes and prices of wheat and meslin export from Ukraine as the sum of the relevant components. This model can be represented as follows:

$$\mathrm{Y}_t = \mathrm{U}_t + \mathrm{V}_t + \sum_{i=1}^{m} \mathrm{S}_{ti} + \xi_t, \tag{2}$$

where:

Y_t is series levels;

U_t is trend component;

V_t is cyclic component;

$\sum_{i=1}^{m} \mathrm{S}_{ti}$ is the sum of *m* seasonal components;

ξ_t is random component, *t* – time.

The technical basis of the method is the singular decomposition of the trajectory matrix, the rows of which are successive segments of a series of length *M* – the main parameter of the method, called the "window length" (Golyandina, 2013; Golyandina, 2018). On the basis of such decomposition we get eigen triples, each of which consists of eigenvalue, eigenvector and factor vector, which are further grouped and the additive component of the original series is restored.

During the analysis, the original one-dimensional time series were transformed into multidimensional ones, their research was carried out using the principal components method and further reproduction of the one-dimensional series, the original series and, if necessary, selected individual components were predicted.

For a detailed consideration of the advantages of this method, the analysis and forecasting of the dynamics of volumes and prices of wheat and meslin export from Ukraine. In the future, to shorten the name of the series, we omit the clarification "and meslin". Calculations were performed using the Caterpillar program based on a time series containing monthly data for the period from January 2000 to June 2021 (UN Comtrade Database). Also, it is possible to use package R (Golyandina, 2018).

The information source of this part of the study is monthly data about wheat and meslin export from Ukraine (HS 2017 commodity code – 1001) from the UN Comtrade Database (UN Comtrade Database). Monthly data sets may mix codes from multiple HS revisions and are provided except for standardization of trade flow and partner information, as well as conversion to U.S. dollars (UN Comtrade Database).

Data for July–December 2021 are used to parameterize the method and assess the adequacy of the model. The window length was chosen a multiple of 12, namely *M*=24, because in a number there is an obvious annual frequency.

To implement the algorithm of the SSA method, four stages were passed through (Golyandina, 2013; Golyandina, 2018). At the first stage, a one-dimensional time series (the graphic image of which is shown in Figure 2) was transformed into a multidimensional one. The chosen time series of wheat export $\{x_i\}_{i=1}^{N=138}$ is formed by a sequence of 138 equidistant values of the function: $f(t): x_i = f[i] = f\big((i-1)\Delta t\big),\ \ i = 1, 2, \dots, 138.$

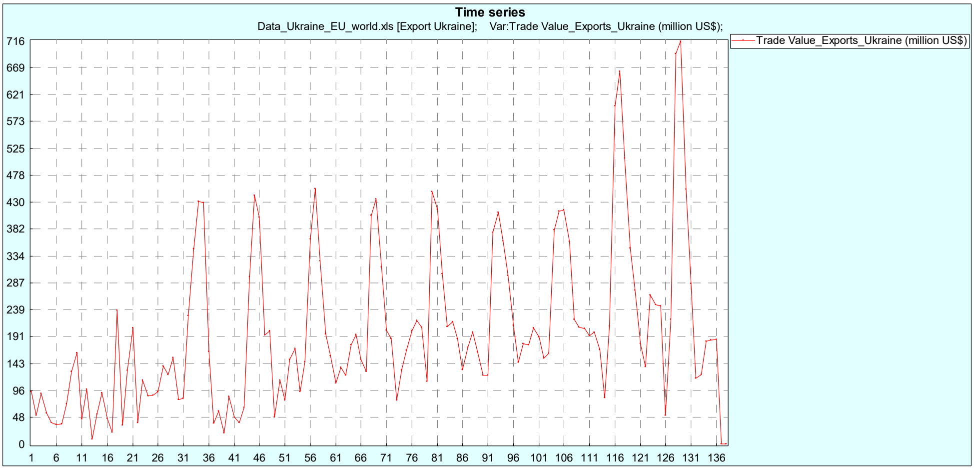


Figure 2. Export of wheat from Ukraine in January 2000–June 2021

Source: authors' calculation using SSA based on UN Comtrade Database.

The first row of the matrix of observations X consists of the first M=24 values of the time series. The second row of the matrix consists of sequences of values from x_2 to $x_{M+1} = x_{25}$ and so on to the last line with the number $k = N - M + 1 = 138 - 24 + 1 = 115$: from value x_k to x_N, where N is the series length. The elements of the matrix are equal to $x_{ij} = x_{i+j-1}$.

At the second stage, the analysis of the principal components is carried out: singular decomposition of the matrix.

The algorithm for implementing the method involves calculating a matrix of second moments:

$$R = \frac{1}{k} X^*(X^*)^T. \tag{3}$$

It should be noted that if the matrix X^* was obtained as a result of centring and normalization, the matrix R will be a correlation matrix with elements

$$r_{ij} = \frac{1}{k} \sum_{l=1}^{k} \frac{1}{S_i S_j} (x_{i+l-1} - \bar{x}_i)(x_{j+l-1} - \bar{x}_j). \tag{4}$$

The correlation matrix of multidimensional series of observations is shown in Figure 3.

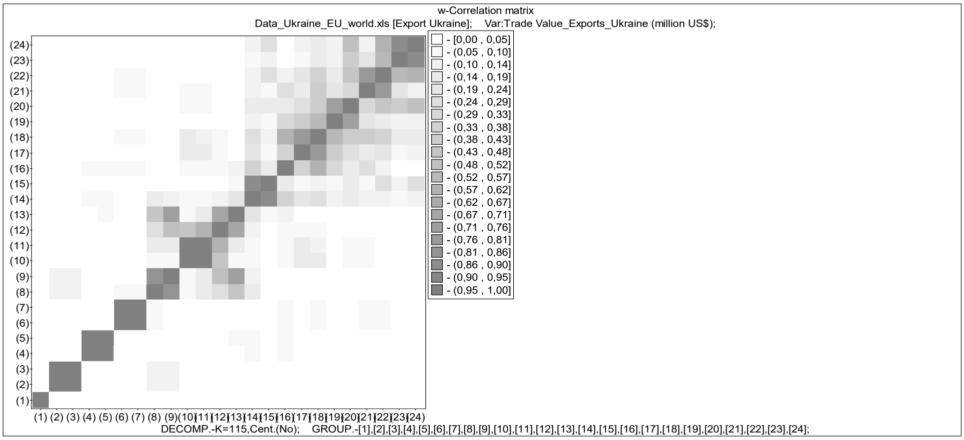


Figure 3. Correlation matrix of multidimensional series of observations

Source: authors' calculation using SSA based on UN Comtrade Database.

A visual mode is chosen for its image, in which larger values are highlighted in a darker colour.

At the third stage, the principal components were selected, which are informative for the selection of components of the time series of wheat export from Ukraine in January 2000–June 2021.

Based on the visual analysis of the diagrams of eigenfunctions and factor vectors, the structure of the studied population was identified, as well as the selection and interpretation of the principal components.

To select pairs of eigenvalues that identify one periodic, logarithms and roots of eigenvalues are shown in Figure 4.

In Figure 4 the "steps" corresponding to 2–3, 4–5, 6–7, 10–11 principal components are noticeable. This suggests that each of these pairs of principal components should be attributed to a separate periodical.

After analysing Figure 4 it is assumed that after about the 13th principal component there is a uniform decrease in very small eigenvalues, i.e. there is noise. Therefore, the first 13 principal components are enough to restore the original series.

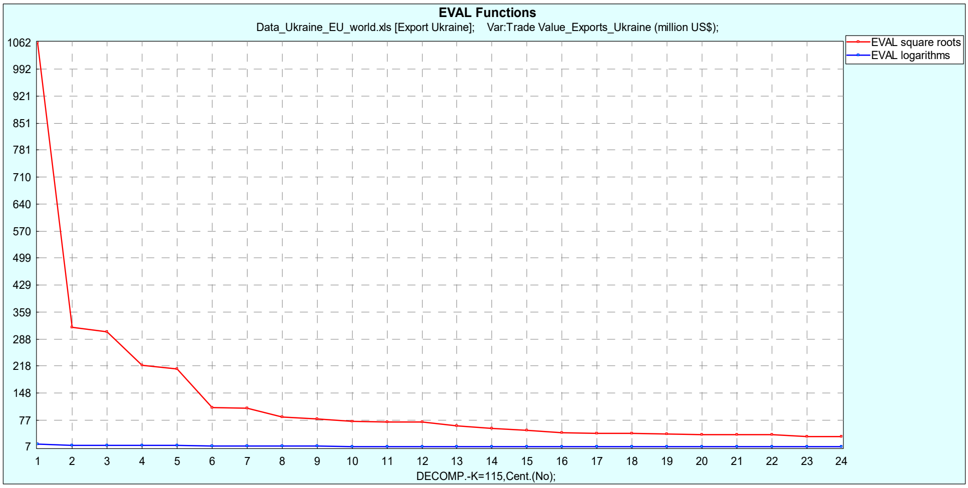


Figure 4. Roots of eigenvalues and logarithms of eigenvalues

Source: authors' calculation using SSA based on UN Comtrade Database.

The formulated assumption is confirmed by one-dimensional graphs of the principal components, shown in Figure 5.

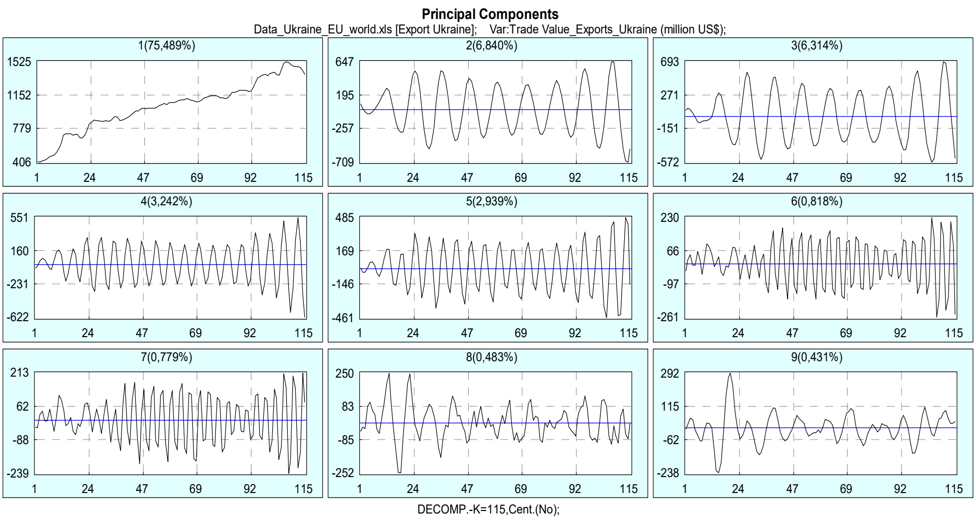


Figure 5. One-dimensional graphs of the principal components

Source: authors' calculation using SSA based on UN Comtrade Database.

After the analysis of the graphs shown in Figure 5, it is assumed that the first principal component is a part of the trend. The second and third principal components can be interpreted as annual periodicals. The fourth and fifth principal components can be interpreted as semi-annual periodicals. You can also see the four-month and three-month periodicity (principal components 6-7 and 10-11, respectively). Two-dimensional graphs of eigenvectors and principal components are constructed to confirm the assumptions about combining principal components into pairs (Figure 6).

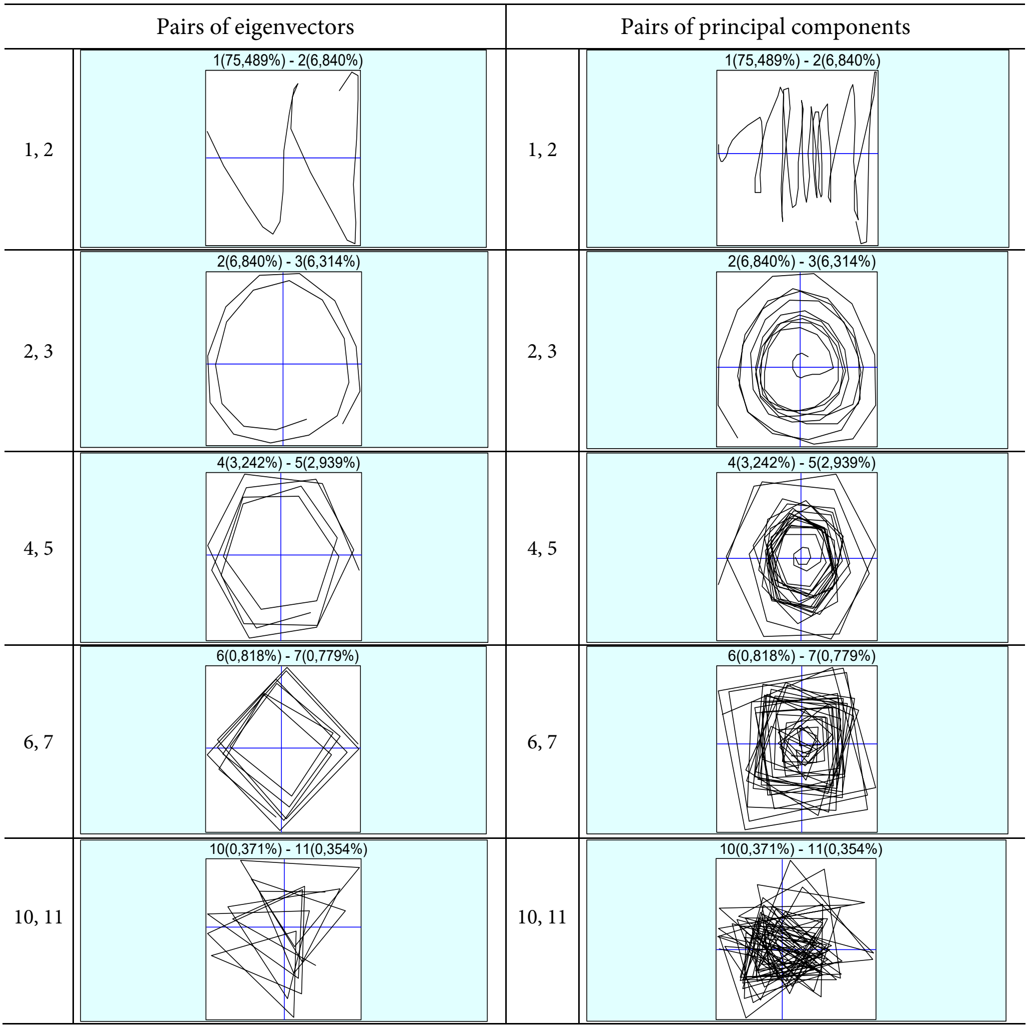


Figure 6. Two-dimensional graphs of eigenvectors and principal components

Source: authors' calculation using SSA based on UN Comtrade Database.

It is known that the two-dimensional image of sine and cosine creates a single circle. Therefore, the harmonic component with the whole period is represented as a regular polygon with the number of vertices corresponding to the length of the period (Golyandina, 2013; Golyandina, 2018). When the amplitude changes, the polygon turns into a spiral. The star-shaped image indicates the presence of periodicals with a small period.

At the fourth stage, the one-dimensional series were reconstructed. Graphic representation of the original and reconstructed series by selected pairs of the principal components of the value of wheat export is presented in Figure 7.

We emphasize the fact that the presence of periodicals with multiple periods means that in this case, analysing the time series of wheat export from Ukraine, we found the annual periodicals and their schedule for sinusoidal components.

1-st PC

on 2-3rd PC

on 4-5th PC

on 6-7th PC

on 10-11th PC

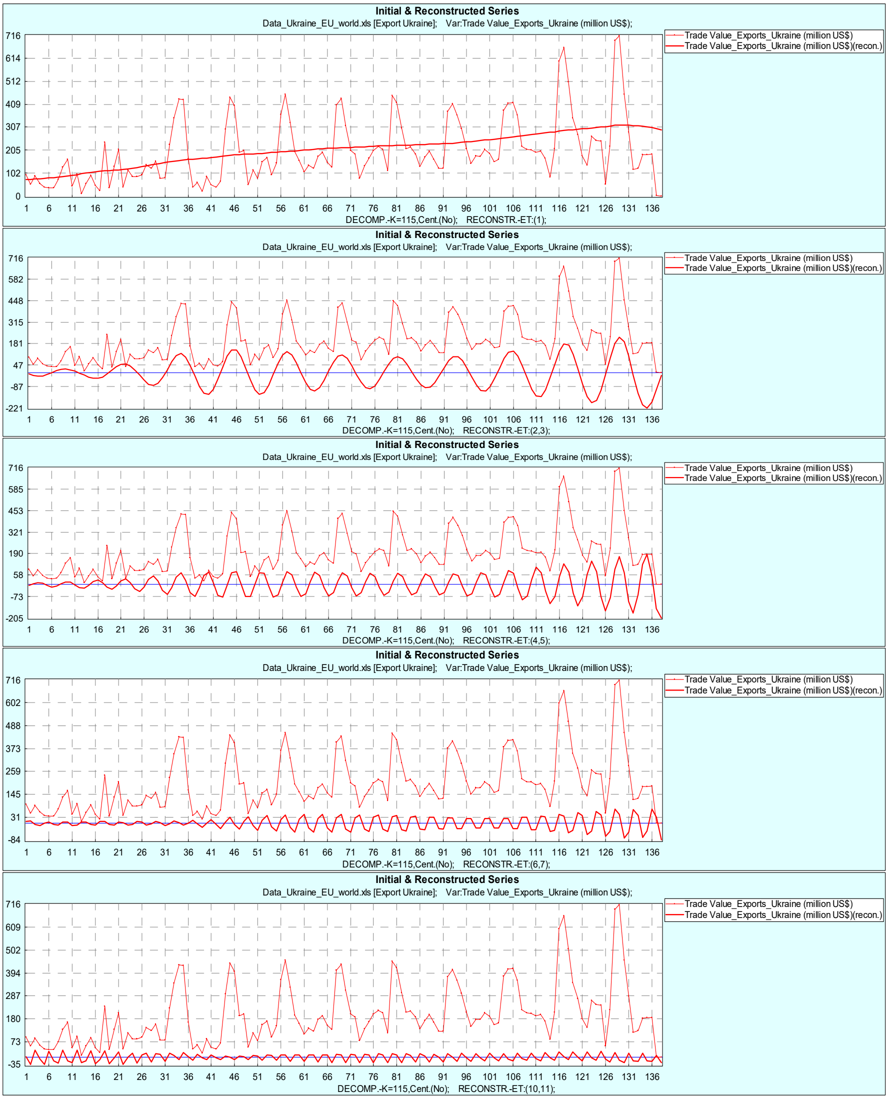


Figure 7. Raw data, reconstructed according to the selected principal components values of the time series of wheat export from Ukraine in January 2000– June 2021

Source: authors' calculation using SSA based on UN Comtrade Database.

The results of the analysis and interpretation of the principal components are presented in Table 9.

Table 9. Interpretation of the principal components

№ of principal components	Interpretation
1	Nonlinear trend
2-3	Annual periodicals
4-5	Semi-annual periodicals
6-7	Four-month periodicals
10-11	Quarterly periodicals

Source: authors' calculation using SSA based on UN Comtrade Database.

Figure 8 presents the outgoing, restored by 13 principal components of the value of a number of wheat export from Ukraine. The same figure shows the predicted values of the series.

Visual analysis of the restored values of the series (Figure 8), which are almost indistinguishable from the original, confirms the conclusion that the first thirteen principal components are enough to model the series.

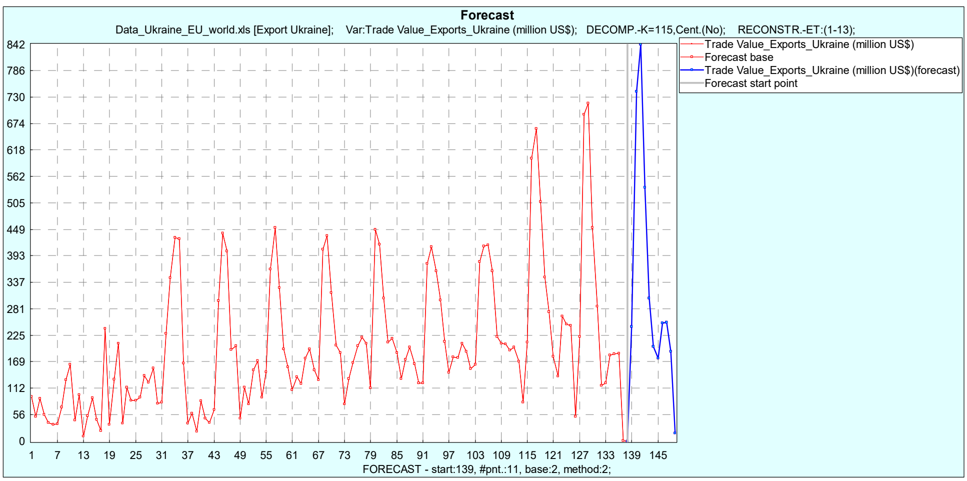


Figure 8. Raw data (January 2000–June 2021), reconstructed on the 1st–13th principal component and forecast values of wheat export from Ukraine (for the second half of 2021 with a forecast base of 138 months)

Source: authors' calculation using SSA based on UN Comtrade Database.

Analysis of the one-dimensional time series, their transformation into multidimensional series, their study using the principal components method and further reconstruction of the one-dimensional series allow modelling of selected individual components of the original series.

Similarly, the SSA method was used to forecast the average monthly export prices of wheat in Ukraine.

Raw time series, reconstructed on 13 principal components and forecast values of the average monthly export prices of wheat in Ukraine are given in Figure 9.

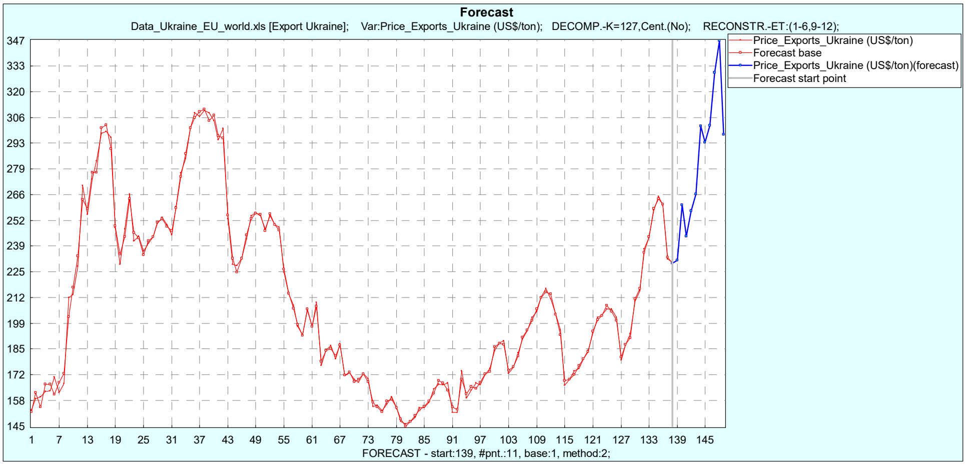


Figure 9. Raw data, reconstructed on the 1st–13th principal component and forecast values of the average monthly export price of wheat in Ukraine in January 2010–May 2022

Source: authors' calculation using SSA based on UN Comtrade Database.

Comparison of the retrospective forecast for 6 months (July–December 2021) with the updated data proves the high degree of accuracy of the applied model, in which the error of the forecast is 9.14% (Table 10).

Table 10. Forecast and actual values of time series of wheat export from Ukraine in July 2021–May 2022

Time period	Wheat export			Price		
	million US$		Forecast error %	US$ per ton		Forecast error %
	retrospective forecast	renewed data		retrospective forecast	renewed data	
2021 July	242.005	220.918	9.55	231.471	229.980	0.65
August	741.096	798.656	7.21	260.455	221.073	17.81
September	842.210	716.744	17.51	244.204	193.182	26.41
October	537.719	908.440	40.81	257.450	266.000	3.21
November	302.755	670.028	54.81	266.259	282.106	5.62
December	200.279	350.094	42.79	301.790	298.303	1.17

Table 10. Forecast and actual values of time series of wheat export from Ukraine in July 2021–May 2022 (cont.)

Time period	Wheat export million US$ retrospective forecast	Wheat export million US$ renewed data	Wheat export Forecast error %	Price US$ per ton retrospective forecast	Price US$ per ton renewed data	Price Forecast error %
2022						
January	175.316			293.104	306.959	4.51
February	250.021			302.054		
March	252.181			329.653		
April	189.433			346.584		
May	16.526			297.230		
MAPE			**28.78**			**9.14**

Source: authors' calculation using SSA based on UN Comtrade Database.

Accuracy degree of the obtained forecast is determined by statistical evaluation of MAPE (5) by comparison with updated data.

$$\text{MAPE} = \frac{100}{n}\sum_t \left|\frac{y_t - \hat{y}_t}{y_t}\right|, \tag{5}$$

where:

$\hat{y}_t$ is forecast value of the time series in the t-th period;

y_t is the actual value of the time series in the t-th period.

The study of the components of time series, such as trend and seasonal fluctuations indicates the need and feasibility of using their statistical estimation in forecasting volumes and prices of wheat export from Ukraine.

The calculated forecast values in July–December 2021 were compared with updated official statistics (Table 10) (UN Comtrade Database).

The calculated statistical estimates of the MAPE allow us to conclude that the models built using this method are highly adequate (Yerina, 2014).

The simulation illustrates that wheat and meslin price export from Ukraine has grown and will continue to grow. Moreover, over time, the rate of change of the studied indicator will also increase. It should also be noted that in the coming periods the average monthly export price of wheat and meslin is expected to increase further. The reliability of the forecast was also confirmed by calculating Henri Theil's coefficients, the so-called Theil's Inequality Coefficients (UI and UII), according to the formula (6) and (7) (see Theil, 1961; Theil, 1966; Cook, 2019).

The first specification for Theil's Inequality Coefficient is shown below:

$$\mathrm{UI}^2 = \frac{\sum(\hat{y}_t - y_t)^2}{\sum y_t^2} \text{ or } \mathrm{UI} = \frac{\sqrt{\sum(\hat{y}_t - y_t)^2/n}}{\sqrt{\sum y_t^2/n}} \tag{6}$$

where:

$\hat{y}_t$ is forecast values,

y_t is raw data.

Depending on the results obtained, the following conclusions can be drawn.

When U=0, the forecast is ideal, i.e. there is an absolute coincidence of actual and forecast values, i.e. $y_t = \hat{y}_t$ for all *t*.

When U=1, the forecast is bad.

When U>1 (this coefficient does not have an upper limit), the forecast is very bad.

Calculating Theil's Inequality Coefficient II, we can judge the severity of the forecast error:

$$\mathrm{UII} = \frac{\sqrt{\frac{1}{n}\sum(\hat{y}_t - y_t)^2}}{\sqrt{\frac{1}{n}\sum \hat{y}_t^2} + \sqrt{\frac{1}{n}\sum y_t^2}}. \tag{7}$$

The inequality coefficient calculated by formula (7) is in the range from 0 to 1. The closer the value of UII is to zero, the better is the forecast. In addition, the denominator of this coefficient is the sum of RMS predicted and introduced changes, and therefore the mismatch factor is determined not only by the standard error of the forecast in contrast to the coefficient formula (6).

That is, the inequality coefficient calculated by formula (6) will better reflect a completely inadequate forecast.

The value of this coefficient was obtained for a number of wheat export from Ukraine coefficient *U*=0.3478. For a number of wheat prices *U*=0.1089.

As you can see, these values are close to zero. According to Theil's theory, this means that the forecast of wheat export from Ukraine is reliable, the price forecast is almost perfect.

Thus, the calculated Theil's Inequality Coefficient (UI) confirmed the conclusion about the very high accuracy of the obtained predicted values.

To estimate the contribution of each component of the level of the series, it is necessary to decompose the series into components so that each component is positive. The use of seasonal wave models based on harmonic analysis is common in the presence of periodic time series fluctuations (Golyandina, 2013; Golyandina, 2018; Hassani, 2007). The dispersion ratio is an estimate of the contribution of harmonics that characterize seasonal waves taking into account the Fourier series expansion.

6. Calculation of the structure of time series by additive components

It is not enough to characterize the contribution of seasonal component fluctuations only to the total variance of the process, because this provision is more correct to use for stationary processes. For non-stationary processes, which are mainly time series that characterize the various elements of the market system, this provision is interpreted ambiguously.

It is proposed to move the coordinate system so that the minimum value of each seasonal component is zero. When transferring the coordinate system for each of the seasonal components there is a difference equal to the deviation of the smallest negative value of the component from zero. To maintain identity, the sum of the modules of these differences must be added to the values of the series restored for the selected components.

In view of the above, the team of authors of this article proposed the following algorithm for calculating the percentage structure of the reconstructed levels of the time series for the selected additive components:

1) use of the Singular Spectrum Analysis (SSA) method;
2) analysis of the principal components and their interpretation;
3) selection of additive components of the raw time series;
4) reconstruction of values of separate components according to the allocated principal components;
5) finding the minimum values of each of the seasonal components;
6) adding the modulus of these values to the corresponding level restored by the component series;
7) calculation of the modulus of the sum of the minimum values of all seasonal components;
8) increasing the corresponding levels of the series reconstructed on the chosen principal components, on the module of the sum of the minimum values of all seasonal components;
9) calculation of the share of components in the total value of the reconstructed series.

Taking into account the proposed algorithm, the formula (2), as well as the fact of the absence in the studied series of cyclic component series is presented as follows:

$$\widehat{\mathrm{Y}}_t + \sum_{i=1}^{m} |\min S_{ti}| = \mathrm{U}_t + \sum_{i=1}^{m} (\mathrm{S}_{ti} + |\min S_{ti}|) + \xi_t. \tag{8}$$

The contribution of the i-th seasonal component to the t-th level of the series is proposed to be calculated according to the formula (9):

$$w_{S_{ti}} = \frac{\mathrm{S}_{ti} + |\min S_{ti}|}{\widehat{Y}_t + \sum_{i=1}^{m} |\min S_{ti}|}. \tag{9}$$

The average contribution of the i-th seasonal component is proposed to be calculated according to the formula (10):

$$\overline{w_{S_i}} = \frac{\sum_{t=1}^{n} \frac{S_{ti} + |\min S_{ti}|}{\hat{Y}_t + \sum_{i=1}^{m} |\min S_{ti}|}}{n}. \quad (10)$$

The average contribution of the trend component is proposed to be calculated according to the formula (11):

$$\overline{w_U} = \frac{\sum_{t=1}^{n} \frac{U_t}{\hat{Y}_t + \sum_{i=1}^{m} |\min S_{ti}|}}{n}. \quad (11)$$

The structure of the time series of wheat export from Ukraine calculated according to the proposed algorithm according to the components selected during the singular spectral analysis is given in Table 11.

Table 11. Structure of components of the time series of wheat export from Ukraine

Time period	Percentage share of additive components of the reconstructed time series of wheat export from Ukraine, %					
	trend	annual	semi-annual	4-month	3-month	others
2010	12.4	33.6	27.0	10.8	4.7	11.5
2011	16.6	33.6	26.6	10.7	4.5	8.0
2012	19.1	29.6	23.1	9.6	4.9	13.7
2013	24.2	29.2	24.0	9.9	4.6	8.1
2014	25.5	27.3	22.6	9.3	4.7	10.6
2015	27.1	27.2	22.3	9.1	4.5	9.8
2016	27.9	26.9	22.0	8.9	4.4	9.9
2017	29.1	27.1	21.8	8.9	4.1	9.0
2018	31.0	26.5	21.1	8.7	4.0	8.7
2019	33.3	24.8	19.5	8.4	4.3	9.7
2020	36.4	24.5	18.6	8.1	3.8	8.6
2021	43.9	12.3	23.1	9.5	4.0	7.2
2010–2021	26.5	27.4	22.7	9.3	4.4	9.7

Source: authors' calculation using SSA decomposition based on UN Comtrade Database.

Graphic representation of the structure of the time series of export of wheat and meslin from Ukraine by selected components is provided in Figure 10.

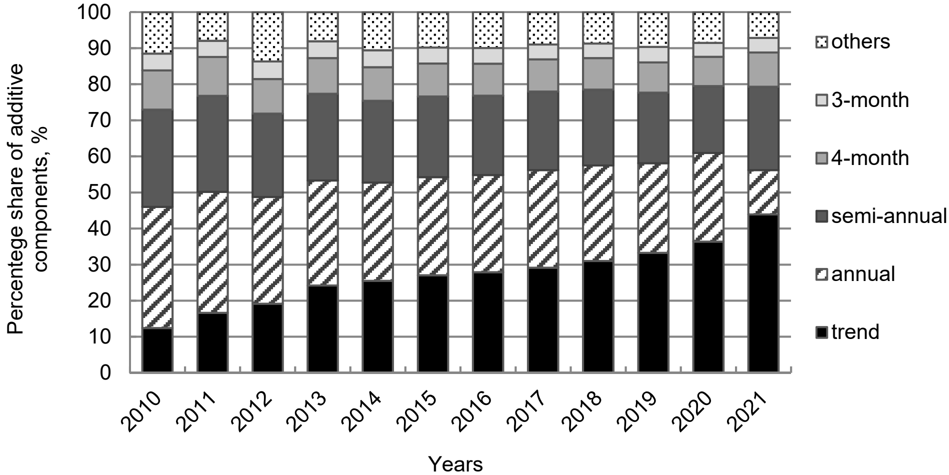


Figure 10. Percentage share of additive components of the reconstructed time series of wheat export from Ukraine in 2000–2021

Source: authors' calculation using SSA decomposition based on UN Comtrade Database.

Analysing the results obtained, we can conclude that the share of the trend has increased significantly (from 12.4% in 2000 to 43.9% in 2021), reducing the share of the annual component (from 33.6% in 2000 to 12.3% in 2021). The shares of the semi-annual, 4-month, and 3-month seasonal components of the levels of the considered time series in 2000–2021 did not change significantly.

Features of using the proposed approach:

– the average structure of the time series for the selected components corresponds to a certain time period;
– the decomposition of the time series into additive components will differ depending on the modelling method and, accordingly, the structure by components will be different.

7. Conclusions

Thus, studying the practice of applying forecasting methods is an important component of managerial activity at all levels of management. Statistical forecasting is considered as one of the methods of predicting the effectiveness of decision-making within the framework of international economic relations, especially in global agribusiness.

The conducted modelling and forecasting illustrates that the volumes of wheat and meslin exports from Ukraine grew over the past decade and will continue to grow in the future. Moreover, over time, the rate of change of the studied indicator will also increase. It should also be noted that further growth in the average monthly export price of wheat and meslin is expected in the coming months.

It has been established that domestic transport and logistics infrastructure, especially during martial law, are constraining factors for the export of wheat and meslin. The contradiction between the pace of development of the wheat market and the transport and logistics infrastructure is becoming an urgent problem and needs to be solved both at the national and international levels in order to ensure global food security. To stabilize the situation and support agribusiness, as recommendations to grain traders, state and local authorities, it is necessary to significantly increase the volume of grain storage capacities (for example, with the use of hermetic plastic sleeves).

The results of the study can also be implemented in the practice of forecasting the export-import potential of countries in global agribusiness. So, for example, it is recommended to conclude long-term international contracts in agribusiness for those products in the composition of time series of which (when decomposed into additive components) the trend has the largest share.

Authors' further research directions within this explored issue are related to the scientifically based selection of the main generalized factors of influence on the volumes and prices of exports of agribusiness products by regions of the world (using principal component analysis and correlation-regression analysis).

References

Agricultural markets and international trade, (2022). EU Science Hub. URL: https://joint-research-centre.ec.europa.eu/scientific-activities-z/agricultural-markets-and-international-trade_en

Armstrong, J., Green, K., (2018). Forecasting methods and principles: Evidence-based checklists. *Journal of Global Scholars of Marketing Science*, *28*(2), pp. 103–159. doi: 10.1080/21639159.2018.1441735.
URL: https://www.researchgate.net/publication/323754973_Forecasting_methods_and_principles_Evidence-based_checklists

Cameron, A., (2022). Agricultural overview: March quarter 2022. URL: https://www.awe.gov.au/abares/research-topics/agricultural-outlook/agriculture-overview

Cook, S., (2019). Forecast Evaluation using Theil's Inequality Coefficients. The Economic Network, University of Bristol.
URL: https://www.economicsnetwork.ac.uk/showcase/cook_theil.

Dama, F., Sinoquet, C., (2021). Time Series Analysis and Modeling to Forecast: a Survey. *arXiv preprint arXiv:2104.00164*. URL: https://arxiv.org/pdf/2104.00164.pdf

Durisin, M., Parija, P., Anghel, I., (2022). The $120 Billion Global Grain Trade is Being Redrawn by Russia's War in Ukraine. URL: https://www.bloomberg.com/news/features/2022-04-05/will-russia-s-war-in-ukraine-cause-wheat-shortages-raise-food-prices-more

Europe and Central Asia Economic Update, Spring 2022 : War in the Region, (2022). URL: https://openknowledge.worldbank.org/handle/10986/37268

Flores, A., Solis, Sheila Y., Álvarez-Aros, É. L., (2020). Contextualization of economic activity and innovation perspective in Reynosa, Mexico: A shift-share analysis 2004-2014. *Revista ESPACIOS*. ISSN 0798 1015. 2020. Vol. 41 (Nº03), 21 p. URL: http://www.revistaespacios.com/a20v41n03/a20v41n03p21.pdf

Golyandina, N., Korobeynikov, A., Zhigljavsky, A., (2018). Singular Spectrum Analysis with R. Use R! Springer, Berlin, Heidelberg. URL: https://vdoc.pub/documents/singular-spectrum-analysis-with-r-5e7h239knpf0

Golyandina, N., Zhigljavsky, A., (2020). Singular Spectrum Analysis for Time Series. Springer Nature, 146 p. URL: https://books.google.lv/books?id=BwgLEAAAQBAJ&dq=GOLYANDINA,+N.,+KOROBEYNIKOV,+A.,+ZHIGLJAVSKY,+A.+(2018).+SSA+Analysis+of+One-Dimensional+Time+Series&hl=uk&source=gbs_navlinks_s

Golyandina, N., Zhigljavsky, A., (2013). Singular spectrum analysis for time series. Springer briefs in statistics. Springer. URL: http://doi.org/10.1007/978-3-642-34913-3

Hassani, H., (2007). Singular Spectrum Analysis: Methodology and Comparison. *Journal of Data Science*. Volume 5, pp. 239–257. URL: https://www.jds-online.com/files/JDS-396.pdf

Isard, W., (1960). Notes on the Use of Regional Science Methods in Economic History. *The Journal of Economic History*, 20 (4). pp. 597–600. URL: https://doi:10.1017/s0022050700109878

Kozera-Kowalska, M., Uglis, J., (2021). Agribusiness as an Attractive Place to Work – A Gender Perspective. URL: https://www.mdpi.com/2077-0472/11/3/202/htm

Kwas, M., Paccagnini, A. and Rubaszek, M., (2022). Common factors and the dynamics of cereal prices. A forecasting perspective. *Journal of Commodity Markets*, 100240. https://doi.org/10.1016/j.jcomm.2021.100240. URL: https://www.sciencedirect.com/science/article/abs/pii/ S2405851321000738

Maeda, E., Abera, T., Siljander, M., Heiskanen, J., (2021). Large-scale commodity agriculture exacerbates the climatic impacts of Amazonian deforestation. URL: https://www.pnas.org/doi/10.1073/pnas.2023787118

Obrębalski, M., Walesiak, M., (2015). Functional structure of polish regions in the period 2004-2013 – measurement via HHI index, Florence's coefficient of localization and cluster analysis. *STATISTICS IN TRANSITION new series*. 2015. Vol. 16, No. 2, pp. 223–242. URL: https://www.exeley.com/statistics_in_transition/pdf/10.21307/stattrans-2015-012

Petropoulos, F., Apiletti, D., Assimakopoulos, V., at al, (2022). Forecasting: theory and practice. *International Journal of Forecasting*. Vol. 38, pp. 705–871. doi: https://doi.org/10.1016/j.ijforecast.2021.11.001. URL: https://www.sciencedirect.com/science/article/pii/S0169207021001758

RAMON – Reference and Management of Nomenclatures. Eurostat's Metadata Server. URL: https://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl= DSP_PUB_WELC

Rubalcaba, L., Gallego, J., Gallo, M. T. and Garrido, R., (2013). Business services location and market factors in major European cities. *Cities*. Vol. 31, pp. 258–266. doi: http://dx.doi.org/10.1016/j.cities.2012.06.022. URL: https://www.pmf.unizg.hr/_download/repository/Business_service_ location.pdf

Sayago-Gomez, J., T., Stair, C., (2017). Location Quotient, Coefficient of Specialization and Shift-Share. Applications on Matlab. *Regional Research Institute*. URL: https://researchrepository.wvu.edu/cgi/viewcontent.cgi?article= 1005&context=rri_tech_docs

State Customs Service of Ukraine. BI – foreign trade Indicators of Ukraine. URL: https://bi.customs.gov.ua/uk/trade/

State Statistics Service of Ukraine, (2022). "Rehiony Ukrainy 2021. Statystychnyj zbirnyk" [Regions of Ukraine 2021. Statistical yearbook] (in Ukrainian), section 2, Derzhkomstat, Kyiv, Ukraine. URL: https://ukrstat.org/uk/druk/publicat/kat_u/2021/zb/12/Regionu_20_pdf.zip

Sukhanova, E., Shirnaeva, N, Zaychikova, N., (2019). Modeling And Forecasting Financial Performance Of A Business: Statistical And Econometric Approach. *TheEuropean Proceedings of Social & Behavioural Sciences.* pp. 487–496. doi: https://dx.doi.org/10.15405/epsbs.2019.03.48. URL: https://www.researchgate.net/publication/331905290_Modeling_And_Forecasting_Financial_Performance_Of_A_Business_Statistical_And_Econometric_Approach

Suvoryj eksport 2022: iak pid chas vijny prodaty ukrains'ke zerno ta zapobihty prodovol'chij katastrofi u sviti? [Strict export 2022: how to sell Ukrainian grain during the war and prevent food catastrophe in the world?] (in Ukrainian). URL: https://latifundist.com/spetsproekt/969-suvorij-eksport-2022-yak-pid-chas-vijni-prodati-ukrayinske-zerno-ta-zapobigti-prodovolchij-katastrofi-u-sviti

Theil, H., (1966). Applied Economic Forecasting. Study in Mathematics & Managerial Economics, v. 4. Amsterdam, North-Holland Pub. Co.; Chicago, Rand McNally. 503 p. URL: https://www.goodreads.com/book/show/ 40666634-applied-economic-forecasting

Theil, H., (1961). Economic forecasts and policy. Amsterdam. The Netherlands: North-Holland, 567 p. URL: http://garfield.library.upenn.edu/classics1990/ A1990CY29300001.pdf

UCGFEA – Ukrainian Classification of Goods for Foreign Economic Activity (2020). Verkhovna Rada. URL: https://zakon.rada.gov.ua/laws/show/674-20#Text

UN Comtrade Database. United Nations. Department of Economic and Social Affairs. Statistics Division. Trade Statistics. URL: https://comtrade.un.org/data/

Wang, S., & Chaovalitwongse, W., (2011). Evaluating and comparing forecasting models. *Wiley Encyclopedia of Operations Research and Management Science*. pp. 1–16. doi: 10.1002/9780470400531.eorms0307. URL: https://www.researchgate.net/publication/313991989_Evaluating_and_ Comparing_Forecasting_Models

Wickham, H., Chang, W., (2017). Devtools: Tools to Make Developing R Packages Easier. URL: http://CRAN.R-project.org/package=devtools

Yerina, A., (2001). *Statystychne modeliuvannia ta prohnozuvannia* [Statistical modeling and forecasting] (in Ukrainian): navch. posib., KNEU, Kyiv, Ukraine, 170 p.

Yerina, A., Yerin, D., (2014). *Statystychne modeliuvannia ta prohnozuvannia* [Statistical modeling and forecasting] (in Ukrainian): navch. posib., KNEU, Kyiv, Ukraine, 348 p.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 199–212, DOI 10.59170/stattrans-2023-011
Received – 01.07.2022; accepted – 05.10.2022

A comparative analysis of the principal component method and parallel analysis in working with official statistical data

Halyna Holubova[1]

ABSTRACT

The dynamic development of the digitized society generates large-scale information data flows. Therefore, data need to be compressed in a way allowing its content to remain complete and informative. In order for the above to be achieved, it is advisable to use the principal component method whose main task is to reduce the dimension of multidimensional space with a minimal loss of information.

The article describes the basic conceptual approaches to the definition of principle components. Moreover, the methodological principles of selecting the main components are presented. Among the many ways to select principle components, the easiest way is selecting the first k-number of components with the largest eigenvalues or to determine the percentage of the total variance explained by each component. Many statistical data packages often use the Kaiser method for this purpose. However, this method fails to take into account the fact that when dealing with random data (noise), it is possible to identify components with eigenvalues greater than one, or in other words, to select redundant components. We conclude that when selecting the main components, the classical mechanisms should be used with caution.

The Parallel analysis method uses multiple data simulations to overcome the problem of random errors. This method assumes that the components of real data must have greater eigenvalues than the parallel components derived from simulated data which have the same sample size and design, variance and number of variables.

A comparative analysis of the eigenvalues was performed by means of two methods: the Kaiser criterion and the parallel Horn analysis on the example of several data sets. The study shows that the method of parallel analysis produces more valid results with actual data sets. We believe that the main advantage of Parallel analysis is its ability to model the process of selecting the required number of main components by determining the point at which they cannot be distinguished from those generated by simulated noise.

Key words: principal components, principal component analysis, factor analysis, Kaiser criterion, parallel analysis, simulation

[1]National Academy of Statistics, Accounting and Audit. Kyiv, Ukraine. E-mail: g_kondrya@ukr.net. ORCID: https://orcid.org/0000-0003-4847-5235

© Halyna Holubova. Article available under the CC BY-SA 4.0 licence

1. Introduction

With the powerful development of the digital economy and the information society as a whole, large amounts of data are produced on a daily basis. Big data is a large data set generated by people using information and communication technologies. Currently, there is no single methodology for generating and summarizing big data that could be used as a universal information base (Osaulenko et al., 2021).

Since socio-economic phenomena and processes are characterized by multidimensionality, which generates large databases, there is a need to summarize, group and concisely identify this information. For the convenience of statistical analysis, it is necessary to determine the main factors or components that form and, accordingly, characterize the phenomenon under study.

For example, the Human Development Index covers a system of statistical indicators that can be summarized in several components: health indicators, education indicators, indicators of material well-being of the population. The assessment of the level of development of information and communication technologies (ICT), which is calculated and published annually by the International Telecommunication Union, is based on 11 indicators, which can be summarized in three sub-indices: availability of infrastructure and access to ICT; intensity of ICT use; ability to use ICT effectively (Korepanov, 2018).

The introduction of experimental statistics, the use of applied statistics methods, the transformation of alternative data sources (for example, departmental statistics) for the production of official statistical information will allow the harmonization of official statistics and achieve comparability of various statistical indicators with international comparisons, classifications, etc.

The most common method of information optimization is the principal components analysis (PCA), which allows one to organize a large array of data. In order to study the internal structure of the object, the dimension of the initial feature set should be compressed, replacing it with a minimum number of components (Ierina, 2014). The main components store all the information about the object of study, Figure 1.

Figure 1 clearly illustrates the transformation of the raw data of multidimensional space into principal components. Currently, there are various methods of implementing PCA. The question arises, which principal components extraction algorithm will work best with the official statistical data? It should be noted that statistical data are usually heterogeneous (there are atypical population units (outliers), a high value of standard deviation, skewness, etc.); are not always subject to the law of normal distribution; some statistical indicators may be incomparable or multidirectional, sometimes incomplete, etc. That is why the choice of the PC method is extremely important, since the formation of the PC serves as the ultimate goal –

for grouping, typology or clustering of data, and an intermediate goal – for advanced statistical research.

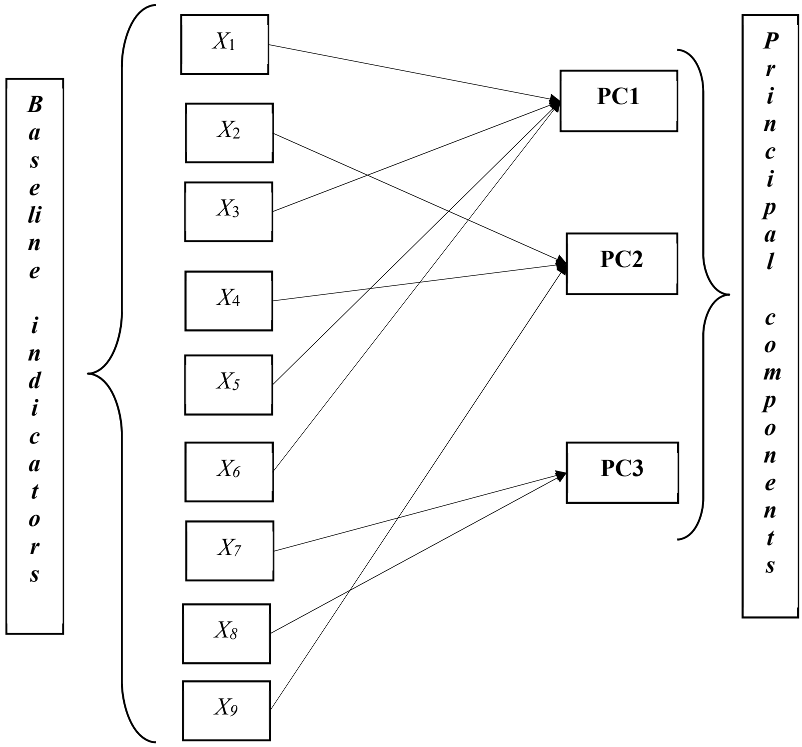


Figure 1. The relationship of the original features (X) and the principal components (PC)

2. Literature review

Applied aspects in the study of methodological principles in separation of the main components are revealed in the works of foreign scholars, including J. L. Horn (1965), H. F. Kaiser (1970), L. W. Glorfeld (1995), R. L. Gorsuch (1983), Ö. Çokluk & D. Koçak (2016), A. V. Silverstein (1987), W. R. Zwick & W. F. Velicer (1986) and many other specialists who studied the mechanism and algorithm in the formation of principal components using various mathematical methods. Domestic scientists use the method of the main components for grouping statistical data, to identify factors influencing the object under study (Holubova, 2013), (Lepeyko & Shcherbak, 2018), (Rosen et al., 2018) or as a tool of public administration (Chinkulyak & Pogrebnyak, 2015), etc.

3. Problem statement

The method of Principal component analysis (PCA) is a powerful research model with the main task to reduce the dimension of multidimensional space with minimal information loss.

This method is useful when working with official statistics, as it is able to group statistics both in statics and in dynamics. Based on the main components it is available to:

- rank and classify objects, countries, regions, enterprises, etc.;
- measure the relationship between primary indicators and key components;
- perform regression analysis, etc.

One of the advantages of using the method of principal components is the ability to get rid of multicollinearity between the original features and perform regression analysis on the principal components.

Assessing the socio-economic development of the country and its regions on a number of statistical indicators, it is possible to identify the economic, social, demographic, political component and so on. Based on the state of the environment according to official statistics, certain types of environmental risks are distinguished according to the degree of risk. By studying demographic statistics, it is possible to classify the population, for example, on type of aging, or to identify factors that affect the decline in birth rates in Ukraine and so on. Analysis of medical statistics data allows, for example, to identify classes of morbidity by age groups or to identify factors that shape the medical system in the country as a whole, and others. Therefore, this method is universal and can be used in various studies using data from both official and administrative statistics.

The PCA method was first proposed in 1901 by K. Pearson, who studied the problem of the best approximation of a set of points by lines and planes.

PCA should not be confused with factor analysis (FA). The latter is a popular method of detecting interpreted linear relationships between variables called factors. In factor analysis, different types of vector rotation are usually used to redistribute variation between factors, while maintaining the total variance of the selected factors. Determining the number of factors is more important than the type of rotation, because the power of factor analysis depends on the ability to distinguish important factors from others. Therefore, it is very important to determine the exact balance between correlations. Determining the number of factors requires close attention, because if the number of isolated factors is greater or less than necessary, it can lead to serious errors that affect the results of the study.

Instead, PCA is a useful method to reduce the number of observed variables to a smaller set of independent components. Therefore, the main goals of PCA are (Holubova, 2020):

1. Data visualization for research analysis, which allows to reveal the latent characteristics of data and interpret the components.
2. Decrease in the number of predictors for future analyzes, such as regression of major components.

4. Methodology

PCA uses elements of linear algebra to determine the basic linear structure inherent in a data matrix. The basis of mathematics in PCA is the decomposition of singular values, which is a generalization of the decomposition of eigenvalues (lambda numbers, λ). The intrinsic value of the principal component is the amount of deviation in the original data, so maximizing the deviation is important because it provides the most information about the actual data. Understanding how these mathematical combinations work is not necessary to understand PCA, but understanding the basic principles of the principal components method is essential when interpreting PCA results.

The study revealed the methodological principles of several methods for selecting the main components (PC). One of the simplest methods for selecting a subset of the PC is to select the first k-number of components with the largest eigenvalues λ. As a result, the main components that best explain the deviations from the data are selected.

The Scree plot stony decline involves the construction of a graph where the abscissa is plotted against the ordinal number of the eigenvalue, and the ordinate – its value. According to R. Cettel (1966), it is necessary to find the point of the greatest slowdown in the decline of eigenvalues and take into account only the factors that correspond to eigenvalues to the left of this point. This criterion is not statistically sound and often leaves not all significant factors in the model.

The criterion of Bartlett's xi-square tests the hypothesis that other eigenvalues are equal, that is, each eigenvalue λ is evaluated sequentially until the null hypothesis is rejected.

One of the classic methods of the PC selection is to study the percentage of total variance, which is explained by each component. Having set a predetermined threshold (usually 75% of the total variance explained), the first k-principal components that collectively explain this variance fraction can be selected as a subset of the components. However, this method of selection, like other methods described above, cannot fully take into account the variance of the data.

Many statistical data packets often use a method that preserves all PC with eigenvalues $\lambda > 1$. It is also called the Kaiser rule, the Kaiser test, or the Kaiser-Gutmann criterion. The basic idea is that with standardized data, the variance of each of the source variables is 1. Therefore, principal components with an eigenvalue of more than one explain more variance than one variable in the source data. This method is popular and practical, but does not take into account the fact that even with random data (noise) you can identify components with eigenvalues greater than one. In these situations, the variance explained by the components is not really useful, as it is due to accidental error or noise.

Parallel analysis (PA) uses multiple data simulations to overcome the problem of random error. The essence of this method is that non-trivial components from real data should have greater eigenvalues than parallel components derived from simulated data that have the same sample size, variance and number of variables. PA is also called the parallel analysis of Horn in honour of its creator J. Horn. The process of performing parallel analysis is based on the Monte Carlo method, namely, it is a simulation of a large number of data sets. Horn argued that the number of iterations should be sufficient, that is, to obtain the most objective results (for example, 1000 or more repetitions), although there are no strict limits. Experiments were recorded when the results did not show a significant difference between one simulation and one hundred iterations. Each simulated data set contains the same number of variables and observations as the original data. For each simulated variable, data are generated by constructing a sample from a multidimensional normal distribution, with the standard deviation equal to the standard deviation of the corresponding actual data variable.

Repeating the steps of n-times gives n-sets of eigenvalues with the calculation of average eigenvalues by sets. This leads to a single set of average eigenvalues $\bar{\lambda}$, with which the eigenvalues obtained from the actual data set are compared. During the development of the PA method, researchers made an assumption that the use of average eigenvalues is similar to setting the error rate of the first type I (α) at 0.50 (instead of the more acceptable level $\alpha = 0.05$), and this may lead to the existence of factors (extra components). With this in mind, L. W. Glorfeld (1995) and R. A. Harshman & J. R. Reddon (1983) proposed the use of a 95 percent threshold for eigenvalues generated from random data. This is also similar to setting α to 0.05, which is a more common standard for type I error. Eigenvalues from actual data are compared with the values of the 95th percentile of generated data, and not with the average eigenvalue $\bar{\lambda}$ (Hayton & Alllen, 2004).

Therefore, the obtained eigenvalue of the PC from the original data should be compared with the upper 95th percentile, calculated from the simulated data sets. If the eigenvalue from the source data is greater than the upper percentile of the simulated data, the component is selected, otherwise it is discarded. The idea is that due to a random error in the data (caused by sample size, sample design, etc.), the PCA generates some components with eigenvalues greater than one. In general, the first eigenvalues generated by noise data will grow with an increasing number of variables and fall with a decreasing number of observations. Preserving only those PC with eigenvalues that exceed the 95th percentile of the simulated eigenvalues ensures that the discrepancies explained by these PC are likely to represent real variance rather than noise variance. That is, parallel analysis is considered more useful in practice than the method of selection of principal components by the Kaiser criterion or other methods of selection.

A.V. Silverstein (1987) compared the Kaiser method and the method of parallel analysis on the example of 24 data sets, and it was found that parallel analysis gives better results. W. R. Zwick & W. F. Velicer (1986) conducted a study comparing five methods used to determine the factors (parallel analysis, the method of minimum mean partial correlation, the graph of Scree plot stony decline, the criterion of Bartlett's xi-square, Kaiser's test) taking into account different conditions (sample size, the number of variables and components and their factor loads, etc.). The researchers concluded that the parallel analysis is consistent with the actual data set used to determine the number of factors, with an accuracy of 92%.

5. The application examples

During the study, the author developed several different data sets that are publicly available on the Internet.

When implementing the principal components method, certain statistical preconditions should usually be followed. All variables should be quantitative (categorical variables are excluded from the analysis) and homogeneous, distribution is symmetrical, and the number of observations should prevail over the number of variables. However, depending on the type of study, there may be exceptions, for example, in medicine, chemistry, biostatistics and other sciences, including working with the real statistical data. Or, for example, in the conditions of laboratory tests or in expensive sample observations, when it is not possible to involve a sufficient number of respondents in the experiment, and so on.

Data set 1 (Glorfeld, 1995)

We have information from 500 Facebook (2016) users on 14 indicators: the number of daily posts; the number of posts per hour; the number of posts about personal life; free time; the number of users who subscribed to your page, the number of people who liked your page, the number of users who liked your photo; the number of comments, likes, distributions, etc. Among all these indicators that can be quantified, it is actually difficult to single out a priori the main components that clearly visualize or typify public activity on Facebook, or in some way can describe the principles of interaction with this social media. A comparative analysis of eigenvalues λ was performed by two methods: the Kaiser criterion and the parallel Horn analysis. The obtained results of eigenvalues are given in Table 1.

Table 1. Eigenvalues of the principal components

Principal components	Eigenvalues λ (initial data)	Eigenvalues λ (Parallel analysis)		
		Average	Upper limit	Lower limit
PC1	5.920	1.289	1.356	1.235
PC2	1.740	1.222	1.270	1.180
PC3	1.658	1.171	1.212	1.136
PC4	1.124	1.126	1.163	1.094
PC5	1.002	1.086	1.117	1.057
PC6	0.835	1.047	1.076	1.016
PC7	0.621	1.009	1.038	0.981
PC8	0.454	0.975	1.002	0.944
PC9	0.304	0.939	0.968	0.913
PC10	0.143	0.905	0.932	0.876
PC11	0.116	0.869	0.899	0.836
PC12	0.069	0.831	0.862	0.798
PC13	0.014	0.790	0.823	0.754
PC14	0.001	0.739	0.779	0.694

As we can see, according to the Kaiser criterion, five main components are selected with the values of more than one (Figure 2), which explains 81.2% of the variation. According to the method of Parallel analysis, only three main components are identified, as evidenced by Figure 3, which visualizes the clipping of three components. The intrinsic value of PC4 is 1.124, which is less than the upper limit of the 95 percent interval (1.163), which gives grounds to exclude this component from further analysis, because its variance is caused by sampling noise, not the real process.

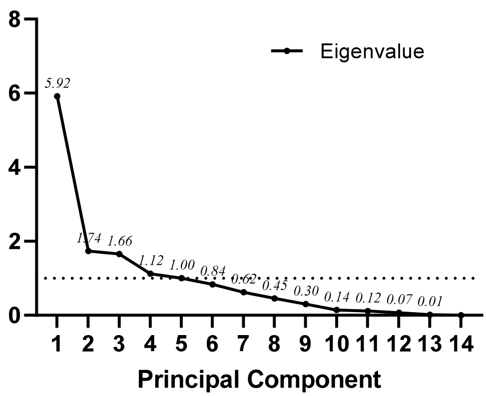


Figure 2. The principal components of the Kaiser criterion

Source: built by the author in the GraphPadPrism packet.

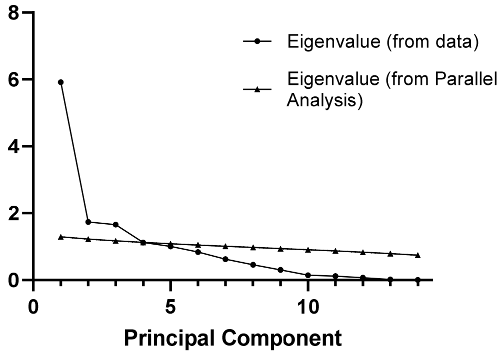


Figure 3. The principal components of the PA method

Source: built by the author in the GraphPadPrism packet.

Data set 2 (Gene Dataset)

For purposes of illustration, a sample of the human gene pool (20 patients and their characteristics of 100 genes, i.e. a matrix of 20 per 100) is considered. Since the variables *(m)* are greater than the observations *(n)*, the maximum number of components that can be selected is n-1, in our example - 19. According to the Kaiser criterion, it is established that there are 19 main components and one main component according to the method of Parallel analysis (Table 2, Figures 4, 5). In addition, the number of iterations (10, 100, 1000 and 5000 simulations were used) did not affect the result, i.e. the isolation of only one component is confirmed.

Table 2. Eigenvalues of the principal components

Principal components	Eigenvalues ƛ (initial data)	Eigenvalues ƛ (Parallel analysis)		
		Average	Upper limit	Lower limit
PC1	17.291	9.749	10.606	9.013
PC2	8.457	8.808	9.449	8.282
PC3	7.587	8.131	8.650	7.678
PC4	6.924	7.575	8.011	7.156
PC5	6.425	7.058	7.477	6.647
PC6	6.157	6.589	6.986	6.229
PC7	5.983	6.151	6.489	5.807
PC8	4.816	5.730	6.050	5.400
PC9	4.524	5.344	5.653	5.019
PC10	4.472	4.969	5.287	4.666
PC11	4.140	4.621	4.924	4.334
PC12	3.983	4.278	4.567	3.984
PC13	3.723	3.954	4.239	3.661
PC14	3.282	3.640	3.920	3.360

Table 2. Eigenvalues of the principal components (cont.)

Principal components	Eigenvalues λ (initial data)	Eigenvalues λ (Parallel analysis)		
		Average	Upper limit	Lower limit
PC15	3.117	3.318	3.600	3.037
PC16	2.745	2.989	3.284	2.712
PC17	2.452	2.686	2.964	2.389
PC18	2.201	2.365	2.655	2.046
PC19	1.720	1.990	2.324	1.651

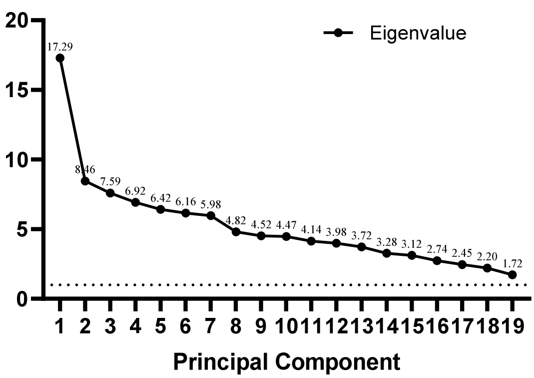


Figure 4. The principal components of the Kaiser criterion

Source: built by the author in the GraphPadPrism packet.

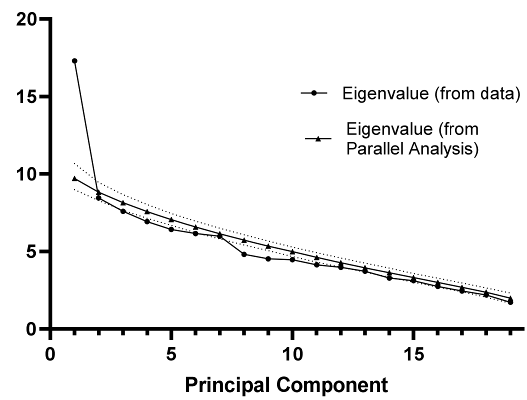


Figure 5. The principal components by the PA method

Source: built by the author in the GraphPadPrism packet.

It is claimed that the number of observations should exceed the number of signs at least twice.

Data set 3 (Decathlon Dataset, 2004).

The author considers a set of decathlon data. These are the results of the 41st athlete in 10 sports at the Olympic Games (2004). The initial data are symmetric, the coefficients of variation for each distribution (ten indicators) in the range of 3-8%, which indicates the homogeneity of the population and the reliability of the average value.

According to the Kaiser criterion, four principal components, which characterize 75% of the variation, are identified. PA allocates only one principal component with the number of simulations 100 and 1000, Table 3. If we use 10 iterations (i.e. only ten correlation matrices are modelled), then two principal components are distinguished. We can assume that this is exactly the case when the number of simulations matters (the more iterations, the more valid the results).

Table 3. Eigenvalues of the principal components

Principal components	Eigenvalues λ (initial data)	Eigenvalues λ (Parallel analysis)		
		Average	Upper limit	Lower limit
PC1	3.272	1.853	2.128	1.642
PC2	1.737	1.559	1.748	1.402
PC3	1.405	1.350	1.493	1.214
PC4	1.057	1.168	1.298	1.061
PC5	0.685	1.016	1.126	0.908
PC6	0.599	0.874	0.972	0.755
PC7	0.451	0.729	0.842	0.623
PC8	0.397	0.605	0.713	0.505
PC9	0.215	0.477	0.581	0.383
PC10	0.182	0.344	0.454	0.240

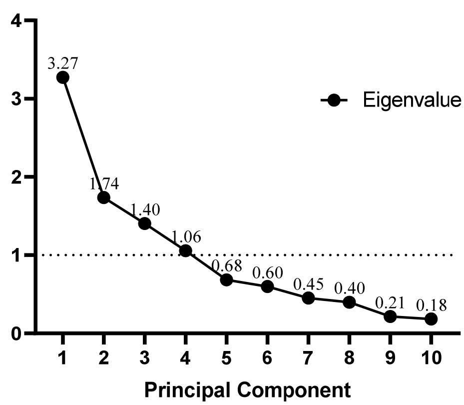


Figure 6. The principal components of the Kaiser criterion

Source: built by the author in the GraphPadPrism packet.

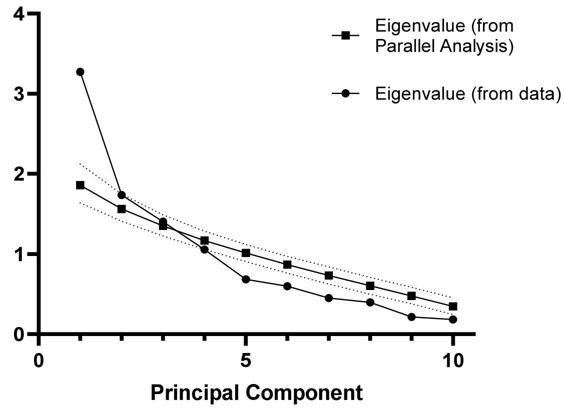


Figure 7. The principal components of the PA method

Source: built by the author in the GraphPadPrism packet.

The results of the study showed that the method of parallel analysis gives consistent results with the actual data sets, taking into account the sample size, symmetry of its distribution, variance and so on. We believe that the method under study gives more objective results in determining the exact number of factors.

6. Conclusions

The data sets used by the author for analysis are not official statistical data. However, these dates clearly characterize the peculiarities of different samples: Facebook Dataset is heterogeneous and has outliers; Gene Dataset is unbalanced in terms of the number of indicators and observations; Decathlon Dataset is indicative, at first glance (homogeneous and symmetrical). On the basis of the Kaiser method, which belongs to classical methods, redundant factors were selected in each of the data sets. According to the results of the Parallel analysis, which is based on multiple simulations, the real number of the main components was determined. Therefore, in our opinion, classical methods of selection of major components should be used with caution. Especially, the statistical data have certain features (heterogeneity, asymmetry, imbalance, etc.), that is why the author considers it appropriate to use Parallel analysis in the context of working with the official statistical data.

The main advantage of Parallel analysis should be the ability to model the process of selecting the number of PCs by determining the point at which the principal components cannot be distinguished from those generated by simulated noise. In our opinion, multiple simulations can protect against erroneous results.

References

Cattell, R. B., (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, pp. 245–276.

Chinkulyak, N. M., Pogrebnyak, L. O., (2015). Statystychnyi analiz yak instrument derzhavnoho upravlinnia [Statistical analysis as a tool of public administration] *Derzhavne upravlinnia – Governance*, 1, pp. 82–88, [in Ukrainian].

Çokluk, Ö., Koçak, D., (2016). Using Horn's parallel analysis method in exploratory factor analysis for determining the number of factors. Educational Sciences: *Theory* & *Practice*, 16, pp. 537–551.

Decathlon Dataset, (2004). Retrieved from: https://malouche.github.io/data_in_class/decathlon_data.html.

Facebook Dataset, (2016). Machine learning repository. Retrieved from: https://archive.ics.uci.edu/ml/datasets/Facebook+metrics.

Gene Dataset. Machine learning repository. Retrieved from: https://archive.ics.uci.edu/ml/datasets.php?format=&task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=table.

Glorfeld, L. W., (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, 55, pp. 377–393.

Gorsuch, R. L., (1983). Factor analysis (2nd ed.). Hillsdale, NJ: LawrenceErlbaum.

Harshman, R. A., Reddon, J. R., (1983). Determining the number of factors by comparing real with random data: A serious flaw and some possible corrections. *Proceedings of the Classification Society of North America at Philadelphia*, pp. 14–15.

Hayton, J., Alllen, D., (2004). Factor Retention. Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. Retrieved from: https://www.researchgate.net/publication/235726204_Factor_Retention_Decisions_in_Exploratory_Factor_Analysis_A_Tutorial_on_Parallel_Analysis/link/5582a85008ae6cf036c1a886/download.

Holubova, H. V., (2013). Statystychnyi analiz osnovnykh faktoriv vplyvu na tranzyt vantazhiv v Ukraini [Statistical analysis of basis factor influence on in-transit freight in Ukraine by regression model] Visnyk Kyivskoho natsionalnoho universytetu im. T. Shevchenka. Ekonomika – Bulletin of Taras Shevchenko National University of Kyiv. *Economics*, 134, pp. 12–16, [in Ukrainian].

Holubova, H. V., (2020). Pryntsypy vyboru holovnykh komponent: osoblyvosti prykladnoho modeliuvannia [Principles of choice of main components: features of applied modeling] Novi dzherela ta metody poshyrennia danykh u statystytsi: materialy XVIII Mizhnarodnoi naukovo-praktychnoi konferentsii z nahody Dnia pratsivnykiv statystyky]. New sources and methods of data dissemination in statistics: materials of the XVIII International scientific-practical conference on the occasion of the Day of Statistics. Kyiv, Informatsiino-analitychneahenstvo. *Information and Analytical Agency*, pp. 155–160, [in Ukrainian].

Horn, J. L., (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 32, pp. 179-185.

Ierina, A. M., (2014). Komponentnyi analiz [Component analysis] Statystychne modeliuvannia ta prohnozuvannia [*Statistical modeling and forecasting*], 348, pp 287–313.

Kaiser, H. F., (1970). A second-generation Little Jiffy. *Psychometrika*, 35, pp. 401-415.

Korepanov, O. S., (2018). Metodolohiia indeksnoho analizu rivnia rozvytku informatsiinoho suspilstva [Methodology of index analysis of the level of development of the information society]. *Statystyka Ukrainy – Statistics of Ukraine*, 1, pp. 6–15, [in Ukrainian].

Lepeyko, T., Shcherbak, A., (2018). Determining factors to ensure the effective formation of the information process in the industrial enterprise management. *Development Management*, 16(4), pp. 88–97.

Osaulenko, O., Holubova, H. and Horobets, O., (2021). Implementing Bid Data in the Public Administration. Stratehiia rozvytku Ukrainy: finansovo-ekonomichnyi ta humanitarnyi aspekty: materialy VIII Mizhnarodnoi naukovo-praktychnoi konferentsii – Strategy of development of Ukraine: financial and economic and humanitarian aspects: materials of the VIII International scientific-practical conference. Kyiv, Informatsiino-analitychne ahenstvo, pp. 219–222, [in English].

Rosen, V. P., Reutsky, M. O. and Demchik, Ya. M., (2018). Zastosuvannia metoda holovnykh component dlia identyfikatsii holovnykh faktoriv vplyvu na velychynu elektrospozhyvannia [Application of the principal components method to identify the main factors influencing the amount of electricity consumption]. *Enerhetyka: ekonomika, tekhnolohii, ekolohiia: naukovyi zhurnal – Energy: economics, technology, ecology: a scientific journal*, No. 3 (53), pp. 81–87, [in Ukrainian].

Silverstein, A. B., (1987). Note on the parallel analysis criterion for determining the number of common factor or principal components. *Psychological Reports*, 61, pp. 351–354.

Zwick, W. R., Velicer, W. F., (1986). Comparison of five rules for determining the numberof components to retain. *Psychological Bulletin*, 99(3), pp. 432–442.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 213–228, DOI 10.59170/stattrans-2023-012
Received – 01.07.2022; accepted – 26.10.2022

A new industrial strategy for Europe – new indicators of the results of its implementation

Oleg Krekhivskyi[1], **Olena Salikhova**[2]

ABSTRACT

This paper discusses the experiences resulting from EU's adoption and implementation of a wide variety of policy measures in response to the COVID-19 crisis. These measures included stimulating the relocation and expansion of manufacturing to reduce vulnerability, depending on imports, ensuring the stability and development of industrial production. Using the example of the pharmaceutical industry in EU27 in the years 1995–2018, the study proposes and tests a new approach to assessing the consequences of relocation policies aimed at developing the local production potential, increasing the value added by activity, and expanding the share of local value added in industry exports. Specifically, the focus is on the formation of statistical analysis tools for assessing the changes of the specialisation and identifying the country's comparative advantages. The authors propose new indicators: RSP – coefficient of Revealed Specialisation of Production, CAVA – coefficient of Comparative Advantage in Value Added by Activity and EVA – coefficient of Comparative Advantages in the Domestic Value Added Exports. Additionally, formulas for their calculation are provided which allow the assessment of the position of Ukraine's industries among a reference group, widening the 'revealed comparative advantage' concept. Finally, a test of the new methodology showed that it can be used to identify the comparative advantages of EU member states supported by state assistance programmes involving the implementation of business projects which aim to develop domestic production.

Key words: indicators, specialization, comparative advantage, industrial production, value added activity, export.

1. Introduction and motivation

The COVID-19 pandemic has affected industrial production in all countries of the world more than any economic shock that has occurred over the past few decades (such as the crises of 1997-1998 and 2008-2009). The closure of foreign counterparty

[1] Secretariat of Cabinet of Ministers of Ukraine, Ukraine. E-mail: o.krekhivskyi@gmail.com. ORCID https://orcid.org/0000-0002-4793-851X.
2 Institute for Economics and Forecasting of NAS of Ukraine, Ukraine. E-mail: salikhova_elena@ukr.net. ORCID: https://orcid.org/0000-0002-7669-6601.

© Oleg Krekhivskyi, Olena Salikhova. Article available under the CC BY-SA 4.0 licence

plants due to a lockdown; the introduction of a ban by the governments of individual countries on the export of goods of strategic importance; the blocking of air and sea transportation caused a shortage in the market of both intermediate and finished industrial goods (Barlow et al., 2021; Kumar et al., 2020; Vo et al., 2021). As the study showed (Tirivangani T. et al., 2021), weak pharmaceutical systems and supply chains turned out to be the most vulnerable in the control of COVID-19 pandemic, that has mostly devastated public health systems and livelihoods in resource-limited countries. The concentration of active pharmaceutical ingredients (API) production and certain medicines in the Asia-Pacific region also has led to interruptions in European health care systems. The crisis situation has clearly demonstrated European's pharmaceutical industries unpreparedness to face such challenges and threats (European Commission, 2020a). Analysis of the root causes of the crisis phenomena in industry has shown that they are a consequence of the delocalization of industrial production in Europe, which has taken place over the past few decades (European Parliament, 2020). The closure of many Europe-based facilities for the production of drugs, health products, API and other intermediate goods has made the health care system and industry hostage to a virtually single producer and supplier of certain goods - on China, significantly weakening the sovereignty of EU member states. The technological, human and manufacturing resources available in the EU are capable of expanding API production and therefore do not have a technological dependency. However, with globalization and offshoring, as well as the active policies of the Chinese government (Wang, 1999), such form of industry dependency developed, which can be called "component dependency".

Its implementation prompted the EU leadership to adopt a number of strategic documents focused on reducing the vulnerability and economy's dependence (Grumiller, 2021). In November 2020, the Pharmaceutical strategy for Europe was adopted, which should ensure synergy with the relevant EU policies on R&D, innovation and industrial production, in accordance with priorities of a New industrial strategy for Europe, adopted in March 2020. The latter was adjusted in May 2021 based on new challenges of the COVID-19 pandemic. In particular, among its new priorities is ensuring sustainability and reducing the dependence of industrial ecosystems, among them – health care, the foundation of which is pharmaceuticals (European Commission, 2020b, 2021). These are being implemented, inter alia, through strengthening the location-specific advantages (the so-called L-advantages) in accordance with the paradigm of OLI-advantages of J. Dunning (Dunning, 1993).

For implementation of priorities, the European Commission introduced, among other things, an investment assistance mechanism for the production of products related to COVID-19 (which allows compensating up to 80% of business costs for the project). This category of products includes not only those related to pharmaceuticals, such as medicines (including vaccines) and their intermediates, APIs and raw materials, but also products of other industries: medical devices, medical equipment and necessary accessories for them; disinfectants and their intermediates, chemical raw materials, etc. Germany, France, Poland and a number of other countries interested in developing their own industry have received the consent of the European Commission to provide financial assistance to companies for R&D projects related to COVID-19, investments in the creation of appropriate research infrastructure, as well as investments in production facilities related to COVID-19 (European Commission, 2020c, 2020d, 2020e). It is obvious that the new EU state aid mechanisms, as well as decisions of business on the implementation of investment and innovation projects will form new capacities in European countries: contribute to increasing gross output by industry, to creating high added value, to raising exports. No doubt, these processes are going to influence structural shifts and specialization in the economy not only in Europe, but also in the world, and therefore new indicators are needed to adequately assess the impact of such policies.

The measurement of post-coronavirus economic transformations is a particular challenge for statisticians. Eurostat and the National Statistical authorities of the European statistical system have developed a number of guidelines and methodological notes in the context of the COVID-19 crisis for monitoring economic processes (Eurostat, 2021a). At the end of 2020, Eurostat launched a new tool related to the COVID-19 crisis in various EU member states. It contains monthly and quarterly indicators on a number of statistical areas that are important for tracking economic and social recovery after the pandemic. These are about 20 indicators covering such aspects as macroeconomic changes, business (including industrial production) and trade, as well as the labour market (Eurostat, 2021b). At the same time, in our view, new approaches are also needed to assess the impact of measures (taken by the EU and Member States) aimed at developing local productive capacities, adding value and increasing the share of local value added in exports, strengthen competitive advantage.

The object of this paper is to provide a toolkit for assessing the positions of countries in the international division of labour: changes in the specialization, in the creation of value added and exports by relying on local resources and location advantages of industries in EU. To analyse the current and future advantages of EU member industries, the author's tools is proposed (the foundation of this tool is laid in the work devoted to new approaches to determining international comparative advantages (Salikhova, 2012).

3. Methodical approaches

Traditionally, RCA (Revealed Comparative Advantage) has been used to measure the comparative advantages identified, calculated from a formula proposed by B. Balassa in 1961 (Balassa, B. 1961, 1989; Dunning, 1992). A number of studies indicate that the development of foreign trade is a proof of the formation of international competitive position and at the same time international competitiveness of the national economy, its key industries, including pharmaceuticals (Falkowski, 2018a, 2018b; Motoryn et al., 2020).

In essence, the approach to calculating RCA is based on the concept of advantages in the Ricardian understanding of economics, which is inspired by free trade and perfect competition. But the modern world economy, and especially activities in fostering local pharmaceutical industry in developing countries, is based on other principles. The market of medicines is characterized by imperfect competition; states intervene in economic processes through various regulators, incentives, covert protectionist instruments. A number of countries have established strong pharmaceutical industries with a high export potential thanks to TNC investment and the transfer of technologies. The classical RCA in that case allows to estimate more visible than Revealed (or real) comparative advantages in the trade of goods, because the indicator does not take into account the local value added in export.

To assess the advantages of the location of firms and industries in the country, the authors propose to introduce into scientific circulation the coefficient of *Revealed Specialization of the country's Production* (*RSP*). *RSP* characterizes the extent of specialization of a country in the production of a particular product (or products of a particular industry) compared to the industry structure of the reference group. It shows structural change in the country's manufacturing specialization and is calculated using the formula:

$$RSP_{ij} = \frac{P_{ij}}{\sum_{i=1}^{I} P_{ij}} \div \frac{\sum_{j=1}^{J} P_{ij}}{\sum_{j=1}^{J} \sum_{i=1}^{I} P_{ij}}, \tag{1}$$

where P_{ij} - production (gross output) of j-th industry of i-th country;

$\sum_{i=1}^{I} P_{ij}$ – total production of j-th industry of I reference group countries (EU, OECD, World) $i = \overline{1, I}$;

$\sum_{j=1}^{J} P_{ij}$ – total production of J industries of i-th country $j = \overline{1, J}$;

$\sum_{j=1}^{J} \sum_{i=1}^{I} P_{ij}$ – total production of J of industries of I reference group countries.

A relative location advantage of a production in a country exists if the value of *RSP* > 1, that is, the country's share in the gross output of a particular industry is greater than the share of the country's total output in the total indicator of the reference group (EU, OECD, World). The RSP indicator measures the extent to which a country has the advantages of an industry location (in line with the OLI advantages paradigm of J. Dunning) compared to other countries in the reference group.

To reflect the value added created as a result of the goods' production, the indicator "value added by activity" is used, measured as the cost of products less the cost of intermediate consumption. To assess the advantages of the country in creating added value in the conditions of relocation of production of goods and components, the authors suggest using the coefficient of *Comparative Advantage in Value Added by Activity* (*CAVA*), calculating it using the formula:

$$CAVA_{ij} = \frac{VA_{ij}}{\sum_{i=1}^{I} VA_{ij}} \div \frac{\sum_{j=1}^{J} VA_{ij}}{\sum_{j=1}^{J} \sum_{i=1}^{I} VA_{ij}}, \quad (2)$$

where VA_{ij} – added value of j-th industry of i-th country;

$\sum_{i=1}^{I} VA_{ij}$ – total value added of j-th industry of I countries of the reference group (EU, OECD, World) $i = \overline{1, I}$;

$\sum_{j=1}^{J} VA_{ij}$ – total value added of J industries of i-th country $j = \overline{1, J}$;

$\sum_{j=1}^{J} \sum_{i=1}^{I} VA_{ij}$ – total value added of J industries of I countries of the reference group.

There is a relative advantage if the value *CAVA* > 1, i.e. the share of value added of a particular industry in a country in the value added of that manufacturing in the reference group countries, exceeds the share of total industry value added that country in the total manufacturing value added of the countries of reference group.

To assess the value added created by the country's activities in the production of goods for export, the OECD accumulates statistics of foreign trade on value added (TiVA database). This database accumulates, among other things; data from the indicator "domestic value added in gross exports" (labelled as EXGR_DVA). This indicator is calculated as the difference between gross output in core prices and intermediate consumption in purchase prices. Internal value added can be decomposed into the following components: employee remuneration; gross operating income; mixed income; production taxes less production subsidies (OECD, 2021). The scientists named this indicator "Value-Added Exports" and used as a measure of the domestic value added embodied in exports (Johnson, 2014).

To assess the comparative advantages of a country in a particular industry export of goods, the authors suggest using coefficient of Comparative Advantage in Embodying Domestic Value Added in Exports (EVA), calculating it using the formula:

$$EVA_{ij} = \frac{ExDVA_{ij}}{\sum_{i=1}^{I} ExDVA_{ij}} \div \frac{\sum_{j=1}^{J} ExDVA_{ij}}{\sum_{j=1}^{J}\sum_{i=1}^{I} ExDVA_{ij}}, \tag{3}$$

where $ExDVA_{ij}$ – domestic value-added exports of j-th industry of i-th country;

$\sum_{i=1}^{I} ExDVA_{ij}$ – total domestic value-added exports of j-th industry of I countries of the reference group (EU, OECD, World) $i = \overline{1, I}$;

$\sum_{j=1}^{J} ExDVA_{ij}$ – total domestic value-added exports of J industries of i-th country $j = \overline{1, J}$;

$\sum_{j=1}^{J}\sum_{i=1}^{I} ExDVA_{ij}$ – total domestic value-added exports of J industries of I countries of the reference group.

Relative advantage of i-th country in embodying domestic value added in exports of j-th industry exists if the value of $EVA > 1$. That is, when the industry export of goods with added value of the country in the structure of industry exports of the reference group is higher than the total industrial export of goods with added value of the country in the structure of total industrial exports of the reference group. This indicates that the country is making more effective use of its comparative advantages in creating added value based on local resources, translating it into export-oriented goods.

3. Empirical analysis

At first glance, the EU27 pharmaceutical industry shows a strong position: in terms of value added in 2019 EU ranked first in the world, ahead of its main competitor, the United States, with \$174,832 million against \$168,517 million (OECD, 2021). The lockdown in early 2020 and supply problems pharmaceuticals (on both the export and import sides) came as a shock to the EU27 pharmaceutical industry. However, at the year-end the industry did not lose but increased export of drugs and medical devices by 5% (Table 1).

Table 1. Dynamics of extra-EU exports of basic pharmaceutical products and preparations from 2019 to 2020, EUR million

Country	2019	2020	2020/2019, %
Germany	48 412	50 086	3.5
Ireland	38 629	40 249	4.2
Belgium	28 910	33 180	14.8
Netherlands	22 437	22 159	-1.2
France	17 794	17 748	-0.3
Italy	15 974	15 578	-2.5
Denmark	6 863	7 345	7.0
Sweden	6 621	6 805	2.8
Spain	4 972	4 995	0.5
Austria	4 524	4 976	10.0
Slovenia	3 100	4 557	47.0
Hungary	1 809	2 041	12.8
Poland	1 151	1 260	9.5
Other EU countries	4 046	4 369	8.0
EU27	205 242	215 348	4.9

Source: Eurostat's data, https://ec.europa.eu/eurostat/data/database.

Taking into account the statistics of trade on added value (OECD, 2021), from Figures (1 and 2) it can be seen that since 1995 EU27 the upward trend in gross pharmaceutical exports. At the same time, the share of foreign value added in it increased from 7.0% (1995) to 17.8% (2018), while this indicator for the main competitors, the United States, China, India, is 10.8%, 11.4% and 15.0%, respectively.

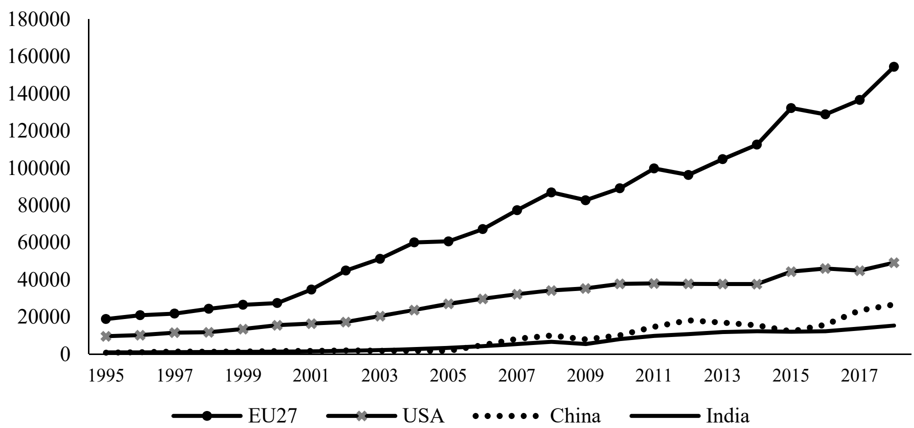


Figure 1. Dynamics of gross exports pharmaceuticals US Dollar, Millions.

Source: TiVA database OECD, www.oecd.org/

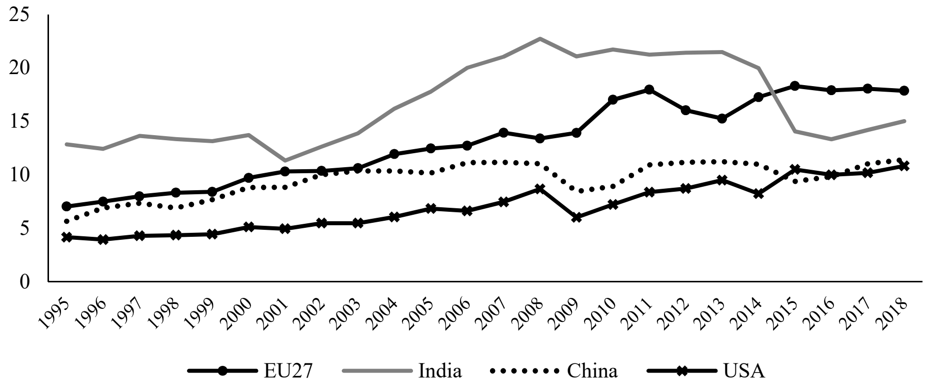


Figure 2. Dynamics of foreign value added share of gross exports, %.
Source: TiVA database OECD, www.oecd.org/

Increase in foreign value added share of gross exports confirms the hypothesis that RCA allows to estimate more visible than revealed comparative advantages in the trade of goods; that RCA cannot be an indicator that uniquely characterizes international competitiveness of the national economy in the face of the globalization with a outsourcing and offshoring of production; and that new approaches are needed to evaluate comparative advantages.

In the calculation of the author's indicators, formulas 1-3 use statistics of indicators from the TiVA database (OECD, 2021) for industry D21 Pharmaceuticals, medicinal chemical and botanical products:

for P_{ij} : PROD indicator: Production (gross output),

for VA_{ij} : VALU indicator: Value added,

for $ExDVA_{ij}$: EXGR_DVA indicator: Domestic value added content of gross exports.

The EU27 countries were selected as the reference group. The calculation results indicate the following.

RSP of pharmaceutical industry in the countries of EU27 shows (Figure 3, the author's calculations of Formula (1)) that in 1995 12 countries of this reference group had an indicator value more than 1, that is, they were specialized in the production of pharmaceutical products. But by 2018, most of them had lost that advantage. Only the industries of Ireland, Denmark, Belgium, Slovenia and Malta remained specialized, and changes in Cyprus' economic policy in terms of increasing the advantages of location to attract foreign investment in the development of pharmaceutical enterprises in the country contributed to obtaining advantages in this industry.

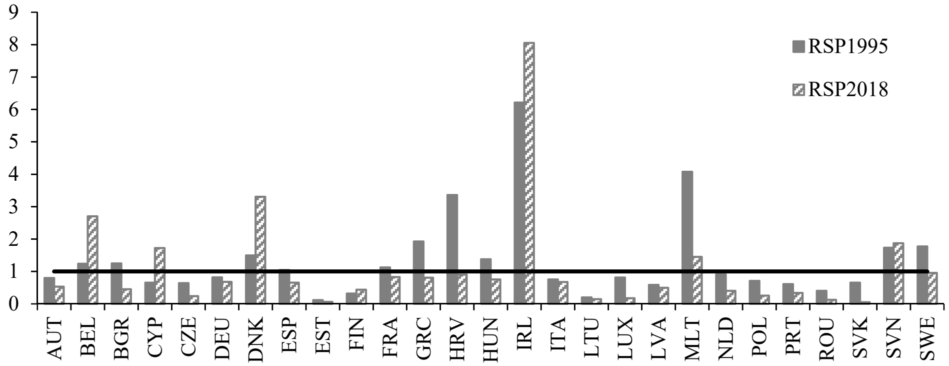


Figure 3. RSP of pharmaceutical industry in the countries of the EU27 reference group for 1995 and 2018.

Source: The author's calculations according to formula (1) based on the OECD TiVA Database.

EU27 countries pharmaceuticals' CAVA shows (Figure 4, the author's calculations of Formula (2)) that in 1995, 10 out of 12 countries that had a production specialization in this industry according to the RSP, showed advantages in creating added value. Two countries (Bulgaria and Spain) did not show any advantages, that is, the local industry relied heavily on foreign pharmaceuticals substances and value-added created abroad. By 2018, Ireland, Denmark, Belgium and Slovenia still had the advantage. Cyprus also joined them.

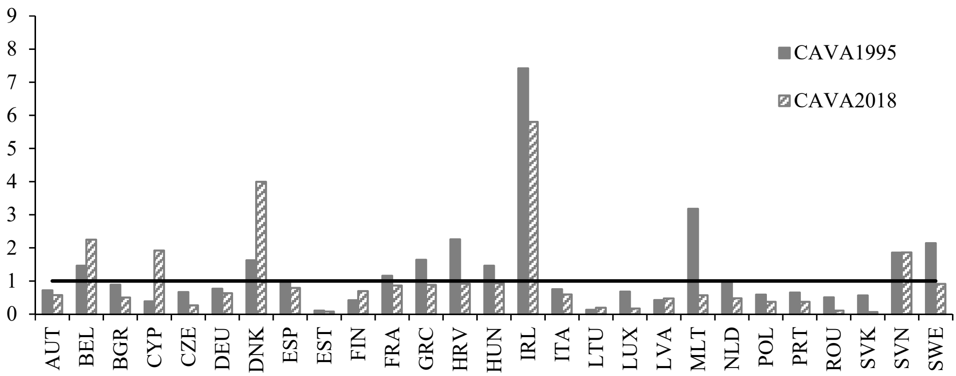


Figure 4. CAVA of pharmaceutical industry in the countries of the EU27 reference group for 1995 and 2018.

Source: The author's calculations according to formula (2) based on the OECD TiVA Database.

But the loss (or absence) of a country's advantage on CAVA of pharmaceuticals in the EU27 reference group does not necessarily mean the loss (absence) of such an advantage in the global economy (Figure 5, the author's calculations of Formula (2)).

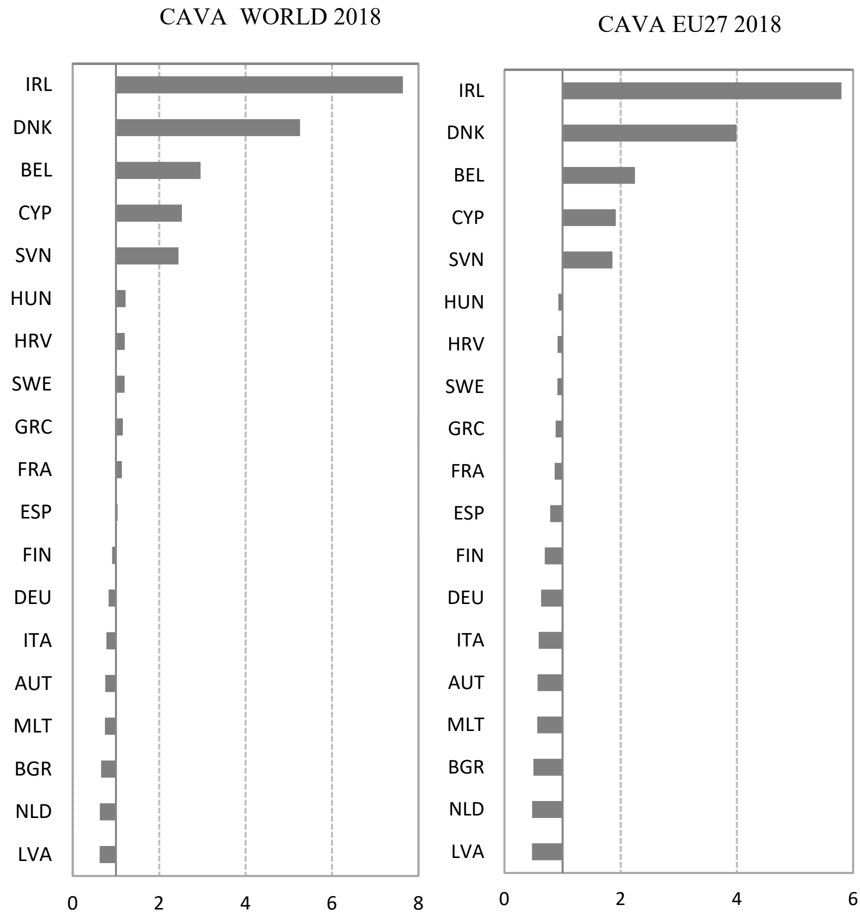


Figure 5. CAVA of pharmaceutical industry in the countries of the reference group: WORLD and EU27.

Source: The author's calculations according to formula (2) based on the OECD TiVA Database.

As an example, France has different values of the CAVA of pharmaceuticals for different reference groups (Figure 6, the author's calculations of Formula (2)). Against the background of countries around the world (the OECD provides statistics in TiVA Database for 66 countries and "other" countries around the world (OECD, 2021), France continues to have advantages in creating added value by pharmaceuticals (from 2013 there was a downtrend), while inside the EU27 the country no longer has advantages.

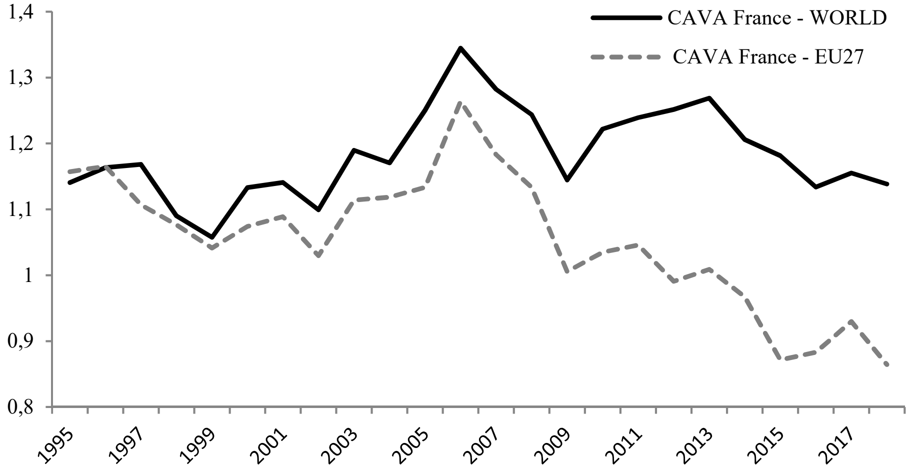


Figure 6. Trend *CAVA* pharmaceuticals industry France by reference groups of WORLD and EU27.
Source: The author's calculations according to formula (2) based on the OECD TiVA Database.

The *EVA* revealed the comparative advantages of 10 EU27 countries in the export of pharmaceutical products in 1995. But by 2018, Bulgaria, Italy, France, Croatia and Sweden reduced the volume of domestic value added embodied in exports, having increased the content of foreign components, which resulted in the loss of their positions (Figure 7, the author's calculations of Formula (3)). Only 6 countries – Ireland, Denmark, Belgium, Slovenia, Malta held and Cyprus increased their comparative advantages in domestic value added export compared to other EU27 countries.

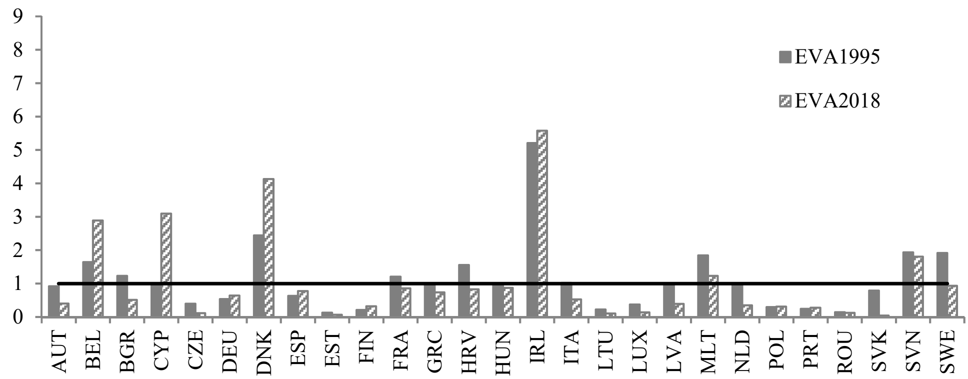


Figure 7. EVA of pharmaceutical industry in the countries of the EU27 reference group for 1995 and 2018.
Source: The author's calculations according to formula (3) based on the OECD TiVA Database.

For a better graphical representation of the calculated coefficients of RSP, CAVA, and *EVA* of pharmaceutical industry in the countries of the EC27 reference group in 1995 and 2018, the natural logarithm of their values was used. Belgium, Denmark, France, Croatia, Ireland, Malta, Slovenia and Sweden had advantages in the production, value added and export of pharmaceuticals in 1995 (Figure 8, the author's calculations of the Formula (1), (2), (3)). A number of countries, the like of Bulgaria, Hungary, and Greece, demonstrated advantages in the pharmaceutical industry only by two coefficients.

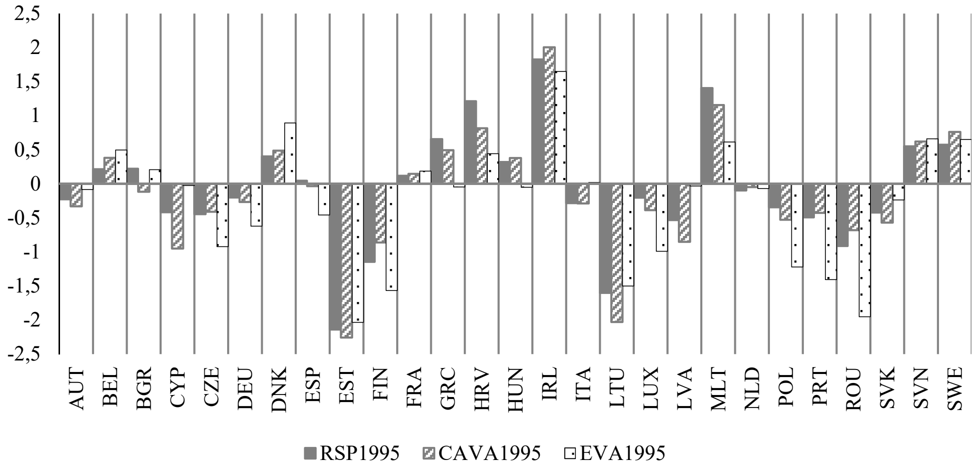


Figure 8. Diagram of RSP, CAVA, EVA of pharmaceutical industry in the countries of the EU27 reference group (1995), filed on a logopharific scale.

Source: The author's calculations according to formula (1-3) based on the OECD TiVA Database.

According to the results obtained (Figure 9, the author's calculations of the Formula (1), (2), (3)), today EU27 has formed a narrow circle (Belgium, Cyprus, Denmark, Ireland, Slovenia) of undisputed leaders in terms of specialization and comparative advantages in pharmaceuticals' production, value added and exports against the background of other industries. Malta has no advantages in creating added value in pharmaceuticals.

Summarizing the above, we can say that on the eve of the crisis phenomena in the economy caused by the pandemic and changes in the European policy with the adoption of "A new industrial strategy for Europe", six EU27 countries showed the specialization of pharmaceuticals industry and the ability of the sector to obtain higher added value (based on local resources) compared to the values of the indicators of other manufacturing of the reference group. Although many issues remain unresolved in the statistical tools for analysing countries' positions in the global economic system, applying the author's approach to assess the consequences of the new mechanisms of EU industrial policy will reveal changes in the specialization and increase of competitive advantages of industries by increasing domestic value added.

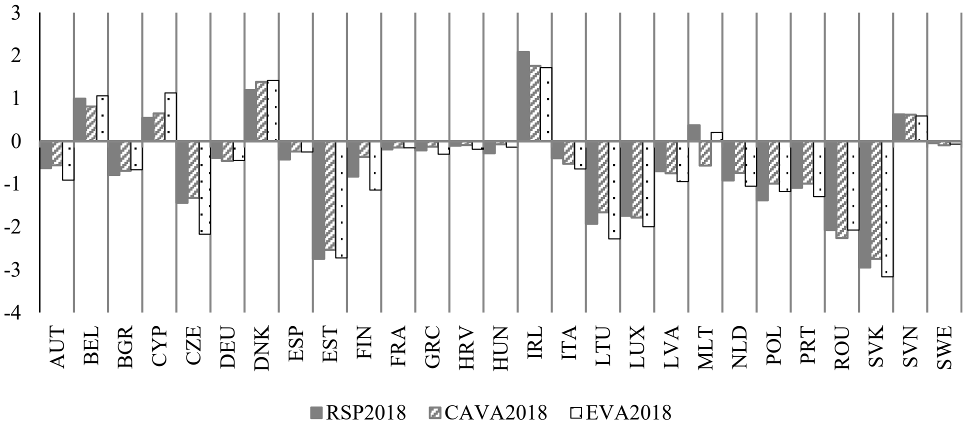


Figure 9. Diagram of RSP, CAVA, EVA of pharmaceutical industry in the countries of the EU27 reference group (2018), filed on a logopharific scale.

Source: The author's calculations according to formula (1-3) based on the OECD TiVA Database.

4. Conclusions

According to the study, the new industrial policy of the EU and member states regarding pharmaceuticals manufacturing and the adoption of a number of relevant strategic documents provide for implementation of mechanisms to eliminate the consequences of the delocalization strategy and reduce dependence on offshore production of both API and finished medicines. The introduced tools for providing state financial assistance for the implementation of investment and innovation projects for the development of production capacities related to COVID-19, encourage the expansion of the "product portfolio" of existing companies, the creation of new enterprises "from scratch" and/or reshoring (in some cases – backshoring) of pharmaceutical industries from Asia to the territory of EU member states. The toolkit created by the authors (the introduction of indicators into scientific turnover: RSP – coefficient of revealed productive specialization, CAVA – coefficient of comparative advantage of the country in the creation of value added and EVA – coefficient of comparative advantage in the domestic value added exports) would permit analysis of the position of industries in the country to assess changes in specialization and location advantages of industries in Europe as a result of the outcome of the priority goals for the EU.

Calculation of RSP, CAVA, EVA indicators for pharmaceuticals industry based on the results of the first 3 years of implementation of the new EU goals on production relocation and comparative analysis of data from 2013 to 2018 are considered by the authors as areas of further research. At the same time, the proposed author's approach should be used in assessing not only pharma, but also other high-tech manufactories that are the basis of industrial ecosystems identified as priorities in "New industrial strategy for Europe" and also to complement it by the approaches to study trends in the globalization of high-tech production through monitoring and analysis of foreign trade in intermediate high-tech goods.

References

Balassa, B., (1961). The Theory of Economic Integration. London, George Allen & Unwin Ltd.

Balassa, B., (1989). 'Revealed' comparative advantage revisited. In B. Balassa (Ed.). Comparative advantage, trade policy and economic development. New York: New York University Press.

Barlow, P., van Schalkwyk, M. C., McKee, M., Labonté, R., and Stuckler, D., (2021). COVID-19 and the collapse of global trade: Building an effective public health response. *The Lancet Planetary Health*, 5, e102–e107. DOI: https://doi.org/10.1016/S2542-5196(20)30291-6.

Dunning, J. H., (1992). The Competitive Advantage of Countries and the Activities of Transnational Corporations, *Transnational Corporations*, No. 2, pp. 11–16.

Dunning, J. H., (1993) Multinational Enterprises and the Global Economy. Addison Wesley, New York.

European Commission, (2020a). Dependency of the EU pharmaceutical industry on active pharmaceutical ingredients and chemical raw materials imported from third countries. Retrieved from: https://ec.europa.eu/health/sites/health/files/files/committee/ev_20200312_795_en.pdf

European Commission, (2020b). Affordable, accessible and safe medicines for all: the Commission presents a Pharmaceutical Strategy for Europe. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2173.

European Commission, (2020c). State aid SA.57100 (2020/N) – Germany – COVID-19 – Federal Framework Scheme "Aid for Covid-19 related R&D, investments in testing infrastructures and production facilities" ("Bundesregelung Forschungs- Entwicklungs- und Investitionsbeihilfen"), Retrieved from: https://ec.europa.eu/competition/state_aid/cases1/202018/285722_2152196_87_2.pdf

European Commission, (2020d). State Aid SA.57367 (2020/N) – France COVID-19: Aid for COVID-19 relevant R&D projects, investment into relevant testing and upscaling infrastructures, and investment into COVID-19 relevant production capacities, Retrieved from: https://ec.europa.eu/competition/state_aid/cases1/202024/286299_2162472_123_4.pdf

European Commission, (2020e). State aid SA.57519 (2020/N) – Poland COVID-19: R&D aid for Covid-19 relevant research and development, investment aid for the construction and upgrade of relevant testing and upscaling infrastructures, and investment aid for investments into production facilities for the production of Covid-19 relevant products, Retrieved from: https://ec.europa.eu/competition/state_aid/cases1/202027/286598_2168965_58_2.pdf

European Commission, (2021). Updating the 2020 Industrial Strategy: towards a stronger Single Market for Europe's recovery. Retrieved from: https://ec.europa.eu/commission/presscorner/detail/en/IP_21_1884

European Parliament, (2020). European Parliament resolution of 17 September 2020 on the shortage of medicines – how to address an emerging problem (2020/2071(INI)). Retrieved from: https://www.europarl.europa.eu/doceo/document/TA-9-2020-0228_EN.html

Eurostat, (2021a). COVID-19: support for statisticians. Retrieved from: https://ec.europa.eu/eurostat/data/metadata/covid-19-support-for-statisticians

Eurostat, (2021b). European statistical recovery dashboard. Retrieved from: https://ec.europa.eu/eurostat/cache/recovery-dashboard/

Falkowski, K., (2018a). Competitiveness of the Baltic States in international high-technology goods trade, Comparative Economic Research. Central and Eastern Europe, ISSN 2082–6737, *De Gruyter*, Warsaw, Vol. 21, Iss. 1, pp. 25–43, https://doi.org/10.2478/cer-2018-0002

Falkowski, K., (2018b). Trade interdependence between Russia vs. the European Union and China within the context of the competitiveness of the Russian economy, Equilibrium. *Quarterly Journal of Economics and Economic Policy*, Institute of Economic Research, vol. 13(4), pp. 667–687. https://doi.org/10.24136/eq.2018.032

Grumiller, J., Grohs H., Reiner, C., (2021). Increasing the resilience and security of supply of production post-COVID-19 – The Case of Medical and Pharmaceutical Products. Working Paper Reihe der AK Wien – Materialien zu Wirtschaft und Gesellschaft 216, *Kammer für Arbeiter und Angestellte für Wien*, Abteilung Wirtschaftswissenschaft und Statistik.

Johnson Robert, C., (2014) Five Facts about Value-Added Exports and Implications for Macroeconomics and Trade Research. *Journal of Economic Perspectives*, 28, 2, pp. 119–142. Retrieved from: http://gesd.free.fr/johnson514.pdf

Kumar A, Luthra S, Mangla S. K., Kazançoğlu Y., (2020). COVID-19 impact on sustainable production and operations management. *Sustain Oper Comput*, 1, pp. 1–7. https://doi.org/10.1016/j.susoc.2020.06.001.

Motoryn, R., Prykhodko, K., Ślusarczyk, B., (2020). Asymmetry of foreign trade turnover between Ukraine and Poland. *Statistics In Transition new series*, June 2020, Vol. 21, No. 2, pp. 143–156, DOI: 10.21307/stattrans-2020-017.

OECD, (2021). Trade in Value Added. Retrieved from: https://www.oecd.org/sti/ind/measuring-trade-in-value-added.htm#access

OECD, (2021). Trade in Value Added. Retrieved from: https://www.oecd.org/sti/ind/measuring-trade-in-value-added.htm#access

Salikhova, O. B. (2012). New approaches to determining international comparative advantages. *Economy and Forecasting*, 1, pp. 101–120. Retrieved from: http://eip.org.ua/docs/EP_12_1_101.pdf

Tirivangani T, Alpo B, Kibuule D, Gaeseb J, Adenuga B. A., (2021). Impact of COVID-19 pandemic on pharmaceutical systems and supply chain–a phenomenological study. *Explorat Res Clin Social Pharm.*, 2, 100037. https://doi.org/10.1016/j.rcsop.2021.100037

Vo, T. D.; Tran, M. D. (2021). The impact of covid-19 pandemic on the global trade. *Int. J. Soc. Sci. Econ. Invent.*, 7, pp. 1–7. https://doi.org/10.23958/ijssei/vol07-i01/261

Wang, Y., Jiang, S., (1999). China: A future star for foreign pharma companies. *Pharmaceutical Executive*, 19(8), pp. 78–87.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 229–258, DOI 10.59170/stattrans-2023-013
Received – 01.07.2022; accepted – 03.11.2022

Assessing the maturity of the current global system for combating financial and cyber fraud

Olha Kuzmenko[1], Hanna Yarovenko[2], Larysa Perkhun[3]

ABSTRACT

The purpose of the article is to assess the maturity of systems for counteracting financial and cyber fraud with the view of their future integration at global-level. The calculations made by the authors were based on indicators for 76 countries, which characterized each country's level of cybersecurity and its ability to combat financial fraud in 2018. After optimising the input data and selecting relevant indicators, the authors built an integrated cybersecurity index using the Sundarovsky convolution method. Sigma-limited parameterisation and Pareto-optimisation were then used to identify the determinants of the ability to counter financial and cyber fraud, which were used as predictors. Nonlinear regression was applied to determine the dependency of the integrated cybersecurity index on the government efficiency index, the ease of doing business and on the crime indices. On this basis, the authors conducted a bifurcation analysis of the maturity of current global system for combating financial and cyber fraud and produced its phase portraits. It was found to be mature ("Government Efficiency Index – Ease of Doing Business" and "Ease of Doing Business – Crime Index") and insufficient mature ("Government Efficiency Index – Crime Index"), with the components' imbalance indicating high system's sensitivity to react on changes. The constructed 'Equilibrium States' phase portraits showed non-equilibrium phase portraits of the 'saddle' type. The obtained results made it possible to identify determinants of a global integrated system's instability to combat financial and cyber fraud.

Key words: financial fraud, cyber fraud, phase portrait, bifurcation analysis.

[1] Sumy State University, Ukraine. E-mail: o.kuzmenko@biem.sumdu.edu. ORCID: https://orcid.org/0000-0001-8520-2266

[2] Sumy State University, Ukraine; University Carlos III of Madrid, Spain. E-mail: hyaroven@inf.uc3m.es. ORCID: https://orcid.org/0000-0002-8760-6835.

[3] MHP SE, Ukraine. E-mail: l.perkhun@mhp.com.ua. ORCID: https://orcid.org/0000-0002-8667-2312.

© Olha Kuzmenko, Hanna Yarovenko, Larysa Perkhun. Article available under the CC BY-SA 4.0 licence

1. Introduction

One of the key tasks facing society today is to create a mechanism of protection against internal threats, which is even more urgently needed in the conditions of war with an external enemy. This is important for two main reasons. First, military conflicts in a country increase its attractiveness as a place for money laundering and the financing of terrorism (Yarovenko, 2021). Secondly, the risk of cyber fraud attacks against various elements of state and private infrastructure is increasing. For example, the number of cyber-attacks started to increase in the world in 2022 even before Russia's military aggression against Ukraine. For example, a large-scale attack against more than 70 government websites was recorded on February 14, 2022 (BBC, 2022); Ukraine's banking institutions were attacked on February 15, 2022 (Euronews, 2022). According to analytical data provided by the Quad9 DNS platform, there was a significant increase in the number of cyberattacks against Ukrainians in March. Of the 121 million malicious events recorded globally as of March 9, 2022, 4.6 million were associated with Ukraine and Poland, where 1.4 million Ukrainian citizens had been displaced by early March 2022 (Krebsonsecurity, 2022).

In addition to information attacks, there were occurrences of cyber financial fraud as well. On February 14, 2022, an 'IcedID' banking Trojan collecting personal banking data of Ukrainians was detected. In April 2022, another case of Internet fraud was discovered, where a fictitious social media page was used to collect financial assistance from EU countries through payments in violation of the confidentiality of payment card data (CyberPeace Institute, 2022).

These examples show that the issue of counteracting financial and cyber fraud is relevant and should be solved at various levels of public administration. In order to achieve it, a systematic approach should be applied, which necessitates the convergence of systems to combat financial and cyber fraud. This, in turn, is possible when their information, technical, software, and organizational integration takes place, both at the level of the state as a whole and at the level of individual business entities and the world. The need for convergence processes in the areas of money laundering and cyber fraud was identified by the US Financial Crimes Enforcement Network (FinCEN, 2009). It should occur at the level of relevant departments responsible for combating money laundering, terrorist financing, cybersecurity, and by businesses themselves. Global consulting companies Deloitte (2019) and PwC (2018) also covered this issue in their reports.

This paper deals with the issue of assessing the maturity of current global system for countering financial and cyber fraud (CFCF) in order to determine its/their potential convergence opportunities in the future. To this end, the authors apply econometric and statistical methods that allow the assessment based on large amounts of data, taking into account temporal, spatial or other characteristics and factors.

2. Literature review

The maturity of CFCS system's depends on the advancement of specific processes and organization practices that are in place in order to ensure the system achieves desired results. In order to reach an appropriate level of maturity, it is also necessary to apply methods that can strengthen the level of protection against financial and cyber threats. Traditional methods are becoming less effective in accomplishing this goal. Matanky-Becker and Cockbain (2022) found that the widely used international three-stage money laundering model is less practical and reliable. It was used in less than a third of cases for a three-year data sample. Therefore, economic, and mathematical methods are increasingly complementing the traditional ones and or replacing them altogether. Two most powerful tools in the fight against financial and cyber fraud are artificial intelligence and machine learning, which rely on models of varying complexity that are constantly trained and retrained, and adapted to new conditions in which the research object functions. Machine learning was used by Hayble-Gomes (2022) to determine attributes needed to generate a suspicious activity report based on the transaction history of U.S. retail banking customers. Neural networks are used to identify and recognize faces of clients of financial institutions (Granados and Garcia-Bedoya, 2022), based on various data sources, such as social networks. AI tools are very effective in detecting financial crimes, including those related to money laundering, because they can be used to develop models which can identify such cases without human intervention (Rouhollahi, et al., 2021). Among different machine learning methods, the Light Gradient Boosting and Extreme Gradient Boosting have high accuracy, reaching more than 99% (Aziz, et al., 2022). Another algorithm, called Random Forest, has also been shown to very effective (94%) in modelling suspicious money laundering transactions (Tundis, Nemalikanti and Mühlhäuser, 2021).

In cases of mass financial and cyber fraud, it is possible to use models to identify group behaviour of individuals in order to identify similar characteristics and detect other cybercrimes associated with similar patterns (Mahootiha, Golpayegani and Sadeghian, 2021). An algorithm for approximating multifunctional behaviour can be used to track actions of users when they access financial transactions at their intermediate nodes (Amala Dhaya and Ravi, 2021). Genetic programming and token competition, proposed by Li and Wong (2021), have shown their effectiveness in determining objective values of individuals, which can be used to distinguishing those who differ from others. Robust regression (Riani, Corbellini and Atkinson, 2018) and logit regression (Yang and Wu, 2021) can be combined with neural networks to detect symptoms of financial and cyber fraud. In combination with dynamic evolutionary glow-worm swarm optimization, Xia et al. (2022) proposed a biological algorithm for predicting the risk of financial fraud. Perkhun, Sorochynskyi and Izosimov (2015) researched the interaction of fraudulent attacks and tools to combat them on the basis of the modified Lotka–Volterra model. Granados and Vargas (2022) looked at how the Foreman-Ricci curvature can be used to build financial networks and quantify sets of suspicious nodes to create a strategy for detecting global financial crime and fraud. The use of data visualization is also a powerful tool for detecting abnormal operations and be used to identify fraud quickly and clearly. Tharani, et al. (2021) proposed a visualization of functions related to transactions in the Bitcoin and Ethereum networks, which helps to quickly identify cases of cyber fraud. The knowledge graph presented by Day (2021) can be used to combat the misuse of electronic payment instruments and cryptocurrencies.

The above-mentioned methods, in most cases, are used directly to detect financial and cyber fraud in individual transactions. These tools are rarely used to study processes, for example to assess the maturity of systems. This article proposes the use of bifurcation analysis and the construction of phase portraits to determine the current state of a global system and the points at which it will reach its equilibrium. It is a viral method used to study dynamic systems. This toolkit was used by Akhramovych, et al. (2022) to review the information security system in social networks and build its linear and dynamic models. Idowu, et al. (2018) described phase portraits to justify the chaos of the financial system. Sierikov and Zubova (2010) built systems of nonlinear differential equations for the market, which were based on the supply and demand model. Bystray, Lykov and Nikulina (2012) developed their own method for identifying macroeconomic risks, which involves the construction of pseudo-phase and phase portraits. Wilkens, Thomas and Fofana (2004) used phase portraits to determine price stability for technology stocks. As can be seen, bifurcation analysis and phase portraits have a wide range of applications for studying states of various systems. In this article, they are used to assess the maturity of the global system for combating financial and cyber fraud.

3. Research Methodology and Data

3.1. Research Methodology

The maturity of the current CFCF system is assessed in several stages.

Stage 1. Indicators are reduced to a single integral cybersecurity index using the Sundarovsky method, which involves the use of formula (1):

$$IS_j = \prod_{i=1}^{n} \left[a_{ij} - a_i^*\right]^{\alpha} \tag{1}$$

where IS_j - integral cybersecurity index for the j-th country;

a_{ij} – actual value of the i-th cybersecurity indicator for the j-th country;

a_i^*- equilibrium value of the i-th cybersecurity indicator for the considered set of countries;

α – constant, exponent.

In order to apply formula (1) to calculate the integral cybersecurity index, we introduce the following assumptions:

1) the absolute value of the difference between the standard deviation and the minimum allowable level is used as the equilibrium level of the constituent indicators:

$$a_i^* = \left|a_{ij} - \sigma_i\right| = \left|a_{ij} - \sqrt{\frac{\sum_{j=1}^{m} \left(a_{ij} - \underline{a_i}\right)}{n-1}}\right| \tag{2}$$

where σ_i - standard deviation of the i-th cybersecurity indicator;

$\underline{a_i}$ - arithmetic mean of the i-th cybersecurity indicator;

2) the ratio of a single value and the number of relevant indicators of cybersecurity is used as a constant value of the indicator of the degree of functional dependence (1). Considering these assumptions, formula (1) takes the form:

$$IS_j = \prod_{i=1}^{n} \left[a_{ij} - \left|a_{ij} - \sqrt{\frac{\sum_{j=1}^{m} \left(a_{ij} - \underline{a_i}\right)}{n-1}}\right|\right]^{1/n} \tag{3}$$

where n - the number of relevant indicators characterizing cybersecurity.

Stage 2. Relevant indicators characterizing the ability of countries to counteract financial fraud are determined by applying sigma-limited parameterization and Pareto optimization. We choose the integral cybersecurity index determined by the Sundarovsky method as an effective factor and indicators characterizing the ability of countries to counteract financial and cyber threats as factors of influence. Sigma-limited parametrization is performed in the form of a one-dimensional test of significance of the influence of indicators on the effective factor, and Pareto optimization is performed by constructing a Pareto diagram of t-values.

Stage 3. A non-linear regression model is built, which describes the dependence of the integral cybersecurity index on the relevant predictors identified at stage 2, and insignificant parameters are eliminated step-by-step. A combination of logarithmic and quadratic functions as well as the multiplicative dependence of the selected indicators should be considered in this process with a view to the following bifurcation analysis of the maturity of the current system for combating financial and cyber fraud and the construction of its phase portraits.

Stage 4. A bifurcation analysis of the maturity of the current CFCF system is conducted and phase portraits of its "maturity" are constructed. This requires intermediate calculations involving the apparatus of differential calculus in order to determine partial derivatives of the function of the dependence of the integral cybersecurity index on relevant predictors and to create a system of differential equations that will serve as the basis for further analysis of the dynamic stability of the system.

Stage 5. A non-linear regression model of the dependence of the integral cybersecurity index on relevant predictors is built using a combination of power, trigonometric and multiplicative dependence of indicators with a view to conducting a bifurcation analysis of the equilibrium states of the current CFCF system and constructing its phase portraits.

Stage 6. A bifurcation analysis of the maturity of the current CFCF system phase portraits of its "equilibrium states" are constructed. This stage is performed similarly to stage 4. .

3.2. Data

To assess the maturity of thecurrent CFCF system, input data from 76 countries were collected and systematized according to two sets of indicators for 2018. One set describes the level of cybersecurity in each country at the national and global levels, and the level of its digitalization and informatization. They relate to the global cybersecurity system of world countries. The second set of indicators characterizes each country's attractiveness for money laundering and is used to make conclusions about its ability to counteract financial threats associated with money laundering and terrorist financing at the macro level. They relate to the global system of combating financial fraud in world countries. Five indicators are included in the first group (e-Governance Academy Foundation, 2022): the Global Cybersecurity Index (GCI), the ICT Development Index (ICT DI), the Network Readiness Index (NRI), the National Cyber Security Index (NCSI), and the Digital Development Level (DDL). The second set includes the Political Stability Index (PSI), the Government Efficiency Index (GEI), the Corruption Perception Index (CPI) (The GlobalEconomy, 2022), the Ease of Doing Business Index (EDB), the Crime Index (CI) (The World Bank, 2022), the Global Terrorism Index (GTI) (OCHA, 2018), and the Financial Secrecy Index (FSI) (Netzwerk Steuergerechtigkeit, 2022).

4. Empirical Results

4.1. Results of Data Analysis

It is necessary to identify causal relationships among the indicators of cybersecurity and indicators that characterize the ability of countries to counteract financial crimes. With this end in mind, a canonical analysis using the Statistica analytical package was conducted and its results are presented in Table 1.

Table 1. Results of the canonical analysis of cause-and-effect relationships between the indicators of cybersecurity and those characterizing countries' ability to counteract financial fraud

Variable	Left Set	Right Set
Variance extracted	100.00%	86.67%
Total redundancy	65.51%	49.39%
Variable 1	Global Cybersecurity Index	Political Stability Index
Variable 2	ICT Development Index	Government Effectiveness Index
Variable 3	Network Readiness Index	Ease of Doing Business
Variable 4	National Cybersecurity Index	Crime Index
Variable 5	Digital Development Level	Corruption Perceptions Index
Variable 6	-	Global Terrorism Index
Variable 7	-	Financial Secrecy Index
Canonical R	0.91	
Chi-sqr(35)	196.50	
p	0.0000	

Source: authors' calculations based on Kuzmenko, Yarovenko and Radko (2021).

As can be seen, 65.51% of the variance in the cybersecurity indicators is explained by the indicators describing countries' ability to combat financial crime. At the same time, only 49.39% of the variance in the indicators characterizing countries' ability to counteract financial threats is explained by the cybersecurity indicators. In addition, the share of variance (variability) explained by the indicators of cybersecurity is 100%, while the share of variance explained by the indicators describing countries' ability to counter financial threats is 86.67%. This means that the latter ones can be treated as the cause, while the former ones as the effect. The canonical correlation R=0.91, which corresponds to the correlation between the first canonical variables, is equal to the maximum canonical root. Its value indicates a strong relationship between groups of variables. The significance of the canonical correlation coefficient is confirmed by the values Chi-Square=196.5 and the level p=0.00.

We optimize the input data array, for which we conduct Chi-Square tests for the statistical significance of canonical roots (Table 2). The first three canonical roots can be considered statistically significant since their values do not exceed the maximum allowable level of 0.05. Roots three and four have a canonical R-sqr value approaching zero and a p-value greater than 0.05, which means they are not statistically significant. Therefore, the first three canonical roots are considered at the next stage when the input data array is optimized.

Table 2. Chi-Square tests of canonical roots

Root Removed	Canonical R	Canonical R-sqr	Chi-sqr.	df	p	Lambda Prime
0	0.9126	0.8328	196.4981	35	0.0000	0.0568
1	0.6730	0.4530	73.9741	24	0.0000	0.3396
2	0.5662	0.3206	32.6488	15	0.0053	0.6209
3	0.2727	0.0744	6.1705	8	0.6281	0.9139
4	0.1128	0.0127	0.8765	3	0.8311	0.9873

Source: authors' calculations based on Kuzmenko, Yarovenko and Radko (2021).

To optimize the array of input data, we conduct a correlation analysis of both sets of indicators (of cybersecurity and of countries' ability to counter financial and cyber fraud). The correlation matrix of cybersecurity indicators is presented in Table 3. The resulting values indicate a significant correlation between the ICT Development Index and the Digital Development Level (the value of the correlation coefficient is 0.96). To optimize the set of input indicators in terms of cybersecurity characteristics, one of the most collinear indicators is removed from further calculations.

Table 3. Correlation matrix of a set of cybersecurity indicators

Variables	GCI	ICT DI	NRI	NCSI	DDL
GCI	1.0000	0.5358	0.7114	0.7094	0.5792
ICT DI	0.5358	1.0000	0.5834	0.6430	0.9607
NRI	0.7114	0.5834	1.0000	0.6813	0.6467
NCSI	0.7094	0.6430	0.6813	1.0000	0.6547
DDL	0.5792	0.9607	0.6467	0.6547	1.0000

Source: authors' calculations based on Kuzmenko, Yarovenko and Radko (2021).

To decide which indicator should be left in the input data array and which should be deleted, we consider the factor structure for the first three statistically significant canonical roots, selected using a piecewise linear plot and Chi-Square tests (Table 4).

Table 4. Factor structure of a set of cybersecurity indicators

Variables	Root 1	Root 2	Root 3	Root 4	Root 5
GCI	0.7935	-0.5738	0.0325	-0.1962	-0.0389
ICT DI	0.8712	0.1721	-0.3910	0.2355	-0.0550
NRI	0.8026	-0.2408	0.3794	0.3538	0.1697
NCSI	0.7257	-0.2962	-0.2189	0.0341	0.5801
DDL	0.9428	0.2574	-0.1977	0.0756	0.0015

Source: authors' calculations based on Kuzmenko, Yarovenko and Radko (2021).

Based on the analysis of the data in Table 4 it can be concluded that the indicator of the Digital Development Level has a more significant impact and should therefore be retained for further calculations.

Let us now consider the correlation matrix of the indicators describing countries' ability to counter financial and cyber fraud (Table 5).

Table 5. Correlation matrix of indicators describing countries' ability to counteract financial and fraud

Variables	PSI	GEI	EDB	CI	CPI	GTI	FSI
PSI	1.0000	0.6575	0.4557	-0.4952	0.7503	-0.6489	0.1353
GEI	0.6575	1.0000	0.8029	-0.6215	0.9037	-0.0476	0.4352
EDB	0.4557	0.8029	1.0000	-0.5826	0.6465	0.0023	0.2687
CI	-0.4952	-0.6215	-0.5826	1.0000	-0.5570	0.1732	-0.2272
CPI	0.7503	0.9037	0.6465	-0.5570	1.0000	-0.1809	0.3449
GTI	-0.6489	-0.0476	0.0023	0.1732	-0.1809	1.0000	0.2143
FSI	0.1353	0.4352	0.2687	-0.2272	0.3449	0.2143	1.0000

Source: authors' calculations based on Kuzmenko, Yarovenko and Radko (2021).

As can be seen, there is a significant correlation between the Government Efficiency Index and the Corruption Perception Index, as evidenced by the value of the correlation coefficient of 0.904. To optimize the set of input indicators in terms of countries' ability to counteract financial and cyber fraud, one of the collinear indicators is removed from further calculations.

To decide which of the two indicators (Government Efficiency Index or Corruption Perception Index) should be retained in the input data array and which should be deleted, we consider the factor structure for the first three statistically significant canonical roots (Table 6). It can be concluded that the Government Efficiency Index has a more significant effect, which is why it is retained for further calculations.

Table 6. Factor structure of indicators of countries' ability to counteract financial and cyber fraud

Variables	Root 1	Root 2	Root 3	Root 4	Root 5
PSI	0.4314	0.6131	-0.5319	0.1180	0.0254
GEI	0.9545	0.1915	-0.1801	0.0406	-0.1027
EDB	0.8542	-0.2170	-0.2463	0.1014	0.2660
CI	-0.5569	0.0130	0.6724	-0.1198	-0.1878
CPI	0.8162	0.5048	-0.2211	-0.1246	-0.0019
GTI	0.1495	-0.6210	0.3207	-0.6026	-0.2745
FSI	0.5055	0.0899	0.3227	-0.3200	0.2829

Source: authors' calculations based on Kuzmenko, Yarovenko and Radko (2021).

4.2. Calculations Results

In the first stage, calculations were performed using formula (3). Their results are presented in Appendix, where the values in the IS column correspond to effective values of the integral cybersecurity index determined by the Sundarovsky method.

In the second stage, sigma-limited parametrization and Pareto optimization were carried out using the Statistica analytical package. The results are shown in Figures 1 and 2.

	Univariate Tests of Significance forIS (Spreadsheet1.sta) Sigma-restricted parameterization Effective hypothesis decomposition				
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	15,932	1	15,932	0,19911	0,656838
Political stability index	60,262	1	60,262	0,75310	0,388504
Government effectiveness index	651,476	1	651,476	8,14157	0,005706
Ease of doing business	1068,591	1	1068,591	13,35430	0,000499
Crime Index	197,796	1	197,796	2,47187	0,120474
Global Terrorism Index	185,399	1	185,399	2,31695	0,132540
Financial Secrece Index	63,668	1	63,668	0,79566	0,375494
Error	5521,278	69	80,019		

Figure 1. One-dimensional test of the significance of the influence of indicators of countries' ability to counteract financial fraud on the integral index of cybersecurity

Source: authors' calculations.

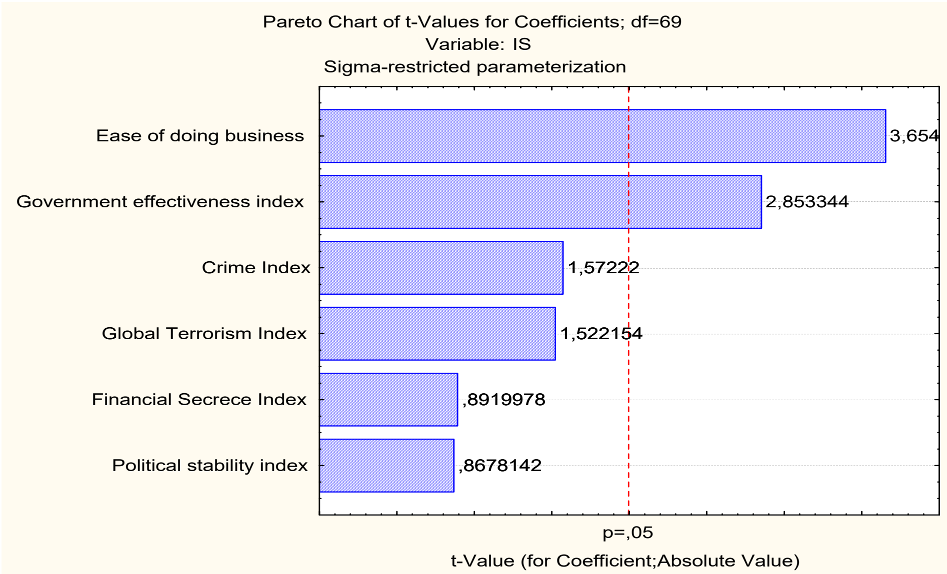


Figure 2. Pareto diagram of t-values of the significance of the influence of indicators of countries' ability to counteract financial fraud on the integral index of cybersecurity

Source: authors' calculations.

Based on the data in Figure 1, it can be argued that only two indicators have a statistically significant influence: the Government Efficiency Index and the Ease of Doing Business, since their significance levels for the Fisher criterion are less than 0.05. The Ease of Doing Business has the largest contribution to the overall model, as indicated by the largest sum of squared deviations SS (1068.59) and the smallest p-value of 0.000499. The next indicator with a statistically significant impact is the Government Efficiency Index, with SS=651.48 and the p-level of 0.0057. A three-dimensional projection is required to build a phase portrait, showing different variations: a node, a focus or a saddle.

The node corresponds to the most balanced system's state characterizing its ability to be in balance and maintain its structure. That is, the components of the cyber security system and the system of countering financial crimes are at the same level of development, which enables their easy integration. A node can be stable or unstable. Stability means the system's ability to return to a state of equilibrium after being brought out of it under the influence of various factors. For example, the emergence of armed conflicts in one country, which will increase the number of migrants, should not significantly affect the rise in crime in the countries where they migrate. Instability is the opposite of stability. A focus also corresponds to a balanced state, but its achievement is more complicated than a node. It can also be stable and unstable. The saddle suits only for unbalanced systems, for which it is impossible to ensure the convergence of the cyber security system and the system for combating financial crimes. Systems for which the phase portrait has the appearance of a "saddle" are immature, that is, not ready for the convergence of the cyber security system and the system for countering financial fraud. "Stable node" corresponds to mature systems, and "unstable node" is mature, but some of their components are unstable. "Stable focus" characterizes systems that are not yet mature enough but have fairly balanced elements. "Unstable focus" corresponds to insufficiently mature systems with individual unbalanced components.

The use of a two-dimensional space would complicate the interpretation of the results. For this reason, further calculations are made using three indicators. The third most important indicator of countries' ability to combat financial fraud is the Crime Index. Although its p-value is equal to 0.12, meaning it is not statistically significant, its nonlinear combination will be used in further calculations, which will be significant for the model. The significance of the factors under consideration can be visually confirmed by the Pareto diagram of t-values showing which indicators of countries' ability to counteract financial threats have a statistically significant impact on the integral cybersecurity index (Figure 2). The Pareto diagram not only shows the statistically significant influence (regressors) of the integral cybersecurity index but also orders them by the power of influence. This statistical toolkit graphically interprets the 80/20 rule, highlighting 80% of the influential environmental factors, particularly the Government Efficiency Index, Ease of Doing Business Index, and the Crime Index, which are relevant and selected for further research.

In the third stage, the Statistica software package is used to determine the specification of the nonlinear regression dependence of the integral cybersecurity index on relevant predictors: the Government Efficiency Index, the Ease of Doing Business Index, the Crime Index. By applying the stepwise inclusion method, a statistically significant dependence is revealed in the form of a square root for the Government Efficiency Index, a logarithmic dependence for the Ease of Doing Business Index, and a quadratic dependence for the Crime Index Index (Figure 3). In the case of the Government Efficiency Index, owing to the presence of negative values in the input data, we consider the dependence of the integral cybersecurity index on this indicator only as part of a multiplicative dependence.

N=58	Regression Summary for Dependent Variable: IS (Spreadsheet1.sta) R= ,80929115 R?= ,65495216 Adjusted R?= ,62891081 F(4,53)=25,150 p<,00000 Std.Error of estimate: 9,2027					
	Beta	Std.Err. of Beta	B	Std.Err. of B	t(53)	p-level
Intercept			-229,789	67,41833	-3,40840	0,001255
LN-V9	0,421814	0,111687	61,729	16,34451	3,77675	0,000404
SQRV8	0,401105	0,126867	17,088	5,40484	3,16163	0,002595
V10**2	-0,187681	0,088620	-0,003	0,00128	-2,11782	0,038898
1/V8	0,150132	0,093801	0,172	0,10727	1,60054	0,115423

Figure 3. Results of regression statistics of dependence of the integral cybersecurity index on relevant predictors: Government Efficiency Index, Ease of Doing Business Index, Crime Index

Source: authors' calculations.

Based on the results of the specification of the dependence of the integral cybersecurity index on relevant predictors, which is expressed as logarithmic, quadratic functions, and the multiplicative dependence of the three indicators, we formalize the indicated nonlinear dependence. The results are presented in Table 7.

Table 7. Results of statistical analysis of the dependence of the integral cybersecurity index on relevant predictors

Specification	Coefficients	Standard error	t-statistics	p-value	Lower 95%	Upper 95%
Y- cross section	-108.6929	56.5889	-1.9207	0.0587	-221.5009	4.1151
ln(EDI)	35.3774	13.2591	2.6682	0.0094	8.9459	61.8089
CI^2	-0.0019	0.0012	-1.6080	0.1122	-0.0037*	-0.00004*
GEI*EDI*CI	0.0028	0.0008	3.5848	0.0001	0.0012	0.0043

** The confidence level is 88%.*

Source: authors' calculations.

Based on data in Table 7 the following regression model can be formulated (4):

$$IS = -108.69 + 35.3774 \cdot ln\,(EDI) - 0.00188 \cdot CI^2 + 0.00277 \cdot GEI \cdot EDI \cdot CI \tag{4}$$

where IS – the integral cybersecurity index;

GEI – the Government Efficiency Index,

EDI – the Ease of Doing Business Index,

CI – the Crime Index.

The statistical significance of ln(EDI) and GEI*EDI*CI was confirmed with p less than 0.05. The p-value for CI^2 exceeds 0.05, but this excess is not critical enough, so statistical significance can be defined for a confidence interval. It was found that the confidence interval with a probability of 88% does not contain a zero value, enabling the use of CI^2 to construct a phase portrait. The coefficient of determination for this model is 62.73%, while the value of the Fisher criterion of 40.40 exceeds the critically acceptable level.

In the fourth stage, the MathCAD application software package was used. The following analysis of the dynamic stability of a CFCF system and the construction of phase portraits of its/their maturity is based on the non-linear function (5), obtained based on the non-linear model (4):

$$f(gei, edi, ci) := -108.693 + 35.37739 \cdot \ln(edi) - 0.00188 \cdot ci^2 + 0.002774 \, gei \cdot edi \cdot ci \quad (5)$$

Based on function (5), we model a system of differential equations (6) that characterize the behaviour of a dynamic CFCF system:

$$\frac{d}{dgei} f(gei, edi, ci) \rightarrow 0.002774 \, ci \cdot edi$$

$$\frac{d}{dedi} f(gei, edi, ci) \rightarrow \frac{35.37739}{edi} + 0.002774 \, ci \cdot gei$$

$$\frac{d}{dci} f(gei, edi, ci) \rightarrow -0.00376 \, ci + 0.002774 \, edi \cdot gei \quad (6)$$

The above three differential equations (6) can be used to establish relationships between variables GEI (Government Performance Index), EDI (Ease of Doing Business), CI (Crime Index) and their first partial derivatives $\frac{d}{dgei} f(gei, edi, ci), \frac{d}{dedi} f(gei, edi, ci), \frac{d}{dci} f(gei, edi, ci)$.

Based on the non-linear approach underlying bifurcation theory, we construct phase portraits of the integral cybersecurity index, where phase trajectories are represented as projections on pairwise planes: the Government Efficiency Index – the Ease of Doing Business, the Ease of Doing Business – the Crime Index, the Government Efficiency Index – the Crime Index. Equations (7) were constructed using the MathCad mathematical analysis software:

$$
\mathrm{Faza}(gei0, edi0, ci0, dt, N) := \left|
\begin{array}{l}
\left(gei_0 \leftarrow gei0 \quad edi_0 \leftarrow edi0 \quad ci_0 \leftarrow ci0\right) \\
\text{for } k \in 0..N \\
\quad \left|
\begin{array}{l}
fff \leftarrow f\left(gei_k, edi_k, ci_k\right) \\
gei_{k+1} \leftarrow \left[gei_k + dt\cdot\left(0.002774 ci_k \cdot edi_k\right)\right] \\
edi_{k+1} \leftarrow \left[edi_k + dt\cdot\left(\dfrac{35.37739}{edi_k} + 0.002774 ci_k \cdot gei_k\right)\right] \\
ci_{k+1} \leftarrow \left[ci_k + dt\cdot\left(-0.00376 ci_k + 0.002774 edi_k \cdot gei_k\right)\right]
\end{array}
\right. \\
\left(gei \quad edi \quad ci\right)
\end{array}
\right. \tag{7}
$$

To visualize formula (7), which represents "phase portraits" of the state of a CFCF system, and subsequently identify its state as one of the three types (saddle, node, or focus), we consider various possible values as factors (Government Efficiency Index, Ease of Doing Business, Crime Index), and the value of the function of the integral cybersecurity index with a given level of accuracy based on the specified number of implementation points:

$$
\begin{aligned}
&(gei1 \quad edi1 \quad ci1) := \mathrm{Faza}(1.6, 80, 42, 0.01, 100) \\
&(gei2 \quad edi2 \quad ci2) := \mathrm{Faza}(1.45, 78, 20, 0.01, 100) \\
&(gei3 \quad edi3 \quad ci3) := \mathrm{Faza}(0.18, 68, 36, 0.01, 100) \\
&(gei4 \quad edi4 \quad ci4) := \mathrm{Faza}(0.43, 56, 51, 0.01, 100) \\
&(gei5 \quad edi5 \quad ci5) := \mathrm{Faza}(-0.32, 50, 52, 0.01, 100) \\
&(gei6 \quad edi6 \quad ci6) := \mathrm{Faza}(-0.45, 57, 70, 0.01, 100)
\end{aligned} \tag{8}
$$

We plug the actual values of the input data (formulas 8) into relationships in order to formalize the phase portraits (7). As a result, we visualize (the first ratio of formulas (8)) the nonlinear dependence of the integral cybersecurity index on relevant factors in the pairwise planes "Government Efficiency Index – Ease of Doing Business" (left fragment of Figure 4) and "Ease of Doing Business - Crime Index" (right fragment of Figure 4).

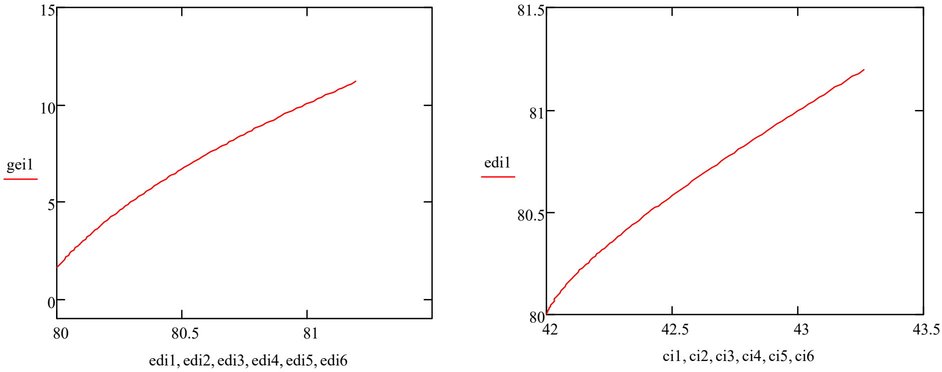


Figure 4. Curves of nonlinear dependence of the integral cybersecurity index on the relevant factors in the planes "Government Efficiency Index - Ease of Doing Business" (left box) and "Ease of Doing Business – Crime Index" (right box)

Source: authors' calculations.

Let us now analyse the phase portraits of a dynamic CFCF system on the entire set of values of input indicators (formulas (8)). First, we consider the system's phase portrait represented in the plane "Government Efficiency Index – Ease of Doing Business" (Figure 5). This phase portrait shows the bifurcation type classified as "unstable focus", i.e., the unstable state of the system. If one parameter changes significantly and the value of another parameter is fixed, the system is in a state of non-equilibrium.

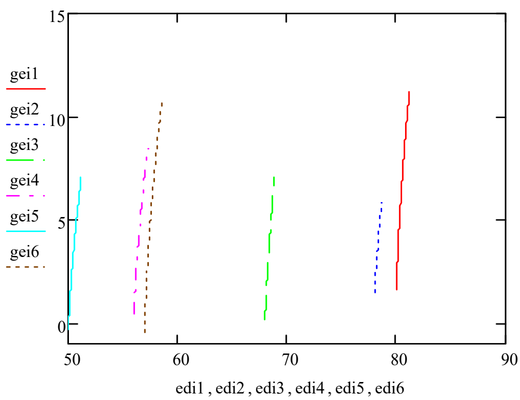


Figure 5. Phase portrait ("unstable focus") of a dynamic CFCF system in a state of non-equilibrium represented in the plane "Government Efficiency Index – Ease of Doing Business"

Source: authors' calculations.

Let us now consider the phase portrait of a dynamic CFCF system represented in the plane "Ease of Doing Business – Crime Index" (Figure 6). Once again, it is in a state of non-equilibrium, classified as "unstable focus".

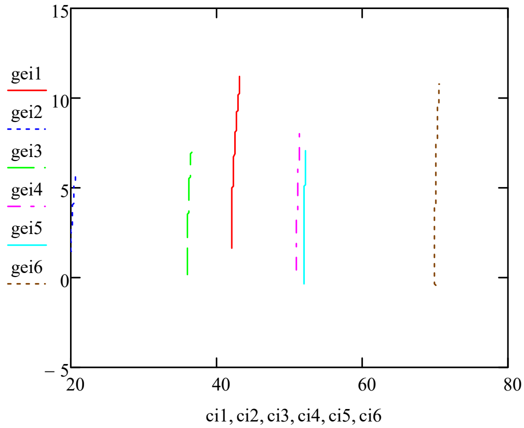


Figure 6. Phase portrait ("unstable focus") of a dynamic CFCF system in a state of non-equilibrium represented in the plane "Ease of Doing Business – Crime Index"

Source: authors' calculations.

The non-equilibrium state of a dynamic CFCF system in the form of a phase portrait classified as "unstable node" can be observed in the plane "Government Efficiency Index – Crime Index", which is shown in Figure 7.

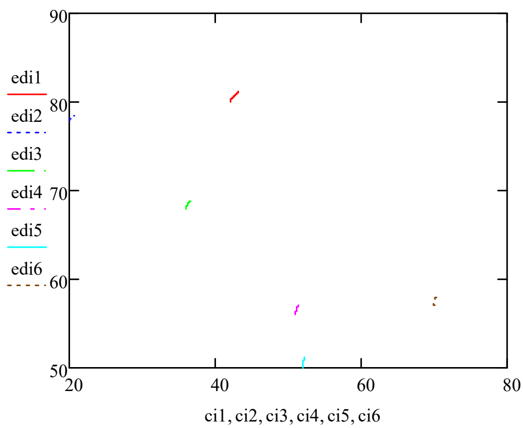


Figure 7. Phase portrait ("unstable node") of a dynamic CFCF system in a state of non-equilibrium represented in the plane "Government Efficiency Index – Crime Index"

Source: authors' calculations.

Bifurcation analysis of the maturity of the current global CFCF system and its phase portraits classified as "unstable focus" and "unstable node" (Figures 5-7) indicate the system under consideration is in a state of maturity, but an incomplete balance of the components in the plane of "Government Efficiency Index – Ease of Doing Business" and "Ease of Doing Business – Crime Index", as well as insufficient maturity and incomplete balance of the "Government Efficiency Index – Crime Index" elements.

In the fifth stage, we determine the specification of the non-linear regression dependence of the integral cybersecurity index on relevant predictors: the Government Efficiency Index, Ease of Doing Business, and the Crime Index. We define the specification from the first relevant feature – the Government Efficiency Index. The integral cybersecurity index determined by the Sundarovsky method is chosen as the resultant feature, and the polynomial (second and third steps), inverse, trigonometric dependencies of the Government Efficiency Index are chosen as factorial ones. The results of the regression analysis are presented in Table 8.

Table 8. Results of statistical analysis of the dependence of the integral cybersecurity index on the Government Efficiency Index

Specification	Coefficients	Standard error	t-statistics	p-value	Lower 95%	Upper 95%
Y- cross section	47.8568	43.4523	1.1014	0.2746	-38.8281	134.5417
GEI*EDI*CI	0.0004	0.0015	0.2402	0.8109	-0.0026	0.0033
GEI^2	-5.1192	18.2560	-0.2804	0.7800	-41.5390	31.3006
GEI^3	1.7633	2.0450	0.8623	0.3915	-2.3163	5.8428
1/GEI	0.08037	0.09631	0.83447	0.40690	-0.1118	0.2725
Sin(GEI)	14.1229	6.3637	2.21929	0.0298	1.4276	26.8182
Cos(GEI)	-16.857	43.9954	-0.3832	0.7028	-104.625	70.9114

Source: authors' calculations.

Given the p-value of 0.0298 (Table 8), the variable sin(GEI) is statistically significant at the alpha level of 0.05. Therefore, a sinusoid is selected as a specification of the dependence of the integral cybersecurity index on the government efficiency index in further calculations.

Next we define the specification of the non-linear dependence of the integral cybersecurity index on the second relevant feature – the Ease of Doing Business. As in the previous case, we treat the integral cybersecurity index determined by the

Sundarovsky method as an effective feature, and polynomial (second and third steps), inverse, logarithmic, square root, trigonometric dependencies of the Ease of Doing Business as factorial ones. Applying the regression analysis tools, we obtain the result presented in Table 9.

Given the p-values in Table 9, it can be argued that there is no statistically significant variable at the alpha level of 0.05. However, a confidence interval calculated for the cubic dependence of the performance attribute on the Ease of Doing Business variable (the smallest p-value equal to 0.1773) does not contain a zero value with a relatively high probability of 82%, which means it can be regarded as statistically significant under the given conditions. Therefore, a cubic dependence is selected as the specification of the dependence of the integral cybersecurity index on the Ease of Doing Business in further calculations.

Table 9. Results of statistical analysis of the dependence of the integral cybersecurity index on the Ease of Doing Business

Specification	Coefficients	Standard error	t-statistics	P-value	Lower 95%	Upper 95%
Y-cross section	-215579.91	167427.8	-1.2876	0.202	-549676.79	118516.97
EDI^2	5.39	4.03	1.3386	0.185	-2.65	13.44
EDI^3	-0.02	0.01	-1.3634	0.177	-0.0373*	-0.0001*
1/EDI	890050.46	692798.4	1.2847	0.203	-492407.17	2272508.10
ln(EDI)	99794.06	77102.97	1.2943	0.200	-54062.51	253650.63
$EDI^{0,5}$	-28814.34	22124.27	-1.3024	0.197	-72962.64	15333.95
Sin(EDI)	0.24	1.78	0.1359	0.892	-3.31	3.80
Cos(EDI)	0.86	1.62	0.5323	0.596	-2.36	4.08

** The confidence level is 82%.*

Source: authors' calculations.

Finally, we define the specification of the nonlinear dependence of the integral cybersecurity Index on the third relevant feature – the Crime Index. We treat the integral cybersecurity index determined by the Sundarovsky method as a resultant feature, and polynomial (second and third degree), inverse, trigonometric dependencies of the Crime Index as factorial ones. Applying the regression analysis tools, we obtain the result presented in Table 10.

Table 10. Results of statistical analysis of the dependence of the integral cybersecurity index on the Crime Index

Specification	Coefficients	Standard error	t-statistics	P-value	Lower 95%	Upper 95%
Y-cross section	7828.5624	5651.85	1.3851	0.1705	-3449.52	19106.65
CI^2	-0.5685	0.40	-1.4118	0.1626	-1.37	0.24
CI^3	0.0027	0.00	1.3904	0.1690	-0.00	0.01
1/CI	-24362.1135	18011.51	-1.3526	0.1807	-60303.52	11579.29
ln(CI)	-4624.8443	3329.88	-1.3889	0.1694	-11269.52	2019.83
$CI^{0,5}$	1679.5597	1203.49	1.3956	0.1674	-721.97	4081.08
Sin(CI)	-3.2419	2.36	-1.3765	0.1732	-7.94	1.46
Cos(CI)	5.7206	2.37	2.4112	0.0186	0.99	10.45

Source: authors' calculations.

Given the p-value of 0.0186 (Table 10), the variable Cos(CI) is statistically significant at the alpha level of 0.05. Therefore, a cosine wave is selected as the specification of the dependence of the integral cybersecurity Index on the Crime Index in further calculations.

Thus, having defined the specification of the dependence of the integral cybersecurity index on relevant predictors (Government Efficiency Index, Ease of Doing Business, Crime Index) in the form of a sinusoid, cubic dependence, cosine wave, respectively, and also introducing an additional variable of multiplicative influence on the performance feature of all three relevant factors, we construct a nonlinear regression dependence. The results are presented in Table 11.

Table 11. Results of statistical analysis of the dependence of the integral cybersecurity index on relevant predictors

Specification	Coefficients	Standard error	t-statistics	P-value	Lower 95%	Upper 95%
Y-cross section	10.8979	4.1283	2.6398	0.0102	2.6663	19.1294
Sin(GEI)	9.9771	5.0902	1.9601	0.0539	-0.1724	20.1266
EDI^3	0.0001	0.0000	5.6383	0.0000	0.0001	0.0001
Cos(CI)	3.4013	1.6051	2.1191	0.0376	0.2009	6.6017
GEI*EDI*CI	-0.0006	0.0012	-0.4738	0.6371	-0.0030	0.0018

Source: authors' calculations.

Based on the coefficients in Table 11, we construct a regression dependence of the integral cybersecurity index on the relevant predictors: Government Efficiency Index, Ease of Doing Business, Crime Index in the form of the following equation:

$$IS = 10{,}8989 + 9{,}9771 \cdot \sin(GEI) + 0.00008 \cdot EDI^3 + 3.4013 \cdot \cos(CI) - 0.00057 \cdot GEI \cdot EDI \cdot CI \tag{9}$$

The validity and accuracy of equation (9) is confirmed based on the following criteria. The coefficients of the variables are statistically significant, since their p-values are below the alpha level of 0.05, except the coefficient before the variable of the multiplicative influence of the three factors. We retain this variable in the model in the bifurcation analysis of the maturity of the current CFCF system and take into account when building phase portraits of its "equilibrium states" since the multiplicative effect of the three factors is required in a qualitative analysis of bifurcations. The coefficient of determination is equal to 70.59%, which means that 70.59% of the variation of the effective feature of the integral cybersecurity index is explained by the variation of the selected factor features, which is a good result.

In the sixth stage of our study, we construct a nonlinear function (10) using equation (9):

$$f(gei, edi, ci) \coloneqq 10.8978783 + 9.977087 \sin(gei) + 7.643510 \cdot 10^{-5} \cdot (edi^3) + 3.40130281 \cos(ci) - 0.00057478 gei \cdot edi \cdot ci \tag{10}$$

Based on function (10), we model a system of differential equations that characterize the behaviour of a dynamic CFCF system to construct phase portraits of "equilibrium states":

$$\frac{d}{dgei} f(gei, edi, ci) \rightarrow 9.97708769 \cdot \cos(gei) + -0.00057478 \cdot ci \cdot edi$$

$$\frac{d}{dedi} f(gei, edi, ci) \rightarrow 0.000229305 \cdot edi^2 + -0.00057478 \cdot ci \cdot gei$$

$$\frac{d}{dci} f(gei, edi, ci) \rightarrow -3.40130281 \cdot \sin(ci) + -0.00057478 \cdot edi \cdot gei \tag{11}$$

The above three differential equations (11) are used to established relationships between variables GEI (Government Efficiency Index), EDI (Ease of Doing Business), CI (Crime Index) and their first partial derivatives $\frac{d}{dgei} f(gei, edi, ci), \frac{d}{dedi} f(gei, edi, ci), \frac{d}{dci} f(gei, edi, ci)$.

Taking the non-linear approach underlying bifurcation theory, we construct phase portraits of the “equilibrium states” of the integral cybersecurity index where phase trajectories are represented as projections on pairwise planes: the Government Efficiency Index – the Ease of Doing Business, the Ease of Doing Business – the Crime Index, the Government Efficiency Index – the Crime Index. The phase portraits are constructed on the basis of the system of differential equations (12) using the MathCad mathematical analysis software:

$$\mathrm{Faza}(gei0, edi0, ci0, dt, N) := \left| \begin{array}{l} \left(gei_0 \leftarrow gei0 \quad edi_0 \leftarrow edi0 \quad ci_0 \leftarrow ci0\right) \\ \text{for } k \in 0..N \\ \quad \left| \begin{array}{l} fff \leftarrow f\left(gei_k, edi_k, ci_k\right) \\ gei_{k+1} \leftarrow \left[gei_k + dt\cdot\left(9.97708769\cos\left(gei_k\right) + -0.00057478ci_k\cdot edi_k\right)\right] \\ edi_{k+1} \leftarrow \left[edi_k + dt\cdot\left[0.000229305\left(edi_k\right)^2 + -0.00057478ci_k\cdot gei_k\right]\right] \\ ci_{k+1} \leftarrow \left[ci_k + dt\cdot\left(-3.40130281\sin\left(ci_k\right) + -0.00057478edi_k\cdot gei_k\right)\right] \end{array} \right. \\ \left(gei \quad edi \quad ci\right) \end{array} \right. \qquad (12)$$

To visualize the phase portraits of the “equilibrium states” of the CFCF system using formula (12) and then identify its state as one of the three types (saddle, node, or focus), we consider various possible values of the three factors (Government Efficiency Index, Ease of Doing Business, Crime Index) and the value of the function that describes the integral Cybersecurity Index with a given level of accuracy based on the specified number of implementation points:

$$\begin{aligned} (gei1 \quad edi1 \quad ci1) &:= \mathrm{Faza}(1.6, 80, 42, 0.01, 100) \\ (gei2 \quad edi2 \quad ci2) &:= \mathrm{Faza}(1.45, 78, 20, 0.01, 100) \\ (gei3 \quad edi3 \quad ci3) &:= \mathrm{Faza}(0.18, 68, 36, 0.01, 100) \\ (gei4 \quad edi4 \quad ci4) &:= \mathrm{Faza}(0.43, 56, 51, 0.01, 100) \\ (gei5 \quad edi5 \quad ci5) &:= \mathrm{Faza}(-0.32, 50, 52, 0.01, 100) \\ (gei6 \quad edi6 \quad ci6) &:= \mathrm{Faza}(-0.45, 57, 70, 0.01, 100) \qquad (13) \end{aligned}$$

We plug the actual values of the input data (formulas 13) into the relationships in order to formalize the phase portraits (12) and determine the equilibrium points represented in the plane “Government Efficiency Index – Ease of Doing Business” (Figure 8). The equilibrium state of the CFCF system corresponds to the following values of its parameters (the intersection points of the graphs shown in Figure 8): the Government Efficiency Index – 1.4838, Ease of Doing Business – 80.183.

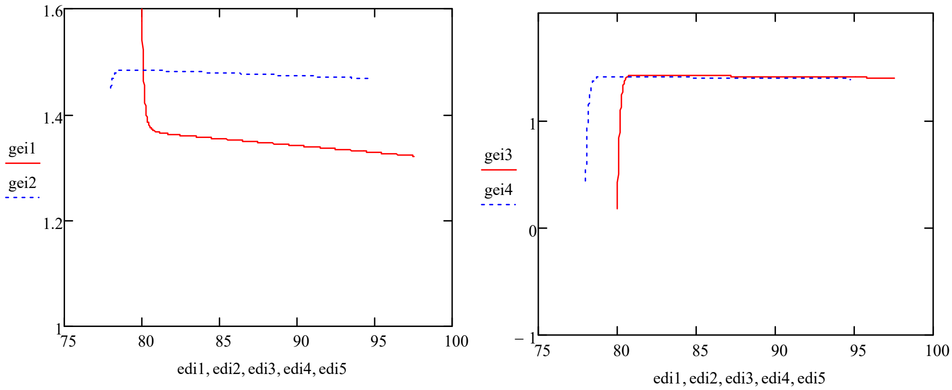


Figure 8. Equilibrium points of the CFCF system represented in the plane "Government Efficiency Index - Ease of Doing Business"

Source: authors' calculations.

Let us now analyse the phase portraits of the dynamic CFCF system on the entire set of values of input indicators (formulas (13)). We first consider the phase portrait of the system represented in the plane "Government Efficiency Index – Ease of Doing Business" (Figure 9). This phase portrait demonstrates the presence of an saddle point characterizing a non-equilibrium state of the CFCF system.

Moving on to the phase portrait represented in the "Ease of Doing Business – Crime Index" plane (Figure 10), we observe that it is in a non-equilibrium state classified as a "saddle". This type of bifurcation indicates an unstable state of the system, i.e., if one parameter changes significantly and the value of another parameter is fixed, the system is in a state of non-equilibrium.

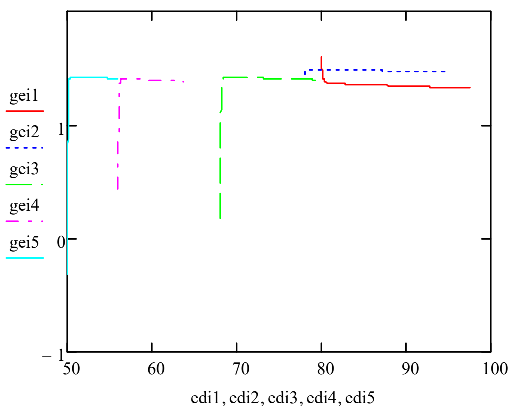


Figure 9. Phase portrait ("saddle") of the dynamic CFCF system represented in the plane "Government Efficiency Index - Ease of Doing Business"

Source: authors' calculations.

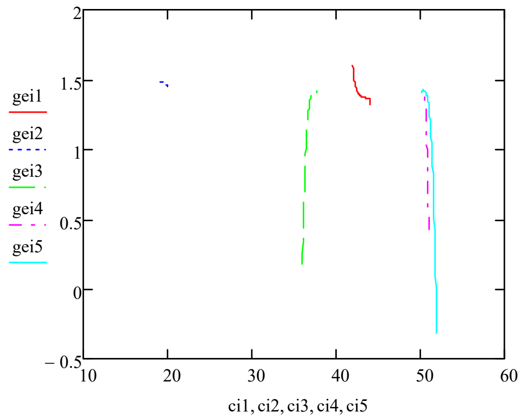


Figure 10. Phase portrait (“saddle”) of the CFCF dynamic system, which is in a non-equilibrium state, represented in the plane “Ease Of Doing Business – Crime Index”

Source: authors’ calculations.

The non-equilibrium state of the dynamic CFCF system as evidenced by a phase portrait classified as “saddle” can also observed in the plane “Government Efficiency Index – Crime Index”, which is shown in Figure 11.

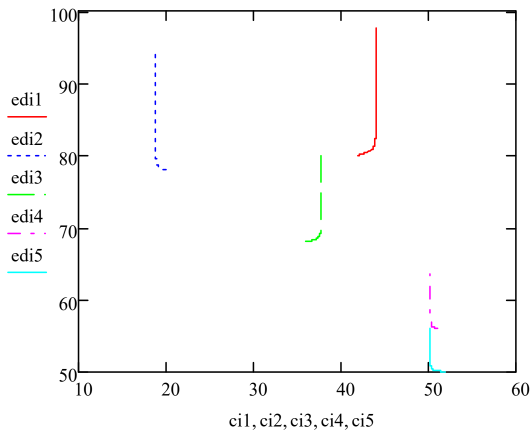


Figure 11. Phase portrait (“saddle”) of the dynamic CFCF system, which is in a non-equilibrium state, represented in the plane “Government Efficiency Index – Crime Index”

Source: authors’ calculations.

The above bifurcation analysis of the maturity of the current CFCF system based on its phase portraits (Figures 9–11) reveals states of non-equilibrium classified as “saddle” for all projections – “Government Efficiency Index - Ease of Doing Business”, “Ease Of Doing Business – Crime Index” and “Government Efficiency Index – Crime Index”. .

5. Conclusions

The purpose of the above analysis was to assess the maturity of the global system for combating financial and cyber fraud with a view to determining its readiness for integration at different levels of state management. Since the studied system is dynamic, i.e. change under the influence of various external and internal factors, a bifurcation analysis was performed involving the construction of phase portraits of its maturity and equilibrium. The resulting phase portraits of the CFCF system's "maturity" were classified as "unstable focus" when represented in the planes "Ease of Doing Business – Crime Index" and "Government Efficiency Index – Ease of Doing Business." A phase portrait classified as "unstable node" was obtained for the plane "Government Efficiency Index – Crime Index".

The results show that the global CFCF system is quite mature according to obtained node and focus phase portraits but unstable. That is, it is significantly affected by the level of crime in a given country, the inefficiency of government decisions, and the lack of opportunities for business development and organization. However, factors such as financial secrecy, political stability, the level of corruption and terrorism do not cause fluctuations. They do not lead to significant changes in the cybersecurity system. The CFCF system based on the integration of cybersecurity and combating financial fraud systems, above all, requires legislative changes, which should improve the living standards of the population and reduce crime in general and financial and cyber fraud in particular. Another strategic factor which must be considered is the creation of opportunities for business development also positively affects the countries' economic processes and foster their economic growth.

The proposed methodology makes it possible to determine points where the system's equilibrium will be reached, but phase portraits classified as "saddle" points indicate that the CFCF system cannot reach a state of equilibrium. (constructed in the context of the respective threeplanes). Changing only one of the parameters will affect this state, provided that another factor has a fixed value. The preliminary conclusions about the system's instability and its lack of its equilibrium are thus confirmed.

To sum up, the detected states of maturity and equilibrium of the CFCF system indicate its sufficient level of maturity, but at the same time, its inability to recover in the planes of "Government Efficiency Index - Ease of Doing Business", "Ease Of Doing Business - Crime Index" and "Government Efficiency Index - Crime Index". That is, there is a need to improve the processes of business regulation and state policy formation in the world countries to counter financial and cybercrimes. In the future, this approach can be recommended to relevant government agencies to form initiatives for developing public financial monitoring and national cybersecurity , as well as to international organizations to improve the strategy of global countermeasures against financial and cybercrimes.

Acknowledgement

This research was funded by the grant from the Ministry of Education and Science of Ukraine (No. s/r 0121U109559, 0121U100467).

References

Akhramovych, V., Shuklin, G., Pepa, Y., Muzhanova, T. and Zozulia, S., (2022). Devising a procedure to determine the level of informational space security in social networks considering interrelations among users. *Eastern-European Journal of Enterprise Technologies*, [e-journal], 1(9–115), pp. 63–74; 10.15587/1729-4061.2022.252135.

Amala Dhaya, M. D. and Ravi, R., (2021). Multi feature behavior approximation model based efficient botnet detection to mitigate financial frauds. *Journal of Ambient Intelligence and Humanized Computing*, [e-journal], 12(3), pp. 3799–3806; 10.1007/s12652-020-01677-w.

Aziz, R. M., Baluch, M. F., Patel, S. and Ganie, A. H., (2022). LGBM: a machine learning approach for Ethereum fraud detection. *International Journal of Information Technology* (Singapore), [e-journal]; 10.1007/s41870-022-00864-6

BBC, (2022). Ukraine cyber-attack: Russia to blame for hack, says Kyiv. [online] Available at: https://www.bbc.com/news/world-europe-59992531 [Accessed 31 May 2022].

Bystray, G. P., Lykov, I. A. and Nikulina, N. L., (2012). Risks assessment and forecasting long time rows of economic indicators. *Economy of Region*, [e-journal] 3, pp. 240–249; 10.17059/2012-3-24.

CyberPeace Institute, (2022). Ukraine: Timeline of Cyberattacks on critical infrastructure and civilian objects. [online] Available at: https://cyberpeaceinstitute.org/ukraine-timeline-of-cyberattacks/, [Accessed 31 May 2022].

Day, M.-Y., (2021). Artificial intelligence for knowledge graphs of cryptocurrency anti-money laundering in fintech. In: Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 13th IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2021. Virtual, Online, 8 November 2021. New York: Association for Computing Machinery; 10.1145/3487351.3488415.

Deloitte, (2019). The connected defense: Elevating the fight against financial crime. Using 4IR technologies to prevent and detect the growing ecosystem of financial crime. [online] Available at: https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/financial-services/deloitte-nl-fsi-elevating-the-fight-against-financial-crime.pdf, [Accessed 31 May 2022].

e-Governance Academy Foundation, (2022). National Cyber Security Index. [online] Available at: https://ncsi.ega.ee/ncsi-index/, [Accessed 31 May 2022].

Euronews, (2022). Ukraine's defence ministry and two banks targeted in cyberattack. [online] Available at: https://www.euronews.com/my-europe/2022/02/15/ukraine-s-defence-ministry-and-two-banks-targeted-in-cyberattack, [Accessed 31 May 2022].

FinCEN, (2009). Mortgage Loan Fraud Connections with Other Financial Crime: An Evaluation of Suspicious Activity Reports Filed By Money Services Businesses, Securities and Futures Firms, Insurance Companies and Casinos. [online] Available at: https://www.fincen.gov/sites/default/files/shared/ mortgage_fraud.pdf, [Accessed 31 May 2022].

Granados, O. and Garcia-Bedoya, O., (2022). Deep Learning-Based Facial Recognition on Hybrid Architecture for Financial Services. In S. Misra, A.K. Tyagi, V. Piuri, L. Garg, ed. 2022. Artificial Intelligence for Cloud and Edge Computing. Springer: Cham. pp. 51-70. 10.1007/978-3-030-80821-1_3.

Granados, O., Vargas, A., (2022). The geometry of suspicious money laundering activities in financial networks. EPJ Data Science, [e-journal], 11(1), article number 6; 10.1140/epjds/s13688-022-00318-w.

Hayble-Gomes, E., (2022). The use of predictive modeling to identify relevant features for suspicious activity reporting. Journal of Money Laundering Control, [e-journal] ahead-of-print; 10.1108/JMLC-02-2022-0034.

Idowu, B. A., Vaidyanathan, S., Sambas, A., Olusola, O. I. and Onma, O. S., (2018). A new chaotic finance system: Its analysis, control, synchronization and circuit design. In: V.-T. Pham, S. Vaidyanathan, C. Volos, T. Kapitaniak, ed. 2018. Nonlinear Dynamical Systems with Self-Excited and Hidden Attractors. Springer: Cham, pp. 271–295; 10.1007/978-3-319-71243-7_12.

Krebsonsecurity, (2022). Report: Recent 10x Increase in Cyberattacks on Ukraine. [online] Available at: https://krebsonsecurity.com/2022/03/report-recent-10x-increase-in-cyberattacks-on-ukraine/, [Accessed 31 May 2022].

Kuzmenko, O., Yarovenko, H. and Radko, V., (2021). Poperednii analiz protsesu konverhentsii system kiberbezpeky ta finansovoho monitorynhu krain *[Preliminary analysis of the convergence process of cyber security systems and financial monitoring of countries. Economy and society]*, [e-journal], 32; 10.32782/2524-0072/2021-32-37. [in Ukrainian].

Li, H., Wong, M.-L., (2021). Grammar-Based Multi-objective Genetic Programming with Token Competition and Its Applications in Financial Fraud Detection. *Natural Computing Series*, [e-journal], pp. 259–285; 10.1007/978-3-030-79553-5_11.

Mahootiha, M., Golpayegani, A.H. and Sadeghian, B., (2021). Designing a New Method for Detecting Money Laundering based on Social Network Analysis. In: 26th International Computer Conference, Computer Society of Iran, 26th International Computer Conference, Computer Society of Iran, CSICC 2021. Tehran, 3 March 2021. New York: Institute of Electrical and Electronics Engineers (IEEE); 10.1109/CSICC52343.2021.9420621.

Matanky-Becker, R., Cockbain, E., (2022). Behind the criminal economy: using UK tax fraud investigations to understand money laundering myths and models. Crime, *Law and Social Change*, [e-journal], 77(4), pp. 405-429; 10.1007/s10611-021-09997-4.

Netzwerk Steuergerechtigkeit, (2022). Financial Secrecy Index - 2018 Results. [online] Available at: https://www.netzwerk-steuergerechtigkeit.de/wp-content/uploads/2018/01/6_fsi-ranking_incl-_eu-tax-havens.pdf, [Accessed 31 May 2022].

OCHA, (2018). Global Terrorism Index (2018). [online] Available at: https://reliefweb.int/report/world/global-terrorism-index-2018, [Accessed 31 May 2022].

Perkhun, L., Sorochynskyi, O., Izosimov, O., (2015). Systema vzaiemodii shakhraiskykh atak ta instrumentiv borotby z nymy. *Scientific Bulletin of National Academy of Statistics, Accounting and Audit*, [e-journal], 1, pp. 50–65; http://194.44.12.92:8080/jspui/handle/123456789/3813. [in Ukrainian].

PwC, (2018). Building a united front on financial crimes. [online] Available at: https://www.pwc.com/gx/en/financial-services/pdf/united-front-financial-crimes-2018-pwc.pdf, [Accessed 31 May 2022].

Riani, M., Corbellini, A. and Atkinson, A.C., (2018). The Use of Prior Information in Very Robust Regression for Fraud Detection. International Statistical Review, [e-journal], 86(2), pp. 205–218; 10.1111/insr.12247.

Rouhollahi, Z., Beheshti, A., Mousaeirad, S. and Goluguri, S., (2021). Towards Proactive Financial Crime and Fraud Detection through Artificial Intelligence and

RegTech Technologies. In: ACM International Conference Proceeding Series, 23rd International Conference on Information Integration and Web Intelligence, iiWAS 2021. Virtual, Online, 29 November 2021 through 1 December 2021. New York: Association for Computing Machinery; 10.1145/3487664.3487740.

Sierikov, A. V., Zubova, O.O., (2010). Marketing as a necessary condition for synergetic management of economic activity. Actual Problems of Economics, 5, pp. 276–283.

Tharani, J.S., Charles, E.Y.A., Hou, Z., Palaniswami, M., and Muthukkumarasamy, V., (2021). In: Proceedings - Conference on Local Computer Networks, 46th IEEE Conference on Local Computer Networks, LCN 2021. Edmonton, 4 October 2021. New York: Institute of Electrical and Electronics Engineers (IEEE); 10.1109/LCN52139.2021.9524878.

The GlobalEconomy, (2022). List of available indicators. [online] Available at: https://www.theglobaleconomy.com/indicators_list.php, [Accessed 31 May 2022].

The World Bank, (2022). World Development Indicators; [online] Available at: https://databank.worldbank.org/reports.aspx?source=2&series=SI.POV.GINI&country=#, [Accessed 31 May 2022].

Tundis, A., Nemalikanti, S. and Mühlhäuser, M., (2021). In: ACM International Conference Proceeding Series, 16th International Conference on Availability, Reliability and Security, ARES. Virtual, Online, 17 August 2021. New York: Association for Computing Machinery; 10.1145/3465481.3469196.

Wilkens, K., Thomas, N. and Fofana, M.S., (2004). Stability of technology stock prices: Evidence of rational expectations or irrational sentiment? Managerial Finance, [e-journal] 30(12), pp. 33–54; 10.1108/03074350410769380.

Xia, P., Ni, Z., Zhu, X., He, Q. and Chen, Q., (2022). A novel prediction model based on long short-term memory optimised by dynamic evolutionary glowworm swarm optimisation for money laundering risk. *International Journal of Bio-Inspired Computation*, [e-journal], 19(2), pp. 77–86; 10.1504/IJBIC.2022.121233.

Yang, Y. and Wu, M., (2021). Explainable Machine Learning for Improving Logistic Regression Models. In: IEEE International Conference on Industrial Informatics, 19th IEEE International Conference on Industrial Informatics, INDIN 2021. Mallorca, 21 July 2021. New York: Institute of Electrical and Electronics Engineers (IEEE); 10.1109/INDIN45523.2021.9557392.

Yarovenko, H., (2021). Informatsiina bezpeka yak draiver rozvytku natsionalnoi ekonomiky | Information security as a driver of national economy development. DSc. Sumy State University. Available at: essuir.sumdu.edu.ua/handle/123456789/83664, [Accessed 31.05.2022]. [in Ukrainian].

Appendix

Final calculations of the integral cybersecurity index, determined by the Sundarovsky method

Country	IS	Country	IS
Australia	57,03	Liberia	3,33
Austria	57,70	Lithuania	60,61
Bahrain	29,76	Luxembourg	58,97
Barbados	12,32	Malaysia	55,41
Belgium	61,32	Malta	39,17
Bolivia	15,94	Mauritius	44,16
Botswana	15,75	Mexico	31,91
Brazil	35,78	Montenegro	33,17
Brunei Darussalam	28,56	Netherlands	65,46
Bulgaria	42,09	New Zealand	53,48
Canada	55,98	North Macedonia	39,64
Chile	37,39	Norway	59,60
China	36,02	Panama	29,08
Costa Rica	12,15	Paraguay	33,70
Croatia	54,06	Philippines	29,63
Cyprus	39,86	Poland	51,48
Czech Republic	50,66	Portugal	53,06
Denmark	63,60	Romania	42,83
Dominica	22,71	Russian Federation	50,74
Dominican Republic	26,45	Saudi Arabia	50,99
Estonia	65,41	Seychelles	14,84
Finland	65,08	Singapore	65,94
France	63,45	Slovakia	51,23
Germany	62,90	Slovenia	46,79
Ghana	22,17	South Africa	27,30
Greece	46,64	Spain	60,68
Grenada	13,80	Sweden	55,84
Guatemala	12,02	Switzerland	61,78
Hungary	49,49	Tanzania	17,16
Iceland	40,08	Thailand	39,41
India	37,28	Trinidad and Tobago	9,65
Indonesia	34,80	Turkey	46,17
Ireland	54,54	Ukraine	43,00
Israel	55,18	United Kingdom	64,85
Italy	53,19	United States	65,35
Japan	58,31	Uruguay	42,02
Kenya	29,73	Vanuatu	13,21
Latvia	52,54	Venezuela	20,47

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 259–276, DOI 10.59170/stattrans-2023-014
Received – 01.07.2022; accepted – 15.12.2022

War-driven wave of Ukrainian emigration to Europe: an attempt to evaluate the scale and consequences (the view of Ukrainian researchers)

Ella Libanova[1], **Oleksii Pozniak**[2]

ABSTRACT

The article aims to evaluate the scale and consequences of the emigration of Ukrainians triggered by the military aggression of the Russian Federation. The paper also attempts to determine the composition of the refugees.

The first weeks of the military aggression saw the most active departure of the population from Ukraine, after that the number of those seeking refuge decreased. According to the estimation of the Ptukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine based on the data from the State Border Guard Service, the number of 'refugees from the war in Ukraine' reached 3 million as of the end of June 2022. The emigrants are mainly residents of Kyiv and Kharkiv, which results in a much higher specific weight of people with higher education than the national average. This fact combined with an orientation towards employment rather than social assistance (a mentality that is relatively close to Europeans), suggests a high probability (especially compared to the same emigrants from other countries, including Syria and Afghanistan) for most Ukrainian women to successfully adapt to life across the border. This is especially true for those who came to Poland, due to the minimal linguistic and cultural differences between the countries.

The potential amount of irreversible migration losses, depending on the military and economic factors, ranges from 600−700 thousand to 5−5.5 million people. Considering the fact that approximately 3 million Ukrainians had already been staying (working) abroad before 2022, the war is likely to result in a demographic catastrophe for Ukraine, whose demographic potential has been utterly exhausted.

Key words: migration, forced migration, refugees, temporary protection, return of migrants, migration policy.

[1] Ptoukha Institute for Demography and Social Studies of the NAS of Ukraine, Director, Academician of NAS of Ukraine, Ukraine. E-mail: libanova@ukr.net
[2] Ptoukha Institute for Demography and Social Studies of the NAS of Ukraine, Head of Sector, Ukraine. E-mail: olex_poznyak@ukr.net.

© E. Libanova, O. Pozniak. Article available under the CC BY-SA 4.0 licence

1. Introduction

The start of a full-scale war waged by the Russian Federation against Ukraine has resulted in millions of Ukrainians losing their homes and jobs, occupation of certain regions of Ukraine that were controlled by the Ukrainian government until February 24, 2022, and significant civilian casualties. These processes have led to the unprecedented levels of the forced migration since the Second World War: the residents of Ukrainian territories adjacent to the war or occupied zone have begun to move en masse to both foreign countries and relatively safe regions of Ukraine. These processes require scientific understanding; the assessment of their scale and consequences becomes extremely important. Such assessments are necessary to identify the effects of large-scale forced migration, to develop recommendations for their regulation and for identifying prospects for the post-war economic recovery.

Ukraine has accumulated considerable experience in assessing migration losses, especially losses from external labour migration and military intervention of the Russian Federation in Donbas since 2014 (Ptoukha Institute, 2018), as well as the COVID-19 pandemic (Libanova, E. M., & Pozniak, O. V., 2020). But the military migration crisis of 2022 is far ahead of all migration phenomena observed in Ukraine during 1991–2021 in terms of scale and level of threat.

The presented article focuses on estimating the number of forced external migrants and the issue of forced relocation within the country as well as labour migration of Ukrainian citizens abroad are outside of the objectives of this study.

The study is based mainly on data from Ukrainian statistical sources whose data formation procedures are well studied by the authors.

The research involves finding answers to the following questions:

1. How the migration behaviour of Ukrainians changed with the start of a full-scale war?
2. What are the real scales of forced migration abroad and the return of forced migrants?
3. How catastrophic the demographic consequences of forced migration will be for Ukraine?
4. What should be the migration policy of Ukraine in new conditions?

2. Methodology

In conditions of active hostilities in a large part of the country, the main problem of evaluation is the lack of information: the production of official statistics is stopped, conducting sample surveys, at least "face to face", is extremely complicated.

The situation is exacerbated by the lack of a census since 2001 and, consequently, the lack of reliable data on the number and composition of the Ukraine's population, especially, its territorial location. Due to the COVID-19 pandemic and especially since the beginning of hostilities, the difficulties in conducting large-scale sample surveys during 2020–2021 have also had an effect, but above all large-scale movements of the population, and outside Ukraine in particular. Estimates of the number of those who left vary from 3 to 6 million people, and according to the expert opinions based on incomplete data from the State Border Guard Service of Ukraine, about a third of them are children. It is clear, that it is extremely difficult to provide a reliable survey of these people through telephone interviews. But a much more significant impact on representativeness may be due to either the lower willingness to participate in polls of the citizens with "pro-Russian" sentiments, or the insincerity of those who have taken part in the poll (given the huge wave of patriotism and prevailing opinion in the society, consciously or even subconsciously, the unwillingness of a large part of the Ukrainian population to give sincere answers). Accordingly, there are good reasons to believe that the artificial increase in the number of respondents who are loyal to their own state is significant. According to the Kyiv International Institute of Sociology, such an increase is at least 4–6% (Kyiv International Institute of Sociology, 2022a). A number of surveys of Ukrainian migrants have been conducted in recipient countries. However, even in these cases, it is difficult to ensure the representativeness of the sample, since different categories of forced migrants may show different degrees of readiness to participate in surveys. That is, in modern conditions, each individual survey cannot be considered as fully representative, moreover, the results of such surveys quickly become outdated.

Theoretically, "big databases" can be used, including the data from the mobile providers, to assess the extent and direction of resettlement and for further sample design, but this requires overcoming a number of barriers, the protection of private data in particular. Moreover, limitation for the application of this method is a significant number of phones that were lost/abandoned by Ukrainian citizens, as well as "squeezed" by the occupants, the need to take into account the possession of two (several) phones by one person, lack of telephones among the representatives of the youngest and older ages. Besides, the amount of the holders of Ukrainian telephone numbers who are abroad will include not only those who left the war, but also external labour migrants, who were working abroad before the active phase of the war. Therefore, we now have to focus on rather eclectic information from open sources, and (in key cases) rely on expert opinions and assessments.

3. Transformations of migratory behavior of the population in wartime

Migration processes have traditionally been an important factor in the human activity, the economy in particular, of every country, that is subject to armed aggression, but they are no less important after the war. Part of the population inevitably leaves their homeland, and sometimes without returning even after the establishment of peace, some move from one territory to another. Russia's military aggression, which began on February 24, 2022, with the bombing of a large number of settlements throughout Ukraine and caused massive destruction of civilian infrastructure, especially housing, has led to large-scale relocations to the safer western regions and beyond.

With the expansion of the area of active hostilities and the aggressor's open desire to cause as much damage to the human and economic potential of Ukraine as possible, the number of evacuees grew: the same cities were both the centres of gravity and centres of mass departure (Odesa, Dnipro, Lviv, and, from the end of March, Kyiv as well). But it seems that large-scale migration movements will not be limited to the war period. Awareness of a large part of the population, entrepreneurs in particular, of the risks posed by the proximity of the borders of an aggressive neighbour is likely to lead to mass relocation to the western and central regions after the end of hostilities.

Priority will be given to building a resilient economy: renovation of the industrial and social infrastructure, and further economic activity while maximizing protection against the risks and threats, posed by the actions of an aggressive neighbour. Besides investment, this will require territorial compliance of labour supply and demand (Libanova, E. M., Pozniak, O. V., & Tsymbal, O. I., 2022).

Undoubtedly, the worst losses are the mass deaths of soldiers and civilians, especially children, the loss of the potential births due to the deaths of pregnant women, and the excess mortality associated with the deteriorating living conditions of many Ukrainians. Moreover, starting from the end of 2022, a significant decrease in the births is expected due to the postponement of births by the large contingents of the population during the crisis. However, migration losses, while not so terrible in the humanistic dimension, are more numerous, measured not even in hundreds of thousands, but in millions of people, who left the country fleeing the horrors of the war, saving their own lives and the lives of their children, parents and grandchildren. The issue of their maximum return from abroad should become one of the cornerstones of the post-war development of both the Ukrainian economy and Ukrainian society as a whole.

The strategy and tactics of this development, in turn, should be based on the assessment of the current migration processes and their subsequent course, above all, on the quality information about the initial scale of departure, and composition of those who were leaving the country, their plans and intentions at the moment of departure, and their change over the time, forecasts of the future returns. The most reliable source of information on the extent of external migration seems to be the border crossing records.

Lack of jobs in the places of resettlement within Ukraine (as the western regions, where the majority of internally displaced persons are concentrated, traditionally differ in this, and that is the main factor in the spread of labour migration there) forces to consider the option of going abroad even those who were not going to.

4. Dynamics of forced departure from Ukraine

According to the State Border Guard Service, over 4.8 million people left Ukraine in the period from February 24 to June 5 (State Border Guard Service, 2022). The data refer to the borders with the EU and Moldova, as the checkpoints on the borders with Russia, Belarus and the Transnistrian border with Moldova have been closed since the beginning of the war, and include foreigners who actively were leaving Ukraine during the first weeks of the war (e.g. among 110 thousand persons who crossed the borders of Ukraine with Moldova, Poland, Romania, Slovakia and Hungary on February 27, 2020, foreigners accounted for about 18%). The data show the number of crossings of the border, so some people can enter the statistics many times. Moreover, 64.7% of those who left crossed the Ukrainian-Polish border, and, say, from March 3 to March 10, this figure exceeded 70%. Of course, not all of these people stayed in Poland, many moved on, first of all, to Germany, but there is no doubt that Poland became the centre of refuge for Ukrainian "refugees from the war": according to various estimates, 40 to 50% of Ukrainians found refuge here, from 10 to 15% stopped in Germany, 5–8% – in the Czech Republic. In general, the final place of evacuation was chosen mostly spontaneously, more often due to the presence of relatives and acquaintances. Also, if the main motive for focusing on Poland is proximity (territorial, cultural, linguistic), Germany was chosen due to better social conditions for refugees, and the Czech Republic – due to the greater employment opportunities (4Service Group, 2022; Ilko Kucheriv Democratic Initiatives Foundation, 2022).

In view of the almost complete lack of data on the population of those who left Ukraine due to the war, the lack of information about those who refused to participate in the survey and the inability to build a correct sample, respondents' answers should not be considered representative and applied to all those who left. At the same time, given the rather large number of contingents of interviewed "military emigrants" (from 0.5 to 3.5 thousand people), it is correct enough to consider them relevant to illustrate the processes taking place, and their changes in particular.

Most of those who went abroad did so in the first weeks of the war. In general, the dynamics of daily departures from Ukraine indicates that there were 3 periods (Figure 1):

- peak departure – February 24 – March 9: (on average, 107 thousand persons were leaving the country per day, and in total 1.5 million persons left in 14 days);
- stabilization – March 10 – May 9: (on average, 28 thousand persons were leaving the country per day and in total, more than 2.5 million persons left in 59 days);
- predominance of the number of entries over the number of exits – from May 10 (on average, 32 thousand persons were leaving the country during May 10 – June 06 per a day, and in total, more than 900 thousand persons left in 28 days).

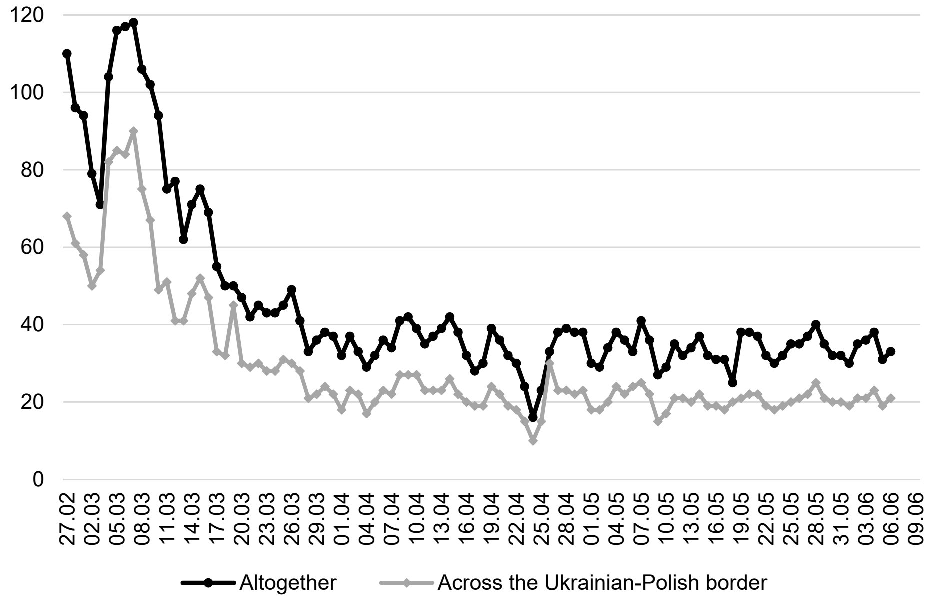


Figure 1. The number of persons who left Ukraine through the Ukrainian-Polish border, thousand persons.

Source: State Border Guard Service of Ukraine.

According to a survey conducted by the Gradus Research Company in May 2022, 37% of city residents with a population of 50 thousand people or more left their homes because of the war. About a quarter of migrants went abroad, while the vast majority relocated within Ukraine (Gradus Research Company, 2022).

Quite naturally and as expected, mostly women, children and the elderly were leaving Ukraine. Namely, in the total number of persons who left Ukraine for the period of February 24 – April 10, the share of women is 66%, persons under 18 years of age –

31%, and over 60 years of age – 13%. According to the structural characteristics of those who left, the first day of the war (February 24) stands out, when the ban on the departure of men aged 18–60, who can be mobilized, had not yet been enacted. Accordingly, among those who left Ukraine on February 24, men of the corresponding age accounted for more than 40.5%, while during the period of February 25 – April 10 – 11.2%.

Due to the fact that the residents of the cities (more than 90%) and persons aged 18–59 years (56%) dominate among those who left, most of them (more than 60%) have a higher education or academic degree. This is, evidently, the basis of their high competitiveness in the labour market, even if they have to work not in the obtained profession.

5. The flow of arrivals of the population to Ukraine

The return flow is also very large: a total of 2.8 million people have entered Ukraine since the beginning of the war, the share of Ukrainian citizens from March 2 to June 2 is 87.8%. At first, in the first 2 weeks, men were actively returning from abroad, their share reached 80% among 200 thousand people who entered Ukraine, i.e. mostly migrant workers returned to participate in the defence of their country, to help the family in these difficult times, etc.; most of them have preserved their Ukrainian roots and, most likely, have close relatives here. However, already at the turn of March-April, the main part of border crossings began to account for various forms of the pendulum migration.

First, a significant part of those who entered Ukraine in the second half of April were returning to celebrate Easter (which in 2022 fell on April 17 and 24). Thus, on Friday-Saturday, on the eve of Catholic Easter, 74 thousand persons entered Ukraine, or 37 thousand on average per day; on the eve of Orthodox Easter – 73 thousand persons, or 36.5 thousand persons per day, while during the period from Monday, April 18 to Thursday, April 21, an average of 32 thousand persons were entering per day (Figure 2).

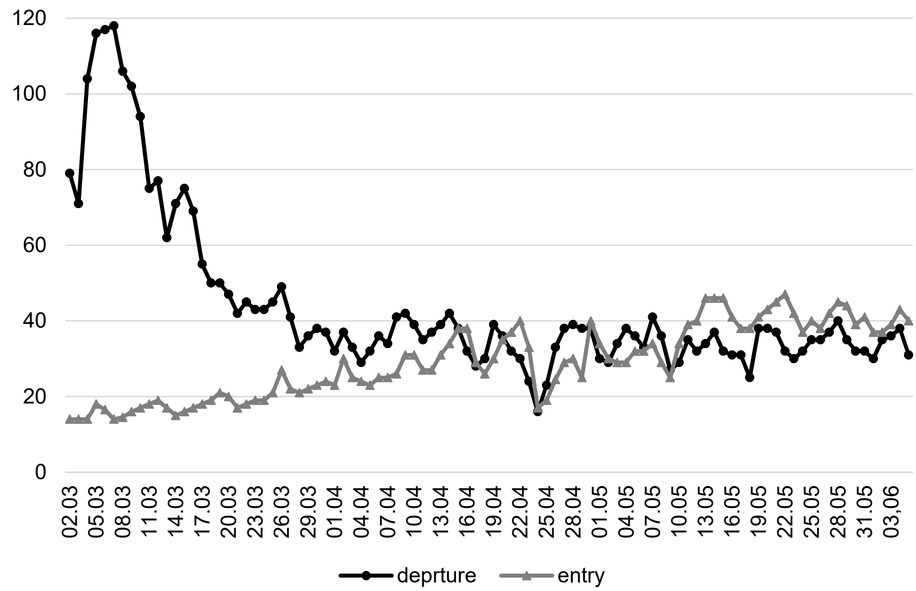


Figure 2. The scale of departures and entries to Ukraine, thousand persons.

Source: State Border Guard Service of Ukraine.

Second, the families evidently desired to see their husbands/parents, and this explains the significantly higher number of entrances to Ukraine on Saturdays (on average, 33 thousand persons entered on Saturday from March 2 to June 6, while, say, on Monday – 26 thousand). The peaks of arrivals on Saturdays began to be recorded as early as Saturday, March 5, and from March 26, the growth of entry into Ukraine on Saturdays increased significantly. The peak values of departures are not so much tied to the days of the week, because the short-term stay of women/children in Ukraine varies from less than a day to several days.

Third, around mid-March a process of shuttle migrations began, in order to withdraw cash abroad from cards of the Ukrainian banks, and from mid-April it became mass; it is a fairly effective commercial tool given exchange rate differences when withdrawing cash in Ukraine and abroad, and, if possible, further resale of foreign currency on the black market. It is clear that it was not so much people who physically crossed the border, as cards issued by Ukrainian banks (according to unofficial estimates, one person crossed the border with dozens of cards). And this flow does not have a significant impact on the overall migration movement, but its existence should not be completely neglected.

Fourth, we cannot ignore the so-called "automobile shuttle migration" related to the import of the used cars for own use and for sale without paying excise duty from

April 5. In particular, in the period of February 27 to March 13, the state border of Ukraine in both directions was crossed by an average of 14.9 thousand cars per day (unfortunately, the State Border Guard Service does not provide information separately on the amount of departure and entry of road transport), while during the period of March 14 – April 13 – 13.7 thousand cars per day, and during April 14 – June 6 – 17.4 thousand cars.

Besides, local border traffic remains widespread (or even has increased) – i.e. crossings of the border by residents of the territory within up to 30 km from the common border with Moldova, Poland, Romania, Slovakia and Hungary, for various social, cultural, family and economic reasons, not related to paid work. In particular, the delivery in small batches (i.e. within the limits not subject to customs duties) of products, and, in the long term, building materials, which became widespread before the war, continues.

Thus, all external migration flows are far from being limited to border crossing by the "war refugees". Those "refugees from the war" who for various reasons could not adapt abroad, whose place of residence in Ukraine was released or became more or less safe due to the cessation of hostilities near it, began to return to Ukraine in late March.

From May 10 (the day after May 9, a sacred date for Russians, analysts and the public expected the possibility of a nuclear strike on that day), 7 thousand more people have been entering Ukraine every day than have been leaving, and the cumulative difference for the period from May 10 to June 6 is 200 thousand persons.

In total, during the war, the total number of Ukrainians ("refugees from the war", labour and shuttle migrants) abroad increased by 2.8 million, with 65% of this increase in the first two weeks.

6. Estimation of the number of forced external migrants from Ukraine

A number of information sources provide higher data on the number of the "war refugees" from Ukraine than the State Border Guard Service. In particular, according to the Office of the United Nations High Commissioner for Refugees, as of June 7, 2022, 7.3 million people left Ukraine, of which just over 1.1 million were deported and left for Russia and Belarus and about 6.2 million left across the western borders with the EU and Moldova (UNHCR, 2022a). At the same time, more than 2.3 million people entered Ukraine during the same period.

In assessing the scale of the military wave of Ukrainian emigration, the data of the State Border Guard Service of Ukraine will be used as reference. It should be noted that almost one-way departure over time has changed to a combination of different types of movements, which were mentioned above (temporary return to see a husband or other relatives, import of a car under the procedure of preferential customs clearance with prior departure, departure to withdraw funds from a card abroad, temporary return of external labour migrants to visit relatives, celebration of Easter, etc.).

According to Ukrainian border guards, the scale of departures from Ukraine is smaller than assessments of UNHCR, with which scientists of the EU countries agree (Mercator Dialogue on Asylum and Migration, 2022). Identifying the reasons for the differences between these data sources requires a detailed study of the procedures for the formation of data sources of information from outside Ukraine, in particular through expert interviews with representatives of the headquarters of international organizations and authorities of the recipient countries. The most obvious reasons for the difference (illegal departure from Ukraine of male citizens and arrival of Ukrainian citizens from the territory of the aggressor countries), cannot explain its size. Solving this problem requires a special study with the involvement of specialists from both Ukraine and the recipient countries and may be a topic or component of an international scientific project.

The first date after which a significant convergence of the number of border crossings at the exit and entry began is March 27, 2022 (the difference between those who left and those who arrived on this day was 19 thousand persons, while on the previous day – 22 thousand persons, the next day – 12 thousand persons, and after that this number did not exceed 15 thousand persons). This happened on the eve of the withdrawal of Russian troops from the Kyiv, Chernihiv and Sumy regions of Ukraine. Accordingly, we will assume March 27 as the date of transition from the dominance of forced departure to the formation of a system of multidirectional and multi-purpose flows. Then, the total number of forced external migrants left for the EU and Moldova was 2.37 million. According to UNHCR, just over 1 million people left or were deported to the aggressor countries, for a total of 3.4 million. At the same time, as it has been already mentioned, at least 200 thousand persons can be considered returnees to Ukraine. Then, the number of forced external migrants from the country is 3.2 million persons, i.e. (taking into account possible estimation errors) about 3.0 million persons as of the beginning of June 2022.

7. The situation of Ukrainians abroad and prospects for return

The prospects for return of the Ukrainian “refugees from the war” from abroad are determined by a number of unpredictable factors, such as: the further course of the hostilities (the duration of the active phase and their spread throughout the country, with time having a greater impact on return decisions and geography on decision to depart); the general economic situation in Ukraine, and the quickness of restoration of the infrastructure and housing after the end of the war (employment opportunities, restoration of infrastructure, especially housing, educational, medical); maintaining ties with relatives in Ukraine, recipient countries' policy towards Ukrainians (opportunities for adaptation of forced migrants from Ukraine: employment, access to educational and medical services, social support, etc., in particular), which is largely determined by the attitude of the local population (Libanova, E. M., Pozniak, O. V., & Tsymbal, O. I., 2022).

A significant part of the forced migrants associate their return to Ukraine with the possibility of pre-accumulation of at least minimal resources for the restoration of destroyed housing and lost property, in order to avoid the need to ask for help.

Those are not going to return in the near future and are wary of such opportunities, who have already left their homes for the second time since 2014.

According to a survey conducted in early March by the authoritative sociological service of France in the three largest EU countries (Germany, France, Italy) and Poland, the fifth largest but geographically closest to Ukraine, the population of these countries show fantastically high solidarity with Ukrainians. Currently, a total of 87% of people agree, for example, with the reception and support of refugees (92% of Poles, 90% of Germans and Italians, 80% of French). For comparison: in 2015 (after the annexation of the Crimea and the Russian invasion of Donetsk and Luhansk regions) 46% of French, 77% of Italians and 79% of Germans supported the reception of refugees from Ukraine. Such support is unlikely to last very long, it will decline over time (probably, depending on the duration of the war as the war progresses), but it is currently extremely high.

Of course, the authorities of the recipient countries have to respond to this situation. That is why 10 EU countries (Austria, Belgium, Hungary, Germany, the Netherlands, Poland, Slovakia, Finland, France, the Czech Republic) and Georgia have introduced, for example, free rail travel for Ukrainians (Today, 2022). Ukraine's western neighbours and Ireland have allowed Ukrainians to enter without a passport (The Browser, 2022a). The EU has activated its Temporary Protection Directive, which may be extended for up to three years (Mercator Dialogue on Asylum and Migration, 2022). Additional benefits and simplifications of procedures are provided by the legislation of most countries (Depo, 2022).

Examples of legislative norms governing the stay of citizens of Ukraine. *According to the law regulating the terms of stay of the Ukrainian refugees in the country and their social support, in Poland (March 13) for all citizens of Ukraine and members of their families without Ukrainian citizenship, including those born in Poland, who entered Poland after February 24, the legality of staying in Poland for 18 months is guaranteed, with the possibility of submitting an application for temporary residence for a period of 3 years after 9 months of stay in the country; students and schoolchildren can continue their studies in Polish educational institutions; all Ukrainian refugees receive a work* *permit and access to medical care, a one-time cash benefit of 300 zlotys (approximately 63 euros) per person. The document also provides benefits for the Poles who will provide food and accommodation for citizens of Ukraine (Public, 2022).*

In the Czech Republic, from March 21, 2022, a special law regulates the provision of protection and visas (Ukrainians who arrive in the Czech Republic fleeing from Russian aggression, receive special visas to live in the country), employment (without special permission) (Ukrainian news, 2022), social security, access to health care and education, the ability to quickly get free housing (those who take in refugees from Ukraine, receive 120 euros per person per month); needy migrants can also count on a cash benefit of 200 euros per month (TSN, 2022).

A simplified temporary refugee program for refugees from Ukraine (CUAET) was launched in Canada on March 17. Ukrainian refugees and their next of kin of any nationality can stay in Canada as temporary residents for up to three years by applying online for a Canadian guest visa and presenting biometric data (fingerprints and photographs). Elementary and high school students have the opportunity to register and start attending school immediately upon arrival in Canada. Anyone can apply for permission to study at a higher education institution. To obtain a work permit for 3 years, Ukrainians have to submit an application at the same time as applying for a visa (The Browser, 2022b).

The UK Government has announced two schemes to support Ukrainian nationals to come in the UK: the Ukraine Family Scheme and the Homes for Ukraine Scheme. The Ukraine Family Scheme allows applicants to join their family members in the UK. To apply to the Ukraine Family Scheme, a person must be Ukrainian or the immediate family member of a Ukrainian national who is applying to the scheme, and have been residing in Ukraine on or immediately before 1 January 2022. The family member in the UK must be one of the following: a) a British national; b) someone settled in the UK – for example, they have indefinite leave to remain, settled status or proof of permanent residence; c) someone from the EU, Iceland, Liechtenstein, Norway or Switzerland who has pre-settled status and started living in the UK before 1 January 2021; d) someone with refugee status or humanitarian protection in the UK. The Homes for Ukraine Scheme will offer a route to those who have no family ties to the UK but instead have someone here willing to provide them with a home. It will enable individuals, charities, community groups and businesses to volunteer accommodation and provide a route to safety for Ukrainians, and their immediate family members, forced to escape their homeland. Those who come under this scheme will be able to live and work in the UK for up to three years, and access public services. Ukrainian nationals or the immediate family members of a Ukrainian national who were resident in Ukraine prior to 1 January 2022 are eligible for the scheme (UNHCR, 2022b).

The most popular status among Ukrainians abroad seems to be that of a temporary protection recipient in the EU, which allows them to visit Ukraine, change their place of residence, etc. It is the focus of almost 70% of polled migrants (4Service Group, 2022). This gives some indication of their intentions to return to Ukraine. But at the same time, this status provides more opportunities for the employment and obtaining of a new, more competitive profession, which will contribute to better prospects for adaptation abroad.

Assessing the prospects for return to Ukraine of those who left after February 24, 2022, we cannot ignore their division into two completely unequal groups: the majority are those who left Ukraine in search of safety, and quite insignificant number are those who decided to take the opportunity to move to a more developed, more prosperous country (most of those who ended up abroad immediately went to other countries). We should not try to bring back people from the second group, they have chosen just such a fate. Instead, of course, it is necessary to strive for the maximum return of mothers with children, and not to get a mass exodus of their husbands after the war. According to various (with an unclear degree of representativeness) estimates, 60 to 90% of migrants seek to return to Ukraine. Thus, according to 4Service (the survey was conducted on March 28 – April 04 and covered 3,027 Ukrainians abroad), 89% of refugees plan to return to Ukraine, but 61% of them believe they may stay abroad if the war lasts long; 67% are going to stay abroad until the end of the war, and only 18% expect the war to continue for several more months. 57% of migrants named safety as a condition for return, while only 9% and 5%, respectively, chose critical importance of employment and housing (4Service Group, 2022).

But we should take into account that such intentions are expressed by people in a state of stress and on an incredible wave of patriotic unity. Refugees, in such circumstances, often cannot admit even to themselves other aspirations, let alone the announcement of such intentions. This, in fact, is a manifestation of the artificial shift of answers in the "pro-Ukrainian direction", which is almost inevitable in modern conditions. Moreover, the mood for return/non-return may change over time, as the longer the war lasts, the more Ukrainian "refugees" will integrate into host communities. The logic of this process suggests that the highest risks of non-return are inherent in, on the one hand, the most vulnerable groups (single mothers with young children, large families, people with special needs) who can count on receiving significant refugee status in EU countries; on the other hand, in the most active and qualified people who are able to quickly find a job and adapt to new conditions. The second group is quite numerous. Thus, according to the data of the Center for Eastern European Studies at the University of Warsaw, almost two thirds of adult refugees from the war have higher education and professional experience, which is in short supply in Poland – in particular, specialists in education, medicine, industrial construction, IT (GALINFO, 2022).

It is obvious that both contingents can be replenished in the future by men of working age, who will reunite with the family as soon as possible. Developed countries need an influx of labour due to the aging and reducing of their permanent population, and the current wage gap combined with destroyed housing and/or lack of decent employment prospects will inevitably create similar risks. This is already openly said by senior IMF officials, assessing the economic losses and prospects of the countries that are actively accepting Ukrainians.

According to a survey conducted by the Kyiv International Institute of Sociology, only 42.9% of Ukrainians plan the future of their children and grandchildren in Ukraine if the hot phase of the war continues. In the event of a truce and the postponement of the war, 54.7% of the Ukrainians plan for the future of their children and grandchildren in Ukraine. The conditions under which more than three quarters of respondents agree to plan the future of their children and grandchildren in Ukraine are as follows: change of power in Russia to democratic (76.8%), Ukraine's accession to NATO or obtaining stronger security guarantees (85.1%), capitulation and disarmament of Russia (84.5%). On the other hand, less than 2% of Ukrainians plan to leave Ukraine for children/grandchildren in any case (Kyiv International Institute of Sociology, 2022b).

Summarizing data from a number of sociological surveys and expert assessments shows that even under the best conditions, at least 20% of forced migrants will not return to Ukraine. If we assume that 3 million people (see above) is the fixed number for the end of June of Ukrainians who went abroad in search of safe environment and did not return, under conditions of rapid de-escalation and cessation of the war (at least its active phase), irreversible migration losses are likely to be 600–700 thousand people. If the number of those forced to leave Ukraine increases to 7–8 million, as expected in the EU and international organizations (European truth, 2022), i.e. if the fighting continues for several more months and spreads throughout Ukraine, more than half of forced migrants will remain abroad, in addition, the husbands of irreversible migrant women will be reunited with their families after the end of martial law. Then, in the long run irreversible migration losses will inevitably increase to 5.0–5.5 million persons.

8. Discussion and conclusions

Russia's armed aggression against Ukraine has led to a catastrophic increase in external migration and to a number of migration problems, such as: the threat of non-return of a large number of people who went abroad, and the risks of departure of the men whose families are abroad after the lifting of the current ban (at the end of the war or its active phase), which will increase the share of non-returnees.

According to estimates by the Ptukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine, based on the data from the State Border Guard Service, as of June 1 2022, the number of “refugees from the war” from Ukraine is about 3 million persons. Possible amount of irreversible migration losses ranges from 600–700 thousand to 5.0–5.5 million persons, in accordance with the military and economic factors. This threatens a demographic catastrophe for the country the demographic potential of which has been exhausted. The total amount of migration losses (including external labour migrants) can reach more than 20% of the population, and the losses of children, women of childbearing age and highly educated persons may be even more intense.

Considering the development of demographic processes, promoting the maximum return of Ukrainians from abroad is one of the main tasks of the post-war strategy of Ukraine's development. All available levers of economic, cultural and social nature must be engaged to solve it.

The Ukrainian government must keep in touch with our compatriots, provide opportunities for children (and not only) to learn Ukrainian language and history, establish online learning of other disciplines in Ukrainian; demonstrate all possible manifestations of the fact that neither the Ukrainian state, nor the Ukrainian authorities have forgotten about them, that even if their homes are destroyed, they will be rebuilt after their return (or even earlier), and the state will take efforts to make wages fairer and more comparable to standards, such as in Poland and the Czech Republic; help in keeping social and employment links with the previous place of work, for example, through remote employment and/or long vacations. An important area of action can be online adult learning for competitive professions, such as IT-related, which can be mastered remotely and which will be in demand in the post-war Ukrainian economy. This training should be conducted by Ukrainian specialists, which will have an additional effect: the preservation of jobs in Ukraine during the war.

Considering that by 2022 3–3.5 million Ukrainians were working abroad, which means that now there are more than 6 million of them there, the issue of the post-war return of our compatriots, a quarter of whom are children, is perhaps the most important in the context of preserving the Ukrainian nation and national security. Preservation of Ukrainian roots, as evidenced by the mass return of male migrant workers who have come back to defend Ukraine, plays a key role in the return decision.

Of course, we will need not only patriotic but also economic motivation, such as: the opportunity to receive a reasonable salary in Ukraine, more rapid economic growth, democratic development, combating legal nihilism and corruption, which has nearly choked us. It seems that many potential and actual migrant workers will be encouraged either to return, or to refuse to leave. As for the "refugees from the war", economic incentives are likely to be important and effective for them, because it will not be easy for everyone to find a suitable job abroad, especially with small children, and financial assistance is nowhere and never sufficient enough to maintain average living standards.

It is also important to strengthen cooperation with the authorities of the recipient countries, in particular in matters of psychological support for forced migrants, their identification, and in the future – in the development and implementation of return programs.

References

4service Group, (2022). Ukrainian Refugees in Europe. A Study of the Behavior and Attitudes of Ukrainians Who Were Forced to Leave for Europe as a Result of the War in Ukraine, https://4service.group/ukrayinski-bizhentsi-v-yevropi-doslidzhennya-povedinki-ta-nastroyiv-ukrayintsiv-shho-vimusheni-buli-viyihali-v-yevropu-vnaslidok-vijni-v-ukrayini/ [in Ukrainian].

Depo, (2022). The Evacuees Were Told about the Rules of Stay in Other Countries, https://news.depo.ua/ukr/news/evakuyovanim-ukraintsyam-rozpovili-shcho-potribno-znati-pro-pravila-perebuvannya-v-inshikh-krainakh-202203101433083 [in Ukrainian].

European Truth, (2022). Germany Expects 8 Million Ukrainian Refugees in the EU. https://www.eurointegration.com.ua/news/2022/03/21/7136341/ [in Ukrainian].

Galinfo, (2022). The Center for Eastern European Studies Has Painted a Portrait of a Ukrainian Refugee in Poland, https://galinfo.com.ua/news/tsentr_shidnoievropeyskyh_studiy_sklav_portret_ukrainskogo_bizhentsya_v_polshchi_385102.html [in Ukrainian].

Gradus Research Company, (2022). Migration and Socio-Political Attitudes during Russia's Full-Scale War Against Ukraine – The First Wave of Research, https://gradus.app/documents/256/GradusEU_survey_wave1_UA_hSorts5.pdf.

Ilko Kucheriv Democratic Initiatives Foundation, (2022). Assessing the Needs and Expectations of the Future of Internally Displaced Persons and Refugees - Analytical Report №1, https://dif.org.ua/uploads/pdf/1108543395628f2db973e3e0.29887914.pdf [in Ukrainian].

Kyiv International Institute of Sociology, (2022a). Perception index of the Russian-Ukrainian war: results of a telephone survey conducted on May 19–24, 2022, https://www.kiis.com.ua/?lang=eng&cat=reports&id=1113&page=1

Kyiv International Institute of Sociology, (2022b). Only Half of Ukrainians Plan for the Future of Their Children and Grandchildren in Ukraine if the military threat persists, https://www.kiis.com.ua/?lang=ukr&cat=reports&id=1118&page=1 [in Ukrainian].

Libanova, E. M., Pozniak, O. V., (2020). External Labor Migration from Ukraine: the Impact of COVID-19. *Demography and Social Economy*, 4(42), pp. 25–40, https://doi.org/10.15407/dse2020.04.025 [in Ukrainian].

Libanova, E. M., Pozniak, O. V. and Tsymbal, O. I., (2022). Scale and Consequences of Forced Migration of the Population of Ukraine as a Result of Armed Aggression of the Russian Federation. *Demography and Social Economy*, 2(48).

Mercator Dialogue on Asylum and Migration, (2022). Refugee protection in the EU: Building resilience to geopolitical conflict. Kiel, Germany, https://www.ifw-kiel.de/fileadmin/Dateiverwaltung/MEDAM-Webseite/Publications/Assessment_Reports/2022_MEDAM_Assessment_Report/MEDAM_assessment_Report_2022.pdf

Ptoukha Institute for Demography and Social Studies of Nas of Ukraine, (2018). Ukrainian Society: the Migration Dimension. National report. Kyiv [in Ukrainian].

Public, (2022). The President of Poland Has Signed a Law on Ukrainian Refugees. What He Predicts, https://suspilne.media/216932-duda-pidpisav-zakon-sodo-perebuvanna-u-polsi-ukrainskih-bizenciv/ [in Ukrainian].

State Border Guard Service of Ukraine, (2022). Facebook, https://www.facebook.com/DPSUkraine [in Ukrainian].

The Browser, (2022a). 6 Countries Where Refugees from Ukraine Can Enter Without a Passport, https://news.obozrevatel.com/ukr/travel/journal/7-krain-kudi-mozhna-vihati-bizhentsyam-z-ukraini-bez-zakordonnogo-pasporta.htm [in Ukrainian].

The Browser, (2022b). Canada Has Launched a Simplified Temporary Stay Program for Refugees from Ukraine, https://news.obozrevatel.com/ukr/abroad/kanada-vidkrila-sproschenu-programu-timchasovogo-perebuvannya-dlya-bizhentsiv-z-ukraini.htm [in Ukrainian].

Today, (2022). It Is Enough to Show the Passport - in Which Countries Travel for Ukrainians is Free, https://economics.segodnya.ua/ua/economics/transport/dostatochno-pokazat-pasport-v-kakih-stranah-besplatnyy-proezd-dlya-ukraincev-1606532.html [in Ukrainian].

TSN, (2022). Wages in the Czech Republic in 2022 and the Average Cost of Living: Everything a Refugee from Ukraine Needs to Know, https://tsn.ua/ato/zarplati-vchehiyi-2022-roku-i-seredni-vitrati-na-zhittya-vse-scho-potribno-znati-bizhencyu-z-ukrayini-2014618.html [in Ukrainian].

Ukrainian News, (2022). Ukrainian Refugees Will Be Able to Work in the Czech Republic Without Obtaining a Special Permit, https://ukranews.com/news/842718-ukrainskie-bezhentsy-smogut-rabotat-v-chehii-bez-polucheniya-spetsialnogo-razresheniya [in Ukrainian].

UNHCR, (2022a). Ukraine Refuge Situation. The Operational Data Portal of UNHCR. https://data2.unhcr.org/en/situations/ukraine

UNHCR, (2022b). UNHCR UK Information and Links on Ukraine, https://help.unhcr.org/uk/unhcr-uk-information-and-links-regarding-ukraine

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 277–294, DOI 10.59170/stattrans-2023-015
Received – 01.07.2022; accepted – 20.11.2022

Using electronic registries to study the COVID-19 pandemic and its consequences

Maryna Puhachova[1], **Oleksandr Gladun**[2]

ABSTRACT

The article analyses systems of electronic information resources (registers and databases) in the field of the healthcare in different countries. Not only do such systems provide information to support the treatment of patients, but they also accumulate large amounts of statistics, thus enabling their qualitative operational analysis. Electronic health systems are particularly popular in the United States, and as regards Europe, in Denmark and Norway. These countries created good conditions for the efficient introduction of new electronic information resources that would store all the accumulated information about the COVID-19 pandemic. The aim of the article is to summarise information on the use of electronic registers and databases to create an information base for the study of the COVID-19 pandemic and its consequences in different countries, and to formulate proposals for the improvement of electronic health systems in Ukraine. On the basis of the results of their research, the authors propose a list of electronic registers that can significantly improve the analysis of both the course and the consequences of the coronavirus disease. The list includes: the register of immunisations, a specialised register of complications after the recovery from COVID-19, a register of tests for the presence of this disease or the recovery, etc., as well as a register of hospitalisation cases.

Key words: COVID-19 pandemic, system of electronic information resources, registers and databases of the healthcare system, electronic health records.

1. Introduction

The large amount of information that accumulates in health care systems needs to be systematized and provided with easy access to specialists for further use in medical practice. Such an opportunity is provided by electronic registers and databases of

1Ptoukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine, Kyiv, Ukraine. E-mail: maryni@ukr.net. ORCID: http://orcid.org/0000-0003-1680-4284.

2 Ptoukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine, Kyiv, Ukraine. E-mail: gladun.ua@gmail.com. ORCID: http://orcid.org/0000-0001-6354-6180.

© Maryna Puhachova, Oleksandr Gladun. Article available under the CC BY-SA 4.0 licence

medical systems, which store huge amounts of information about patients. Actually, the so-called electronic health records (EHRs) are created to store and use this information, which allows information about each person from different registries to be collected together. The spread of the new virus has necessitated not only the processing of increased amounts of information to be used to treat patients (disease follow-up, use of certain drugs, vaccination, death registration), but also the creation of new and more active use of existing electronic information resources outside the purview of health care. These are, for example, systems for registration of businesses and vehicles, educational systems, information resources of pension and social funds, in which individuals and legal entities will be able to register and receive information remotely (Kayumova and Meunier, 2021). Thus, the pandemic, that is currently raging in the world, has accelerated the digitalisation of not only in the field of health care but also in other areas of society.

The aim of the article is to summarize information on the use of electronic registers and databases to create an information base for the study of the COVID-19 pandemic and its consequences in different countries and to develop proposals for its improvement in Ukraine.

The research methodology is based on the application of the following methods: system analysis – to determine the role and relationship of registers in the general health care system of each country and Ukraine in particular; scientific generalization and comparative analysis, which made it possible to group the studied registers and databases on the basis of common features and to identify similar international and national registration systems; research methods of complex systems (analytical and synthetic) – for the analysis of systems containing information related to the pandemic, in individual countries and international organizations, in terms of both the functional composition of these systems and their sufficient unification.

2. Electronic health information resources during the pandemic

As mentioned earlier (Gladun et al., 2021; Puhachova and Gladun, 2021), developed countries have formed systems of electronic information resources that cover various areas of state functioning and provide governments and inhabitants with the necessary information. In these countries, health care systems are the most advanced. At present, during the pandemic, they have become even more important, as they not only quickly provide the necessary information, which is constantly accumulating in large quantities, but also provide assistance in treating patients with COVID-19, and often allow tracking disease chains (virus carriers, who directly transmit it to others) for rapid response and interruption of these chains.

Health care systems usually have a number of specialized registers and databases that contain information about patients, diseases, vaccinations, medicines, medical staff and medical equipment, etc. Electronic medical records are the most convenient and widespread tool for registering information about patients in such systems, which is used in many developed countries, and in recent years – in Ukraine. This is especially important during the pandemic. EHRs provide an opportunity not only to obtain generalized information about each patient, but also to build scenario forecasts for the country / region on the growth or decline (the so-called waves) in morbidity, the course of the disease in different patient groups, complications, mortality based on information collected in the registers.

Particular researchers (Sudat et al., 2021; Madhavan et al., 2021; Satterfield et al., 2021), for example, described some problems with the use of electronic card information to study the spread of coronavirus in different US states and highlighted certain shortcomings inherent in EHR data. It is emphasized that the quality of data on the disease and the early deadlines for their registration significantly increase the effectiveness of crisis response. Data quality is critical to the success of the analysis (from the simplest to the most complex), and in the context of COVID-19, large-scale data sharing in the United States and around the world has played an important role in how medical systems responded to pandemics. It is the availability of electronic information resources that allows to quickly register information about diseases and treatments and share the results. The authors emphasize the usefulness of electronic registers, which store information about patients that can be used and scaled to improve clinical care, research and decision-making in health care.

Undoubtedly, the worst consequence of coronavirus disease is death, and the availability of as much information as possible about the course of the disease and the various factors that affect it (comorbidities, treatment, sex, age, etc.) helps to better predict treatment outcomes. And it is the availability of the necessary information in electronic registers which allows one to create a quality information base for the study, and the ability and ability to process this information are the main prerequisites for obtaining a reliable forecast. This is stated in the article (Schwab et al., 2021), the authors of which argue that the use of clinical prognostic models based on EHR information can help make better decisions about the treatment of patients. Researchers have introduced an early warning system for the incidence of COVID-19, which allows the risk of death associated with the disease to be assessed in advance. Thus, models are able to predict in advance which patients are at higher risk of dying, and can help provide appropriate resource priority for these individuals. In addition, as more observational data are collected, these models can be used to identify both new risk factors and interactions between existing factors, providing better understanding and opportunities for appropriate intervention.

The authors of the article (Estiri et al., 2021) argue that the ability to operate with large amounts of information allowed not only to predict death from COVID-19, but also to understand differences in risk factors (such as diabetes, cancer, lung disease, smoking) for different age groups. And the ability to calculate accurate risk assessments at the individual level solely on the basis of EHR is crucial for qualitative forecasts and efficient allocation of resources, according to the researchers.

Moreover, it is known that the National COVID Cohort Collaborative database has already been created in the United States, which accumulates information on coronavirus incidence and its consequences. Articles have been published (e.g. Brown et al., 2021) in which this database is used as an information platform for in-depth research and forecasting. In this paper, the authors presented a multifactor model of logistic regression, which revealed links between age, male gender, liver disease, dementia, African or Asian descent, obesity and severe consequences of COVID-19.

A secure OpenSAFELY analytical platform has been developed to provide detailed primary care records for a significant proportion of patients in the country, and to study COVID-19 morbidity and mortality in the UK's National Health Service (Williamson et al. 2020). On this platform, electronic patient medical records are pseudonymically linked to data from the COVID-19 confirmed inpatient death notification system. It is known that in early 2020, this ability to combine information from different sources made it possible to study an array of data on almost 17.5 million adults in the country. Thus, with the further accumulation of data from various sources, the platform provided an opportunity to create a large information base for the study of the pandemic and its consequences.

Some researchers point out that people with mental disorders may be at increased risk for COVID-19 infection and more likely to have adverse effects. For example, a nationwide EHR database of 61 million adult patients from 360 hospitals and 317,000 health professionals in 50 U.S. states by July 29, 2020 was analysed (Wang et al., 2021). As a result, people with a newly diagnosed mental disorder were found to be at increased risk of infection (which is even more prevalent among Afro-Americans and women) and have higher rates of some adverse effects of infection.

As far back as October 2019, all American countries approved the Plan of Action for the Strengthening of Information Systems for Health, which is based on the joint efforts of the Pan American Health Organization and the Inter-American Development Bank and promotes adoption of international standards to ensure the exchange of information, knowledge and data (Pan American Health Organization, 2021). In other words, American countries have made efforts to improve their medical information systems and, if possible, to coordinate their work by expanding the opportunities for information exchange. This decision, in our opinion, was very timely given the pandemic covered the world's population a few months later.

Moreover, information from electronic health registries is used to track the spread of the virus involving geolocation. Thus, according to Singaporean researchers, the combination of data from electronic medical records and maps helped to track contacts not only within local communities, but also among physicians who became infected both in the workplace and in their communities. Based on this information base, a system of personnel surveillance was established to detect clusters of acute respiratory infections among health professionals, which allowed to quickly break the chains of virus spread (Sim et al., 2021).

With the pandemic in the background, a major problem in maintaining electronic information resources in working order was their overburden. Researchers and physicians in many countries report that not only elements of health care systems (hospitals, infirmaries, ambulance crews, medical staff, etc.) are overloaded during peak morbidity, but also electronic registers of these systems. This undoubtedly affects the quality of information registration and, consequently, the maintenance of an appropriate disease information base (Pan American Health Organization, 2021; Ashley, 2020; Slone, 2020; AbouZahr et al., 2021). In addition, other systems more or less health-related are becoming more overloaded during a pandemic, i.e. those directly affected by government pandemic response and recovery measures – the social sphere (unemployment, patient financial support during the illness, accounting and reimbursement of medicines, death, birth, etc.), loans for small business, statistical system and systems of other spheres of state functioning (Kayumova and Meunier, 2021). All relevant information should be collected, protected and stored to ensure the provision of basic administrative services. Failures in the US electronic health system have made it difficult to get vaccinated in the early stages of vaccination (Popperl et al., 2021). In addition to the system registers, which contain EHR of patients with information about routine examinations, diagnoses, prescriptions, tests, vaccinations, etc., the country created registers that should contain information about the incidence of COVID-19 and vaccination against it. This has greatly complicated the process of immunization, monitoring side effects and ensuring a fair and efficient distribution of vaccines.

Another problem that arises when entering data into electronic resources and their subsequent proper use is the lack of a patient ID, according to some physicians and IT professionals. Related to this is the ease of use of electronic patient records and their safety. In other words, the presence of a unique identifier from a technical point of view allows the following:

- avoid duplication of registry entries for the same patient, and therefore not artificially increase the amount of information and time to search for it;
- minimize the possibility of errors in information about the health of patients;
- facilitate the exchange of information between different medical institutions.

However, even in the US health care system, such identifiers are not yet used (Slone, 2020).

Mandatory electronic case reporting (eCR) has been introduced in the United States as part of the Centers for Disease Control and Prevention since January 1, 2022 for all hospitals and infirmaries in the country to obtain information on COVID-19 (Centers for Disease Control and Prevention, 2021). This will allow to automatically create and transmit disease reports from an electronic medical record to health authorities for consideration and action. The system can quickly, securely and seamlessly transfer EHR data in health care facilities to state, local and territorial health care facilities and provide information to health care professionals. Therefore, all information on the incidence of coronavirus in the country should be collected together and can be processed quickly.

In some countries, electronic registers are also used to register persons when accessing premises. For example, the New South Wales Government (Australia) has developed an electronic registration tool called COVID Safe Check-in. Businesses and organizations required by the relevant health care order must register as COVID Safe and keep electronic records of the people who enter their premises. Such records make it possible to track vital contacts when someone is diagnosed with COVID-19 (NSW Government, 2022). When a company or organization registers as COVID Safe, it is provided with a unique QR code to reflect this action. When visiting the premises, customers scan the QR code with their phone's camera and register either with a special Service NSW application or by entering their data in the appropriate web form. After that, persons who visited the premises will be recorded, and the special medical service will be able to contact these people if necessary.

A similar approach has been introduced at the University of Greifswald (Mecklenburg-Vorpommern, Germany), where the central system of the University Computer Center collects and stores contact details of all those who visit its libraries and participate in all kinds of activities and classes (e.g. seminars, lectures, sessions of university bodies, etc.). This information may, if necessary, be passed on to the Department of Health (University Greifswald, University Computer Center, 2022).

Similar systems have been developed and are operating in Japan, South Korea and some other countries. Thus, recording such information allows, on the one hand, to monitor and break the chains of virus spread, and on the other hand, to accumulate a variety of information (on the possibility of detecting the disease at an early stage, asymptomatic carriers, etc.). Clearly, these registers are an additional contribution to the creation of an information base for the COVID-19 study.

It is worth mentioning electronic information resources that store data about clinical trials on coronavirus and the publications of scientists and researchers dedicated to them. These registers and databases have been created both by international organizations and in individual countries. For purposes of illustration the following resources (European Clinical Research Infrastructure Network, 2022; European Network of Centers for Pharmacoepidemiology and Pharmacovigilance, 2021; Stanley et al, 2021) can be presented:

1) WHO International Clinical Trials Registry Platform (ICTRP);
2) Database of clinical trials conducted worldwide and funded by private and public funds (ClinicalTrials.gov, National Library of Medicine);
3) EU Clinical Trials Register;
4) Cochrane Central Register of Controlled Trials (CENTRAL). Cochrane is an international network headquartered in the United Kingdom, a registered non-profit organization and a member of the National Council of Voluntary Organizations of the United Kingdom;
5) Pan African Clinical Trials Registry (PACTR);
6) The EU Electronic Register of Post-Authorization Studies (EU PAS Register);
7) Register of Information Storage Information on ISRCTN (International Standard Randomized Controlled Trial Number) of the Port of Springer Nature;
8) European Network of Centers for Pharmacoepidemiology and Pharmacovigilance (ENCePP);
9) COVID-19 Research Database – Platform C19.

The European Clinical Research Infrastructure Network (ECRIN) website also provides detailed information on national electronic information resources used by countries on all continents to register information related to the incidence of COVID-19 (European Clinical Research Infrastructure Network, 2022).

3. Registers and databases in the medical systems of Nordic countries

As we mentioned earlier, in Europe, special attention is paid to the creation and use of electronic health records in Nordic countries, with Norway and Denmark having the most extensive medical information systems (Puhachova, 2021). According to our information, the health care system of Denmark uses 28 electronic resources, Norway – 21, other northern countries – within 10. Table 1 (compiled by the authors on the basis of information (eSundhed, 2021; The National Center or Register Based Research, 2021; Norwegian Institute of Public Health, 2021; Ministerie van Volksgezondheid, Welzijn en Sport, 2021; TEHIK, 2021; Socialstyrelsen, 2021; Finnish Institute for Health and Welfare, 2021) shows, as the example, electronic registers / databases that may be directly related to the incidence of COVID-19 and possible consequences of the disease. It is clear that these registers are not a complete list of all medical registers available in countries – the virus causes complications to various body systems. In addition, for example, the Births Register and the In Vitro Fertilization Registry (introduced in Denmark) could also be classified as disease-related due to the decline in fertility caused by the pandemic. Yet, in Table 1 it can be seen that if patient registers are available in all northern European countries, then such important systems that contain registers of surveillance of infectious and viral diseases and registers for monitoring the resistance to antimicrobial and antiviral drugs, are created only in Norway, according to our information. This also applies to the immunization registers and the rehabilitation register.

Table 1. Registers of health care systems in Nordic countries

№	Common name of registers in countries	Country					
		Denmark	**Netherlands**	**Norway**	**Estonia**	**Sweden**	**Finland**
1.	Patient register	The National Patient Register, Register of Preventive Measures	The National Register of Hospital Care	Norwegian Patient Register	Estonian National Health Information System, Estonian Health Statistics Database	National Patient Register	Care Register for Health Care, Register of Primary Health Care visits
2.	Register of cardiovascular diseases			Norwegian Cardiovascular Disease Registry		Myocardial Infarction Database	
3.	Register of abortions	The Register of Legally Induced Abortions		Registry of Pregnancy Termination	Estonian Abortion Registry	Register of Abortions	
4.	Birth register	The Medical Birth Register, The Medical Register of Births and Deaths		Medical Birth Registry of Norway	The Estonian Medical Birth Registry	Swedish Medical Birth Register	Medical Birth Register
5.	Registers of surveillance of infectious and viral diseases			Register of Norwegian Surveillance System for Communicable Diseases, Norwegian Intensive Care and Pandemic Registry,			

				Register of the Norwegian System for Monitoring the Use of Antibiotics and Related Infections, COVID-19 Emergency Preparedness Register			
6.	Registers for monitoring the resistance to antimicrobial and antiviral drugs			Register of Norwegian Surveillance System for Antimicrobi al Drug Resistance, Norwegian Surveillance System for Antiviral Resistance			
7.	Immunization registry			Norwegian Immunisati on Registry			
8.	Register of causes of death	The Register of Causes of Death, The Mortality and Occupation Database			Estonian Causes of Death Registry	Causes of Death Register	Causes of Death Register
9.	Register of lung diseases	The Danish Lung Cancer Register			The Estonian Tuberculosis Registry		
10.	Register of diagnoses established in hospitals	The Register of Hospital Discharges				BD In-patient care diagnosis	
11.	Register of patients undergoing rehabilitation			Rehabili-tation Register			
12.	Hospital Discharge Register	The National Hospital Discharge Register					

13.	Register of recipes			Norwegian Recipe Database		National Prescribed Drug Register	
14.	Health insurance register	The Public Health Insurance Register					
15.	Register of registered specialists	The Register of Professional Mobility, The Authorization Register	The National Register of Medical Workers, Dutch Register of Physicians, Register of social hygiene	Register of Medical Employees		Statistical database of health professionals	
16.	Municipal health services register	The Register of Municipal Health Services					
17.	Hospital utilization register	The Hospital Utilisation Register					
18.	Register of state reimbursements for medical care			Database on Control and Payment of Medical Care Compensation			
19.	Register of Medicines and Drugs Statistics	The Register of Medicines and Drugs Statistics	The Register of Medicines and Drug Statistics				
20.	Hospital Errors Register				Hospital Errors Register		
21.	Register of medical services of the armed forces			Registry of the Norwegian Armed Forces Medical Services			

In contrast to the growing pandemic in some countries, specialized registers have been created, for example, the Norwegian Institute of Public Health (NIPH) introduced a COVID-19 Emergency Preparedness Register in 2020 (Beredt C19) (Lindman, 2021). NIPH emphasizes that national surveillance systems and health registries established before the epidemic gather information to 287ravel the purpose of each such registry. However, to get a quick and reliable overview of the situation related to COVID-19, you need to have all the necessary information, so it is better to refer to one register, rather than “extract” and compile data from different sources during research, which significantly slows down the analysis flow. Information on the Beredt C19 is collected from the Norwegian Surveillance System for Communicable Diseases (it provides daily information on confirmed COVID-19 cases and age, sex, site of infection, country of birth), the Norwegian Patient Registry and the Norwegian Intensive Care and Pandemic Registry (see Table 1); at the same time it is possible that in the future, if necessary, new sources of information may be identified. Thus, Beredt C19 provides an operational analysis of how the pandemic and the measures taken affect the health of the population, as well as how medical services are provided.

Quite a large amount of data to be stored on electronic resources relates to information about border crossings by domestic citizens and foreigners. That is, in the specialized registers of border services, in addition to the usual information about travellers, data on the availability of vaccinations against COVID-19, vaccine brands, test results, etc., should be stored. Thus, the coronavirus pandemic has prompted not only health systems around the world, but also other areas of government to introduce new electronic information resources and improve those used previously. The efficiency and comprehensiveness of using diverse information in North European countries in the COVID-19 context is determined by the system approach to creating registers and the availability of a unique identifier of a person for all the registers.

4. The system of electronic medical information resources in Ukraine and its improvement

In Ukraine, state electronic information resources are currently created and operate both in the field of health care and in other areas: social, financial, tax, etc.

An Electronic Health Care System in Ukraine (eHealth) was launched in 2017 in trial mode and in 2018 in operative mode (Electronic health care system in Ukraine, 2022). Its operation is maintained by the National Health Service of Ukraine. It is a sophisticated and hierarchical system with the two components: Central Database (CDB) and Medical Information Systems (MIS). The interactions of users (medical employees, pharmacists, and patients) with CDB are carried out via MIS.

CDB is an information and telecommunication system containing legally established registers, program modules, an information subsystem required for the implementation of government financial guarantees, and other components.

Presently, the following registers are being operated in CDB:

- Register of patients;
- Register of declarations on the choice of a doctor who provides primary medical care;

- Register of business entities in the health protection sector;
- Register of medical specialists;
- Register of medical workers;
- Register of contracts on medical service of the population;
- Register of contracts on reimbursements;
- Register of medical records, records on referrals and prescriptions;
- Register of medical conclusions.

The functions of CDB enable creation, review, information and documentation exchange between medical registers, MIS and government electronic information resources of other systems.

By now, the following MIS modules have been connected to CDB:

- Administrative module of a provider of medical services of the primary medical aid;
- Workplace of a doctor of the primary medical aid;
- Administrative module of a drugstore;
- Workplace of a pharmacist;
- Administrative module of a provider of medical services of the specialized medical aid;
- Workplace of a doctor of the specialized medical aid;
- Working with patient records;
- Data access.

Hence, MIS enabling to computerize the operation of business entities in the field of health care sector supports creation, updating, review and exchange of information in electronic form, with CDB in particular.

Presently, eHealth contains 27.7 million registered patients, with entering data on individuals about the fact of vaccination from the fixed nomenclature of deceases, which lays the legal ground for issuance of respective certificates.

A problem that is faced by Ukraine and needs to be dealt with in creating new registers is the availability of a person unique identifier for all the electronic information resources. Another serious problem in Ukraine is the completeness of data filling and the consistency of data from various registers. Thus, the Unified State Demographic Register of Ukraine contains data only about two thirds of the Ukrainian population.

It is obvious that the pandemic of COVID-19 has created new challenges for the health protection system on the whole and for operating of its electronic information resources in particular. Like in Western countries, they are provoked in Ukraine by the occurrence of the system overloading, lack of personnel with good training, large numbers of ill persons among both medical workers and population, etc.

Another problem imposing heavy limitations on applications of registers for analysing COVID-19 morbidity by territorial location (district, residential area) is that the actual residence in Ukraine quite often differs from the registered one. The possibility of using data from mobile network operators for geolocating purposes calls for solution of a series of normative and legal and methodological issues.

An intersectoral Working Group on mathematical modelling of problems related with the pandemic of SARS-CoV-2 in Ukraine was created in April 2020 to predict the development of the situation of COVID-19 diseases. The group was joined by the National Academy of Sciences of Ukraine, Kyiv Taras Shevchenko National University, Vinnytsia National Technical University, and the National Academy of Medical Sciences of Ukraine. One of the operative areas of the Working Group is to elaborate a projection of the evolution of COVID-19 pandemic. Its information base is data from the Center of Public Health at the Ministry of Health of Ukraine and the National Health Service of Ukraine. But this base cannot be referred to as a register.

The Center of Public Health at the Ministry of Health of Ukraine, with the sponsorship of the United Nations Children's Fund (UNICEF) and the U.S. Agency on International Development (USAID) and in the partnership with the international organization REACH, has created an information base from which analytical dashboards and online maps could be developed. They enable for a detailed statistical multifactor and geospatial analysis of COVID-19 morbidity and mortality, operation of hospitals and quarantine restrictions in schools at the level of country and individual settlements. But this information base is not a register in the strict sense, either. The public access to this resource was closed after the beginning of full-scale aggression of Russia Federation against Ukraine. Once a full-fledged register is available, it will enable to apply quasi-experimental approaches to investigating the efficiency of using various methods of medical treatment.

Hence, CDB contains information about coronavirus patients. But, as we think, both for this purpose, and for supervision of other diseases which are complications of COVID-19 or the most widespread in our country, it is possible to offer some more these resources (as it was proved earlier (Puhachova and Gladun, 2021; Puhachova, 2021), implemented in the Danish, Norwegian and Estonian systems:

1. Register of cardiovascular diseases – due to high mortality from them (it ranks first in the causes of death in Ukraine), including as a complication due to the incidence of coronavirus.
2. Register for monitoring the use of antibiotics and related infections. It is known that the population alone, without the advice of doctors, takes antibiotics. In view of this, in Ukraine it was planned to switch to full prescription release of these drugs in pharmacies. We believe that the availability of a register will help streamline the monitoring of antibiotic use and its consequences.

3. Register of surveillance for resistance to antiviral drugs – due to uncontrolled consumption of such drugs by the population, especially during the COVID-19 pandemic.
4. Immunization register – to register the received (including mandatory, planned, booster) vaccinations by the country's residents, as well as to track the volume of vaccine use and related financial costs.
5. Hospitals utilisation register, which is extremely relevant during epidemics.

To this list, in our opinion, the Specialized Register of Complications after Recovery from COVID-19 and the Register of Tests for the Presence of the Disease should be added (it should be noted that testing is conducted by both public and many private laboratories, so information on test results is scattered) or recovery after it, bearing in mind that sufficient statistics have already been accumulated on the disease and its consequences.

In addition, as mentioned above, it would be appropriate to supplement the electronic system of the State Migration Service with an additional specialized register (sub-register) containing information on individuals crossing the border related to the COVID-19 pandemic.

The creation of new registers will require additional resources, but it can allow to systematize the respective information and accelerate its access. It is important not only in view of measuring its social and economic impact, but also for gaining experiences in creating analogous registers in time of other epidemics or infectious diseases.

5. Conclusions

In more than two years of the COVID-19 pandemic, physicians and scientists have amassed vast amounts of data on morbidity, treatment, consequences, and mortality from the disease; there have also been many research results. Of course, all this information must be aggregated, sorted and be stored electronically with the ability to be quickly accessed by specialists (it is about impersonal data sets). This is exactly the work that is being done in the health systems of many developed countries. In Ukraine, thanks to the introduction of an Electronic Health Care System in 2018, there are electronic registers that contain a lot of information about patients, including those who have recovered (died) from coronavirus. But the vast amounts of data that will accumulate and continue to be collected until the end of this pandemic are prompting experts to consider whether the data are stored in a way that is convenient enough to use, including in medical statistics. In view of this, and taking into account the experience of developed countries, we have proposed the introduction of several new electronic registers, including those directly related to COVID-19 disease.

In our further research it is planned to analyse the impact of the COVID-19 epidemic on demographic processes in Ukraine on the basis of information from electronic information resources.

References

AbouZahr, C., Bratschi, M. W., Cercone, E., Mangharam, A., de Savigny, D. and Dincu, I. et al., (2021). The COVID-19 Pandemic: Effects on Civil Registration of Births and Deaths and on Availability and Utility of Vital Events Data. *American Journal of Public Health*. Retrieved from: https://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2021.306203.

ACCESS template protocol for safety of COVID-19 vaccines. (2021). *European Network of Centers for Pharmacoepidemiology and Pharmacovigilance*. Retrieved from: https://www.encepp.eu/encepp/viewResource.htm?id=39362.

Ashley, L., (2020). The tsunami of data and the hard pivot to remote work and home school has stretched government information management capabilities to their limits. *Nextgov*. Retrieved from: https://www.nextgov.com/ideas/2020/08/covid-19-threat-or-opportunity-electronic-government-records-capture-access-and-preservation/167432/.

Brown, J. S., Bastarache, L.andWeiner, M. G., (2021). Aggregating Electronic Health Record Data for COVID-19 Research – Caveat Emptor. *JAMA Network Open*, 4 (7):e2117175. doi: 10.1001/jamanetworkopen.2021.17175.

COVID-19 Electronic Case Reporting for Healthcare Providers. (2021). *Centers for Disease Control and Prevention*. Retrieved from: https://www.cdc.gov/coronavirus/2019-ncov/hcp/electronic-case-reporting.html.

COVID-19 Registration System. (2022). *University Greifswald. University Computer Centre*. Retrieved from: https://rz.uni-greifswald.de/en/services/study-teaching/elearning/covidregister/.

COVID Safe check-in. (2022). *NSW Government*. Retrieved from https://www.service.nsw.gov.au/transaction/covid-safe-check.

COVID-19 Trials Registries. (2022). *ECRIN – European Clinical Research Infrastructure Network*. Retrieved from: https://ecrin.org/covid-19-trials-registries.

Danish Registers, (2021). The National Centre for Register Based Research. *econ.au.dk.* Retrieved from: https://econ.au.dk/the-national-centre-for-register-based-research/danish-registers/.

Database reporting, (2021). Finnish Institute for Health and Welfare. *thl.fi*. Retrieved from https://thl.fi/en/web/thlfi-en/statistics-and-data/statistics-by-topic/database-reporting#Primary_care.

Electronic Health Records (EHR) and Interoperability: Understanding two key concepts for a better Public Health response, (2021). *Pan American Health Organization*. Retrieved from: https://iris.paho.org/handle/10665.2/52003.

Electronic health care system in Ukraine, (2022). *ehealth.gov.ua*. Retrieved from https://ehealth.gov.ua/.

Emergency preparedness during COVID-19, (2021). *Socialstyrelsen*. Retrieved from: https://www.socialstyrelsen.se/en/.

Estiri, H., Strasser, Z. H., Klann, J. G., Naseri, P., Wagholikar, K. B. and Murphy, Sh. N., (2021). Predicting COVID-19 mortality with electronic medical records. *NPJ Digital Medicine, 4, 15*. Retrieved from: https://doi.org/10.1038/s41746-021-00383-x

Gladun, O. M., Puhachova, M. V. and Vynogradova, M. V., (2021). Electronic registers: European experience of establishment and use. Kyiv: IDSS. Retrieved from: https://idss.org.ua/arhiv/registers.pdf [in Ukrainian].

Health Registers, (2021). *NIPH-Norwegian Institute of Public Health*. Retrieved from: https://www.fhi.no/en/hn/health-registries/

Kayumova, M. and Meunier, F., (2021). Digitizing business registries. *World Bank blogs*. Retrieved from: https://blogs.worldbank.org/developmenttalk/digitizing-business-registries.

Lindman, A. E. S., (2021). Emergency preparedness register for COVID-19 (Beredt C19). *Norwegian Institute of Public Health*. Retrieved from: https://www.fhi.no/en/id/infectious-diseases/coronavirus/emergency-preparedness-register-for-covid-19/.

Madhavan, S., Bastarache, L., Brown, J. S., Butte, A. J. et al, (2021). Use of electronic health records to support a public health response to the COVID-19 pandemic in the United States: a perspective from 15 academic medical centers. *Journal of the American Medical Informatics Association,* 15, 28(2), 393–401. doi: 10.1093/jamia/ocaa287.

Popperl, S., Weiner, L. and King, N., (2021). Electronic Health Records May Be Delaying COVID-19 Vaccinations. *NPR – National Public Radio*. Retrieved from: https://www.npr.org/sections/coronavirus-live-updates/2021/02/10/966279323/electronic-health-records-may-be-delaying-covid-19-vaccinations.

Producten en diensten, (2021). Ministerie van Volksgezondheid, Welzijn en Sport. *www.cibg.nl.* Retrieved from: https://www.cibg.nl/producten-en-diensten [in Dutch].

Puhachova, M. V., (2021). Organization and functioning of electronic registers of health care systems of Northern European countries and Ukraine. Proceeding from Modern statistics: problems and prospects of development: XIX International Scientific and Practical Conference.(pp. 29–33). Kyiv: NASOA. Retrieved from: http://194.44.12.92:8080/jspui/handle/123456789/6165 [in Ukrainian].

Puhachova, M. V., Gladun, O. M., (2021). Registers of Healthcare Systems of Northen European Countries. *Demography and social economy*, 4 (46), 21–37. Retrieved from: https://doi.org/10.15407/dse2021.04.021 [in Ukrainian].

Registre. *eSundhed,* (2021). Retrieved from: https://www.esundhed.dk/Registre/ [in Danish].

Satterfield, B. A., Dikilitas, O. and Kullo, I. J., (2021). Leveraging the Electronic Health Record to Address the COVID-19 Pandemic. *Mayo Clinic Proceedings*, Vol. 96, *Issue 6,* pp. 1592–1608. Retrieved from: https://doi.org/10.1016/j.mayocp.2021.04.008.

Schwab, P., Mehrjou, A., Parbhoo, S., Celi, L. A., Hetzel, J. and Hofer, M. et al, (2021). Real-time prediction of COVID-19 related mortality using electronic health records. *Nature Communications,12,* 1058. Retrieved from: https://doi.org/10.1038/s41467-020-20816-7.

Sim, J. X. Y., Conceicao, E. P., Wee, L. E., Gan, W. H., Ling, M. L. and Venkatachalam, I. et al, (2021). Utilizing the electronic health records to create a syndromic staff surveillance system during the COVID-19 outbreak. *American Journal of Infection Control, Vol. 49, Issue 6,* pp. 685–689. Retrieved from: https://doi.org/10.1016/j.ajic.2020.11.003.

Slone, S., (2020). COVID-19 Pandemic Demonstrates Need for Electronic Health Record Interoperability. (2020). *The Council of State Governments*. Retrieved from: https://web.csg.org/covid19/2020/04/24/covid-19-pandemic-demonstrates-need-for-electronic-health-record-interoperability/.

Stanley, B., Davis, A., Jones, R., Julious, S., Ryan, D., Jackson, D. et al, (2021) Characteristics of patients in platform C19, a COVID-19 research database combining primary care electronic health record and patient reported information. *PLoS ONE,* 16 (10): e0258689. Retrieved from: https://doi.org/10.1371/journal.pone.0258689.

Sudat, S. E. K., Robinson, S. C., Mudiganti, S., Mani, A. and Pressman, A. R., (2021). Mind the clinical-analytic gap: Electronic health records and COVID-19 pandemic response. *Journal of Biomedical Informatics*, Vol. 116. Retrieved from: https://doi.org/10.1016/j.jbi.2021.103715.

Uleriigiline digiregistratuur kaivitub joudsasti, (2021). *TEHIK*. Retrieved from: https://www.tehik.ee/uudis/uleriigiline-digiregistratuur-kaivitub-joudsasti [in Estonian].

Wang, Q. Q., Xu, R. and Volkow, N. D., (2021). Increased risk of COVID-19 infection and mortality in people with mental disorders: analysis from electronic health records in the United States. *World Psychiatry, Vol. 20, Issue 1,* pp. 124–130. Retrieved from: https://doi.org/10.1002/wps.20806.

Williamson, E., Walker, A. J., Bhaskaran, K., Bacon, S., Bates, Ch. and Morton, C. E. et al., (2020). OpenSAFELY: factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients. *medRxiv*. Retrieved from: https://doi.org/10.1101/2020.05.06.20092999.

STATISTICS IN TRANSITION new series and STATISTICS OF UKRAINE, February 2023
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 295–320, DOI 10.59170/stattrans-2023-016
Received – 13.07.2022; accepted – 29.11.2022

Under military war weapon support the economic bond level estimation using generalized Petersen graph with imputation

Deepika Rajoriya[1], **Diwakar Shukla**[2]

ABSTRACT

Several countries of the world are involved in mutual and collaborative business of military equipments, weapons in terms of their production, sales, technical maintenance, training and services. As a consequence, manufacturing of boms, rockets, missiles and other ammunitions have taken structured and smooth shape to help others where and when needed. Often the military support among countries remain open for information to the media, but sometime remain secret due to the national security and international political pressure. Such phenomenon (hidden or open support) is a part of military supply chain and could be modeled like a Petersen graph considering vertices as countries and edges as economic bonds. For a large graphical structure, without sampling, it is difficult to find out average economic bonding (open & secret) between any pair of countries involved in the military business or support.

This paper presents a sample based estimation methodology for estimating the mean economic bond value among countries involved in the military support or business. Motivation to the problem is derived from current Russia-Ukraine war situation and a kind of hidden support to war by NATO countries. A node sampling procedure is proposed whose bias, mean-squared error and other properties are derived. Results are supported with empirical studies. Findings are compared with particular cases and confidence intervals are used as a basic tool of comparison. Pattern imputation is used together with a new proposal of CI-Imputation method who has been proved useful for filling the missing value, specially when secret economic support data from involved countries found missing. The current undergoing war between Ukraine and Russia and secret weapon, economic support from NATO countries is an application of the proposed methodology contained in this paper.

Key words: Graph, Petersen Graph, Estimator, Bias, Mean Squared Error (MSE), Optimum Choice, Confidence intervals (CI), Nodes (vertices), Pattern Imputation, CI-Imputation (LL-imputation and UL-imputation), Economic Bonds, Military War, Weapon Support.

1. Introduction

The Russian war in Ukraine is a kind of complecated political event prolonged over time frame. After the pass of many months it is hard to predict about the ultimate date of war end from either side. The Russian invasion was started in Feb, 2022 and by April, 2022 as per United Nations High Commissioner for Humen rights report more than 2800 death of civilians occurred in Ukraine. There is big difference between military capacity of NATO, Russia and Ukraine as per record of 2022.

[1]Dr. Harisingh Gour Centeral University, Sagar (M.P.) India. E-mail: deepikarajoriya2112@gmail.com
[2]Dr. Harisingh Gour Centeral University, Sagar (M.P.) India. E-mail: diwakarshukla@rediffmail.com.

© Deepika Rajoriya, Diwakar Shukla. Article available under the CC BY-SA 4.0 licence

Table 1.1 World Nuclear Forces [see link the independent resource on global security]

Country	Deployed warheads*	Other warheads**	Total 2021	Total 2020
USA	1800	3750	5550	5800
Russia	1625	4630	255	6375
UK***	120	105	225	215
France	280	10	290	290
China	–	350	350	320
India	–	156	156	150
Pakistan	–	165	165	160
Israel	–	90	90	90
North Korea****	–	[40-50]	[40-50]	[30-40]
Total	3825	9255	13080	13400

In view of report China support to Ukraine [see Link], the Russia asked China for military assistance of equipments and economic support. It is customary to ask for business deal, financial and military support from either country during the war period from neighbouring coutries, China support to Ukraine and Nine big question answered by Russia [See links].

Assume several countries of the world involved with each other in trading of war-weapons. They are having economic bonding among themselves in terms of export, import, supply and manufacturing of war-weapons etc. As an example, several NATO countries are involved in mutual colllaboration and exchange of weapons during the current war of Ukrain and Russia. All Europian countries can be treated as a group involved in supply of open and secrete war weapons to countries involved in fight to save the own territory. One can visualize the current war scenario as under:

(a) Type I: Between war group countries, the open and accountable war-weapon business.

(b) Type II: Within a country accountable war-weapon business.

(c) Type III: Secret (unaccountable) war-weapon business andsupport between countries.

The table 1.2 reveals such the structure of type I, II, III in terms of numerical values for only five countries A, B, C, D and E (treating a_{ij} as business value, i, j =1,2,3,4).

Table 1.2 Countries and War Period Exchange Economic

Countries	Type I (units)	Type II (units)	Type III (units)
$A \to B$	a_{11}	a_{12}	a_{13}
$B \to C$	a_{21}	a_{22}	a_{23}
$C \to D$	a_{31}	a_{32}	a_{33}
$D \to E$	a_{41}	a_{42}	a_{43}

(d) External Economic Bonding: It is defined as the accountable weapon trade between two countries which is auditable.

(e) Internal Economic Bonding: It is the internal accountable war-weapon trade within country among army, defence, security forces and internal manufacturing companies.

(f) Secret Economic Bonding: The trade of military products between and within countries who are secret (un-accountable) like many NATO countries are supporting Ukraine providing secret war weapons (as per reports).

Remark 1 *The information about type-I, type-II and type-III business (Economic bonds) can be obtained through the National Audit reports, United Nations reports (like IMF reports, Security Council reports, media and spying agencies reports etc.) either immediately or after long time when war is over. For intermediate Economic bonds, within country, the ordinance factories who are producing gun, tanks, arms and ammunitions and supplying those to own army, Paramilitary Forces, Private Security agencies within the country may be considered. For secret Economic bonding, information about only few units in sample is required which may available, at any instant, through authentic media sources.*

1.1. Objective

In view to Ukraine-Russia war, interest of data analyst is to evaluate the average amount of internal economic and secret economic bond together existing between any two countries using sampling techniques and imputation method if secret economic data found missing.

1.2. Motivation

The Europian country organizations (like EU or NATO) have open and free trade policies among them in currency EUROS. During war and military action, the secret economic and infrastructrue exchange is an obvious possible internal factor. A Petersen graph can be used as a model tool to represent such real situation where vertices (inner and outer) be countries and edges (weapon deal) be the trade among them during war period. Outer edges are for accountable weapon business between countries, intermediate edges are for within country and inner edges represent secret business. The current war and hidden weapon supply (with financial support) have motivated to model the real war situation like a Petersen graph.

The generalized Petersen graph G(n,k) was introduced by Coxeter et al. (1950) and named by Watkins (1969) from very interesting family of trivalent graphs that can be described by only two integer parameters. They include Hamiltonian and non-Hamiltonian graph, Bipartite and non-Bipartite graphs, vertex transitive and non-vertex transitive graphs, cayley and non-cayley graphs of girth 3,4,5,6,7 or 8 [Krnc, M. et al. (2018)]. A generalized Petersen graph G(n,k) is a family of cubic graph who is 3-regular graph. Following notations of Watkins et al. (1969) for a given integer n and $k < \frac{n}{2}$ one can define a Petersen graph $G\ (n,k)$ as a graph of vertex set $(\mu_0, \mu_1, ..., \mu_{n-1}, \nu_0, \nu_1, ..., \nu_{n-1})$ and edge set partitioned into three equal parts $(\mu_i\mu_{i+1}, \mu_i\nu_i, \nu_i\nu_{i+k} \mid 0 \leq i \leq n-1)$ where subscripts are to be read modulo n. The $G\ (3,1)$ and $G\ (4,1)$ are given below as examples (fig 1.1 and 1.2).

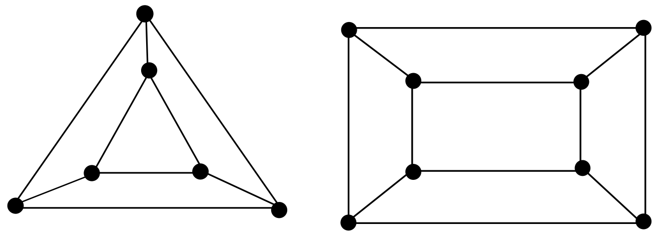


Fig 1.1 Petersen Graph

Let $\mu = (\mu_1, \mu_2, \mu_3, ...)$ denotes a set of vertices and $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3, ...)$ is a set of edges. The $G = (\mu, \varepsilon, R)$ consitutes a graph, in general, where R is a set of relations.
Example of five NATO countries linked like a Petersen graph (see fig 1.2) as under:
Vertices (μ_1 *and* ν_1)$\rightarrow$ Poland
Vertices (μ_2 *and* ν_2)$\rightarrow$ Hungary
Vertices (μ_3 *and* ν_3)$\rightarrow$ Bulgaria
Vertices (μ_4 *and* ν_4)$\rightarrow$ Romania
Vertices (μ_5 *and* ν_5)$\rightarrow$ Turkey.
The shape of graph can be extended to 30 or more NATO countries with similar edge-connectivity in inner and outer form (fig 1.2).

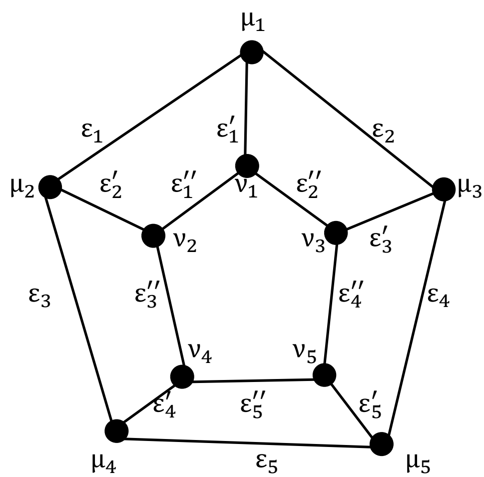


Fig 1.2 Petersen Graph G(5,1)

Table 1.3 Relation of Vertices and Edges in Petersen Graph

S.No.	Set μ	Set ν
1.	$\mu_1 = (\varepsilon_1, \varepsilon_2, \varepsilon_1')$	$\nu_1 = (\varepsilon_1'', \varepsilon_2'', \varepsilon_1')$
2.	$\mu_2 = (\varepsilon_1, \varepsilon_3, \varepsilon_2')$	$\nu_2 = (\varepsilon_1'', \varepsilon_3'', \varepsilon_2')$
3.	$\mu_3 = (\varepsilon_2, \varepsilon_4, \varepsilon_3')$	$\nu_3 = (\varepsilon_2'', \varepsilon_4'', \varepsilon_3')$
4.	$\mu_1 = (\varepsilon_3, \varepsilon_5, \varepsilon_4')$	$\nu_4 = (\varepsilon_3'', \varepsilon_5'', \varepsilon_4')$
5.	$\mu_1 = (\varepsilon_4, \varepsilon_5, \varepsilon_5')$	$\nu_5 = (\varepsilon_4'', \varepsilon_5'', \varepsilon_5')$

Note 1.1 The set of vertices $\mu = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$ denotes countries for external economic level where as set $\nu = (\nu_1, \nu_2, \nu_3, \nu_4, \nu_5)$ denotes some countries for secret economic level. The paired set of vertices $w = \{(\mu_i, \nu_i) : i = 1,2,3,4,5\}$ represents some countries for internal economic level.

Table 1.4 Node-Edge Matrix of Petersen Graph

	ε_1	ε_2	ε_3	ε_4	ε_5	ε_1'	ε_2'	ε_3'	ε_4'	ε_5'	ε_1''	ε_2''	ε_3''	ε_4''	ε_5''	*row total*
μ_1	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	3
μ_2	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	3
μ_3	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	3
μ_4	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	3
μ_5	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	3
ν_1	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	3
ν_2	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	3
ν_3	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	3
ν_4	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	3
ν_5	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	3

1.3. Pattern Imputation

In light of fig 1.2 and table 1.4, for large number of outer vertices N and large number of inner vertices N, the general relationship R is

At i=1 $\mu_1 \rightarrow (\varepsilon_1, \varepsilon_2, \varepsilon_1'); \nu_1 \rightarrow (\varepsilon_1'', \varepsilon_2'', \varepsilon_1')$

$\mu_i \rightarrow (\varepsilon_{i-1}, \varepsilon_{i+1}, \varepsilon_i'); \nu_i \rightarrow (\varepsilon_{i-1}'', \varepsilon_{i+1}'', \varepsilon_i'), i = 2,3...n-1$

At i=N $\mu_N \rightarrow (\varepsilon_{N-1}, \varepsilon_{N+1}, \varepsilon_N'); \nu_N \rightarrow (\varepsilon_{N-1}'', \varepsilon_N'', \varepsilon_N')$.

Under large N, for external set of vertices μ, secret set of vertices ν and internal set ω, the pattern imputation is proposed as under:

Step I At i=2 take $\mu_i \rightarrow (\varepsilon_{i-1}, \varepsilon_{i+1}, \varepsilon_i'); \nu_i \rightarrow (\varepsilon_{i-1}'', \varepsilon_{i+1}'', \varepsilon_i'), i = 2,3...N-1$

Step II At i=1 impute in step I, ε_0 by ε_1, ε_0'' by ε_1'' and take $\mu_1 \rightarrow (\varepsilon_1, \varepsilon_2, \varepsilon_1'); \nu_1 \rightarrow (\varepsilon_1'', \varepsilon_2'', \varepsilon_1')$

Step III At i=N, impute in step I, ε_{N+1} by ε_N and ε_{N+1}'' by ε_N'' and take $\mu_N \rightarrow (\varepsilon_{N-1}, \varepsilon_N, \varepsilon_N')$;
$\nu_N \rightarrow (\varepsilon_{N-1}'', \varepsilon_N'', \varepsilon_N')$.

To note that imputation of ε_0 *by* ε_1, ε_{N+1} *by* ε_N and ε_0'' by ε_1'', ε_{N+1}'' by ε_N'' is like a specific imputation just to maintaine a pattern so it is called pattern imputation. In general, it may random imputation also like ε_0 to replace by any ε_i, ε_{N+1} by any ε_i, ε_0'' by any ε_i'' ε_{N+1}'' by any ε_i'' randomly chosen. The pattern imputation is closed to the nearest neighbour imputation, but earlier maintains a pattern but later do not.

1.4. Economic Bond Structure Between Countries

Looking at fig 1.2 and assuming large N, the Generalised Petersen Graph G(N,k) can be expressed having edge weights as different economic level bonds between vertices (countries).

(a) **Single Economic Bonding:** The bonding is between any vertex pair (μ_i, μ_{i+1}) at external level, any pair (ν_i, ν_{i+1}) at secret level and any vertex pair (μ_i, ν_i) at internal level. The symbols δ_i, δ_i', δ_i'' represent value of corresponding bonding as shown in fig 1.3.

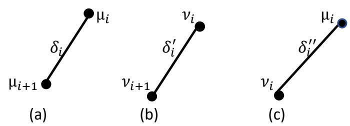


Fig 1.3 Single Economic Bonding

(b) **Double Economic Bonding:** This bonding is between one external and one internal pair of vertices or one internal with one secret pair of vertices. The α_i and α_i' are edge-weights revealed in fig 1.4. Double economic bond may be taken as external+internal as one part (one variable) and (internal+secret) as another part (other variable).

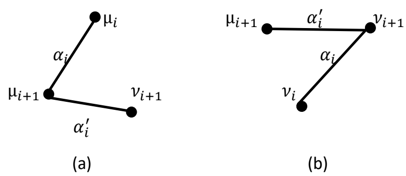


Fig 1.4 Double Economic Bonding

(c) **Triple Economic Bonding:** This consitutes bonding among two vertex pairs at external and secret level and one pair at internal level. The β_i, β_i', β_i'' are edge weights as economic levels shown in fig 1.5.

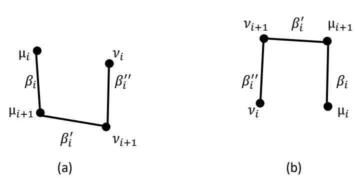


Fig 1.5 Triple Economic Bonding

2. Estimation

In view of Ukraine-Russia war situation and secret military help by NATO countries, authors considered the case of double economic bond estimation only in the content of this paper assuming large N. Define $U_i = \varepsilon_i + \varepsilon_i'$ as external+internal edges and $Z_i = \varepsilon_i'' + \varepsilon_i'$ as secret+internal edges.

$\bar{U} = \frac{\sum_{i=1}^{N} U_i}{N} = \frac{\sum_{i=1}^{N}(\varepsilon_i+\varepsilon_i')}{N}$; $\bar{Z} = \frac{\sum_{i=1}^{N} Z_i}{N} = \frac{\sum_{i=1}^{N}(\varepsilon_i''+\varepsilon_i')}{N}$ (Population means)

$S_U^2 = \frac{\sum_{i=1}^{N}(U_i-\bar{U})^2}{N-1}$; $S_Z^2 = \frac{\sum_{i=1}^{N}(Z_i-\bar{Z})^2}{N-1}$ (Population mean square)

$C_U = (\frac{S_U}{\bar{U}})$; $C_Z = (\frac{S_Z}{\bar{Z}})$ (Population coefficient of variation)

$S_{UZ} = \frac{\sum_{i=1}^{N}(U_i-\bar{U})(Z_i-\bar{Z})}{N-1}$; $\rho_{UZ} = \rho_{ZU} = \frac{S_{UZ}}{S_U.S_Z}$ (Population correlation coefficient)

Let a simple random sample of large size n ($n < N$) containing vertices like (μ_j, ν_j), j=1,2,3...n is drawn from N vertices using without replacement procedure.

Sample statistic are:

$$\bar{u} = \frac{\sum_{j=1}^{n} u_j}{n} \quad (sample\ mean\ of\ external + internal\ edges) \tag{1}$$

$$\bar{z} = \frac{\sum_{j=1}^{n} z_j}{n} \quad (sample\ mean\ of\ internal + secret\ edges) \tag{2}$$

$$s_u^2 = \frac{\sum_{j=1}^{n}(u_j-\bar{u})^2}{n-1};\ s_z^2 = \frac{\sum_{j=1}^{n}(z_j-\bar{z})^2}{n-1} \quad (sample\ mean\ square\ of\ U\ and\ Z) \tag{3}$$

$$c_u = (\frac{s_u}{\bar{u}});\ c_z = (\frac{s_z}{\bar{z}}) \quad (sample\ coefficient\ of\ variation) \tag{4}$$

$$s_{uz} = \frac{\sum_{j=1}^{n}(u_j-\bar{u})(z_j-\bar{z})}{n-1} \tag{5}$$

$$\rho_{uz} = \frac{s_{uz}}{s_u.s_z} \quad (sample\ correlation\ between\ u\ and\ z) \tag{6}$$

The out mean $\bar{U}$ is assumed known but inner mean $\bar{Z}$ is unknown and the aim of this paper is to estimate $\bar{Z}$ using known ($\bar{u}$,$\bar{z}$, $\bar{U}$) by an appropriate efficient estimation strategy along with imputation for missing.

2.1. Proposed Estimation Strategy

To estimate unknown $\bar{Z}$ in the internal+secret double economic bond and support obtained between any two war involved countries, the proposed estimation strategy $[using\ \bar{z}, \bar{U}, \bar{u}]$ is:

$$E = (\bar{z})[\phi_1(\bar{u}, \bar{U})][\phi_2(\bar{u}, \bar{U})]^{-1}$$

where,

$$\phi_1(\bar{u},\bar{U}) = [(A+C+D)\bar{U}+gB\bar{u}]$$

$$\phi_2(\bar{u},\bar{U}) = [(A+gB+D)\bar{U}+C\bar{u}]$$

$A=(q-1)(q-2); B=(q-1)(q-4); C=(q-2)(q-3)(q-4); D=(q-1)(q-2)(q-3)(q-4)(q-5), g=\frac{n}{N},\ 0<q<\infty$

The proposed is in accordance with shukla et al. (2014) but as a part of new structure, a term is added D which is in power five in q . At q=4, as a special case, the proposed strategy converts to the internal+secret sample mean based economic bond value estimation through a sample.

3. Setting Approximations

For two real numbers h_1 *and* h_2, $|h_1|<1$ *and* $|h_2|<1$, assuming both N, n large, one can express approximations as per Singh et al.(2003), Chochran et al. (2005), Shukla et al. (2014) and Rajoriya et al. (2021).

$$\bar{z} = \bar{Z}(1+h_1) \tag{7}$$

$$\bar{u} = \bar{U}(1+h_2) \tag{8}$$

Let $E^*(.)$ denotes expected value of random variables $\bar{z}$ and $\bar{u}$, then one can get the followings Deo et al. (2001), Shukla et al. (2002), Shukla et al. (2018), Shukla et al. (2020), Donga et al. (2021) and Rajoriya et al.(2021).

$$E^*(h_1) = E^*(h_2) = 0 \tag{9}$$

$$E^*({h_1}^2) = \frac{(N-n)}{Nn}C_Z^2 \tag{10}$$

$$E^*({h_2}^2) = \frac{(N-n)}{Nn}C_U^2 \tag{11}$$

$$E^*(h_1h_2) = \left(\frac{(N-n)}{Nn}\right)(\rho_{ZU}.C_Z.C_U) \tag{12}$$

Theorem 1 *Under large sample approximations, the proposed E could be expressed as:*

$$E = \bar{Z}\left[(1+h_1)+\Delta^*\left\{(h_1+h_1h_2)-\frac{Ch_2^2}{\Delta}\right\}\right]$$

where, $\Delta=(A+gB+C+D)$; $\Delta^*=[\frac{(gB-C)}{\Delta}]$

proof:

$$E = (\bar{z})[\phi_1(\bar{u},\bar{U})][\phi_2(\bar{u},\bar{U})]^{-1}$$

where,

$$\phi_1(\bar{u},\bar{U}) = [(A+C+D)\bar{U} + gB\bar{u}]$$

$$\phi_2(\bar{u},\bar{U}) = [(A+gB+D)\bar{U} + C\bar{u}]$$

Using (7) and (8), $|h_1| < 1$, $|h_2| < 1$

$$\phi_1(\bar{u},\bar{U}) = [(A+C+D)\bar{U} + gB\{\bar{U}(1+h_2)\}] \tag{13}$$

$$\phi_2(\bar{u},\bar{U}) = [(A+gB+D)\bar{U} + C\{\bar{U}(1+h_2)\}] \tag{14}$$

Then $\phi_1(\bar{u},\bar{U})$ could be expressed as:

$$\phi_1(\bar{u},\bar{U}) = [\bar{U}(A+gB+C+D)]\left[1+\frac{(gBh_2)}{(A+gB+C+D)}\right] \tag{15}$$

Since $|h_2| < 1$, therefore $|\frac{gBh_2}{(A+gB+C+D)}| < 1, \ \forall\, g > 0, q > 0$

Moreover, for $\phi_2(\bar{u},\bar{U})$ using expansion of $(1+x)^{-1}$, one gets

$$[\phi_2(\bar{u},\bar{U})]^{-1}\,[(A+gB+C+D)\bar{U}+\bar{U}Ch_2]^{-1}$$

$$= (\bar{U})^{-1}[A+gB+C+D]^{-1}\left[1+\frac{Ch_2}{(A+gB+C+D)}\right]^{-1}$$

$$= (\bar{U})^{-1}[A+gB+C+D]^{-1}\left[1-\frac{Ch_2}{(A+gB+C+D)}+\frac{C^2h_2^2}{(A+gB+C+D)}\cdots\right]$$

Define $\Delta = (A+gB+C+D)$, then one can express proposed E as

$$E = (\bar{z})[\phi_1(\bar{u},\bar{U})][\phi_2(\bar{u},\bar{U})]^{-1}$$

$$E = \bar{Z}(1+h_1)\left[1+\frac{gBh_2}{\Delta}\right]\left[1-\frac{Ch_2}{\Delta}+\frac{C^2h_2^2}{\Delta^2}\cdots\right]$$

$$E = \bar{Z}(1+h_1)\left[1-\frac{Ch_2}{\Delta}+\frac{C^2h_2^2}{\Delta^2}+\left\{\frac{gBh_2}{\Delta}-\frac{gBCh_2^2}{\Delta^2}+\frac{gBC^2h_2^3}{\Delta^3}\cdots\right\}\right]$$

$$E = \bar{Z}\left[(1+h_1)+\frac{(gB-C)}{\Delta}\left\{(h_2+h_1h_2)-\frac{Ch_2^2}{\Delta}\right\}\right]$$

which is expressed after ignoring terms $(h_1^s.\ h_2^t), (s+t) > 2,\ s,t = 0,1,2,3,4...$ because of having high power on h_1 and h_2. The denominator Δ is high for $g > 0$, therefore, one can narrate that contribution of these terms in estimation will be very low (negligible).

Define $\Delta^* = \frac{(gB-C)}{\Delta}$. Then $E = \bar{Z}\left[(1+h_1)+\Delta^*\left\{(h_1+h_1h_2)-\frac{Ch_2^2}{\Delta}\right\}\right]$

Theorem 2 *The bias of estimator E under (7), (8) using theorem 1 is:*

$B[E] = Bias[E] = \bar{Z}\left[\Delta^*\left\{\frac{N-n}{Nn}\right\}\left\{\rho_{ZU}.C_Z.C_U) - \frac{C}{\Delta}C_U^2\right\}\right]$
where $\rho_{UZ} = \rho_{ZU}$ is correlation coefficient between double economic bond variables U and Z in Petersen graph.

proof: The $E^*(.)$ denotes expected value of the proposed estimator E and $B[E] = [E^*(E) - \bar{Z}]$

Now $E^*(E) = E^*\left[\bar{Z}(1+h_1) + \bar{Z}\Delta^*\left\{h_1 + h_1h_2 - \frac{Ch_2^2}{\Delta}\right\}\right]$

$= \left[\bar{Z} + \bar{Z}E^*(h_1) + \bar{Z}\Delta^*\left\{E^*(h_1) + E^*(h_1h_2) - \frac{CE^*(h_2^2)}{\Delta}\right\}\right]$

$= \left[\bar{Z} + \bar{Z}\Delta^*\left\{E^*(h_1h_2) - \frac{C}{\Delta}E^*(h_2^2)\right\}\right]$ Using (7) and (8) and theorem 1

$B[E] = [E^*(E) - \bar{Z}]$

$= \bar{Z}\left[\Delta^*\left\{E^*(h_1h_2) - \frac{C}{\Delta}E^*(h_2^2)\right\}\right]$

$= \bar{Z}\left[\Delta^*\left(\frac{N-n}{Nn}\right)\left\{(\rho_{ZU}C_zC_U) - (\frac{C}{\Delta})C_U^2\right\}\right]$

Corollary 1 *The estimator E is almost unbiased under condition*

$(\rho_{ZU}C_zC_U) = (\frac{C}{\Delta})C_U^2 \implies \frac{C}{\Delta} = \rho_{ZU}(\frac{C_Z}{C_U}) = M\ (Let)$

$\implies \frac{C}{(A+gB+C+D)} = M$

$$\implies M(A+gB+C+D) + C(M-1) + MD = 0 \tag{16}$$

Note 3.1 The equation (17) is having highest power five in terms of q. Therefore, it may has maximum of five roots satisfying the equation. Best root will be that having lowest mean square error (MSE).

Theorem 3 *The mean squared error (MSE) of the proposed strategy is*

$MSE[E] = \bar{Z}^2\left[(\frac{N-n}{Nn})\left\{C_Z^2 + (\Delta^*)^2C_U^2 + 2\Delta^*\rho_{UZ}C_UC_Z\right\}\right]$
proof: $MSE[E] = E^*[E - \bar{Z}]^2$

$= E^*\left[\bar{Z}(1+h_1) + \Delta^*\left\{h_1 + h_1h_2 - \frac{Ch_2^2}{\Delta} + ...\right\} - \bar{Z}\right]^2$

$= E^*\left[\bar{Z}(h_1 + \Delta^*h_2)\right]^2$ ignoring terms $(h^sh^t), (s+t) > 2, s,t = 1,2,3,4,5$

$= \bar{Z}^2\left[E^*(h_1^2) + (\Delta^*)^2E^*(h_2^2) + 2\Delta^*E^*(h_1h_2)\right]$

$$\implies MSE[E] = \bar{Z}^2\left[(\frac{N-n}{Nn})\left\{C_Z^2 + (\Delta^*)^2C_U^2 + 2\Delta^*\rho_{UZ}C_UC_Z\right\}\right] \tag{17}$$

Theorem 4 *The minimum (optimum) mean squared error is attained when* $\Delta^* = -M$

where $M = \rho_{UZ}(\frac{C_Z}{C_U})$.

proof: Differentiating MSE[E] with respect to the term Δ^* and equate to zero, one gets;

$$\frac{MSE[E]}{\Delta^*} = 0 \implies \Delta^* = -\rho_{UZ}(\frac{C_Z}{C_U}) = -M \tag{18}$$

Corollary 2 *The optimuum MSE expression (19) could be expressed as*

$$\frac{(gB-C)}{(A+gB+C+D)} = -M$$

$$\implies AM + gB(M+1) + C(M-1) + DM = 0 \tag{19}$$

Note 3.2 Equation (20) of optimum MSE is having highest power five on term q, therefore, there will be maximum of five roots of equation (20). The best q will be that containing lowest bias value. The proposed strategy E attains the optimum level of MSE and reduces the bias too. This is a novel feature of proposed estimation procedure E.

4. Numerical Illustration

Remark 2 *It is difficult to get real and reliable data of secret Economic bond immediately during the war (like Russia & Ukraine and support of NATO countries). But data of Internal Economic Bond among various ordinance factories within countries could be obtained when the audit and assessment reports, by Auditors, are available. It takes several years to come and to get published. The current Russia-Ukraine war be treated as an application of the proposed whose data will be published after long time. In absence of that, an artificial data set is used just to test the proposed methodology and to demonstrate the suggested procedure to the article readers.*

Remark 3 *There may uneven economic distribution support (as open & secrete) by various involved countries. But, one can assume nearly homogeneous support by most of NATO countries to the Ukraine, specially at the starting duration of war. Later on, as the war progresses, the open and hidden, both kinds of economic support may convert into heterogeneous distributions. In this paper, the almost homogeneous economic support, as was in beginning period of the war is assumed. It is a restriction also in the content of the paper.*

Remark 4 *The size N, if large, will not affect the properties of the proposed methodology using Petersen graph model. In fact, the Petersen graph is a closed network of vertices which can accommodate any number of additional vertices, as and when requried, without loosing structure and properties.*

Define F= Secret Economic Bond $=\varepsilon_i^{''}$; G= External Economic Bond $=\varepsilon_i$; H= Internal Economic Bond $=\varepsilon_i^{'}$. Consider the generalized Petersen structure with N=150. The assumed economic bond values are considered below:

Table 4.1 Military War Weapon Assumed Data of N=150 Countries as Population

S.No.	F = ε_i''	G = ε_i	H = ε_i'	S.No.	F = ε_i''	G = ε_i	H = ε_i'
1.	25 units	43units	86units	76.	41 units	87units	34units
2.	53 units	81units	64units	77.	75 units	32units	66units
3.	34 units	14units	86units	78.	48 units	32units	71units
4.	43 units	61units	74units	79.	87 units	92units	56units
5.	37 units	28units	69units	80.	49 units	22units	76units
6.	91 units	23units	41units	81.	65 units	86units	56units
7.	34 units	48units	72units	82.	45 units	33units	31units
8.	92 units	43units	21units	83.	49 units	64units	88units
9.	35 units	63units	71units	84.	93 units	21units	65units
10.	27 units	83units	34units	85.	75 units	83units	89units
11.	51 units	63units	86units	86.	46 units	26units	18units
12.	63 units	72units	65units	87.	68 units	37units	28units
13.	39 units	84units	42units	88.	88 units	63units	29units
14.	52 units	26units	75units	89.	28 units	44units	75units
15.	84 units	35units	42units	90.	39 units	42units	56units
16.	28 units	39units	67units	91.	37units	47units	76units
17.	56 units	42units	63units	92.	82 units	56units	96units
18.	81 units	33units	26units	93.	17 units	47units	89units
19.	29 units	57units	76units	94.	76 units	44units	28units
20.	85 units	38units	43units	95.	45 units	63units	60units
21.	91 units	34units	78units	96.	77 units	42units	63units
22.	38 units	49units	65units	97.	29 units	51units	36units
23.	57 units	63units	84units	98.	39 units	53units	56units
24.	19 units	43units	96units	99.	78 units	88units	40units
25.	65 units	36units	73units	100.	20 units	75units	64units
26.	48 units	96units	21units	101.	73units	37units	58units
27.	43 units	65units	92units	102.	84 units	73units	36units
28.	45 units	39units	17units	103.	95 units	43units	21units
29.	83 units	91units	26units	104.	58 units	68units	28units
30.	57units	48units	21units	105.	71 units	39units	50units
31.	23 units	58units	61units	106.	47 units	40units	19units
32.	47 units	82units	53units	107.	85 units	73units	26units
33.	27 units	63units	73units	108.	60 units	53units	44units
34.	98 units	34units	61units	109.	28 units	49units	81units
35.	45 units	23units	54units	110.	35 units	63units	66units
36.	81 units	53units	66units	111.	48 units	28units	39 units
37.	22 units	93units	81units	112.	56 units	54units	87 units
38.	55 units	42units	76units	113.	41 units	40 units	81 units
39.	29 units	63units	66units	114.	45 units	63 units	21 units
40.	68 units	41units	96units	115.	35 units	71 units	66 units
41.	25 units	93units	46units	116.	88 units	23 units	86 units
42.	63 units	71units	32units	117.	35 units	43 units	88 units
43.	73 units	61units	24units	118.	69 units	40 units	66 units
44.	58 units	83units	46units	119.	38 units	33units	96units
45.	48 units	43units	22units	120.	68 units	43units	56units
46.	31 units	48units	69units	121.	21 units	84 units	26 units
47.	47 units	33units	26units	122.	25 units	49units	77units
48.	35 units	87units	76units	123.	48 units	64 units	92 units
49.	63 units	71units	36units	124.	20 units	63units	29 units
50.	85 units	53units	46units	125.	28 units	33 units	83 units
51.	76 units	29units	36units	126.	77 units	62 units	55 units
52.	32 units	61units	59units	127.	60 units	43 units	56 units
53.	47 units	93units	73units	128.	65 units	74 units	78 units
54.	93 units	84units	64units	129.	48 units	66 units	58 units
55.	55 units	84units	29units	130.	94 units	47units	76units
56.	48 units	19units	36units	131.	59 units	31 units	63 units
57.	71 units	94units	68units	132.	76 units	93 units	84 units
58.	92units	83units	57units	133.	95 units	73 units	66units
59.	28 units	59units	28units	134.	70 units	83 units	56 units
60.	38 units	47units	71units	135.	46 units	29 units	46units
61.	93 units	72units	65units	136.	79 units	92 units	36units
62.	35 units	83units	57units	137.	54 units	54 units	47 units
63.	45 units	84units	91units	138.	80 units	43units	98 units
64.	46 units	52units	29units	139.	95 units	46units	19 units
65.	15 units	73units	82units	140.	39 units	63 units	93 units
66.	37 units	87units	62units	141.	97 units	76 units	34 units
67.	93 units	13units	96units	142.	85 units	94 units	33 units
68.	75 units	84units	56units	143.	76 units	33 units	57 units
69.	39 units	83units	92units	144.	79 units	65 units	88 units
70.	72units	65units	86units	145.	83 units	60 units	59 units
71.	47 units	41units	68units	146.	90 units	22 units	86units
72.	85 units	38units	21units	147.	79 units	39 units	88 units
73.	68 units	91units	26units	148.	46 units	55 units	39 units
74.	45 units	38units	56units	149.	98 units	68 units	88 units
75.	30 units	43units	82units	150.	29 units	85 units	89 units

Remark 5 *Define Secret level+ Internal level Economic Bond* $=\varepsilon_i'' + \varepsilon_i' = Z_i$ *; Internal level + External level Economic Bond* = $\varepsilon_i' + \varepsilon_i = U_i$.

Table 4.2 Double Economic Bond (in U and Z) Data of N=150 Countries as Population (from table 4.1)

$Z_i=\varepsilon_i''+\varepsilon_i'=F+G$	$U_i=\varepsilon_i'+\varepsilon_i=G+H$	$Z_i=\varepsilon_i''+\varepsilon_i'=F+G$	$U_i=\varepsilon_i'+\varepsilon_i=G+H$
111	129	75	121
117	145	141	98
120	100	119	103
117	135	143	148
106	97	125	98
132	64	121	142
106	120	76	64
113	64	137	152
106	134	158	86
61	117	164	172
137	149	64	44
128	137	96	65
81	126	117	92
127	101	103	119
126	77	95	98
95	106	113	123
119	105	178	152
107	59	106	136
105	133	104	72
128	81	105	123
169	112	140	105
103	114	65	87
141	147	95	109
115	139	118	128
138	109	84	139
69	117	131	95
135	157	105	94
62	56	116	64
109	117	86	96
78	69	121	89
84	119	66	59
100	135	111	99
100	136	104	97
159	95	109	130
99	77	101	129
147	119	87	67
103	174	143	141
131	118	122	121
95	129	66	84
164	137	101	137
71	139	174	109
95	103	123	131
97	85	135	106
104	129	134	129
70	65	124	99
100	117	47	110
73	59	102	126
111	163	140	56
99	107	49	92
131	99	111	116
112	65	132	117
91	120	116	99
120	166	143	152
157	148	106	124
84	113	170	123
84	55	122	94
139	162	160	177
149	140	161	139
56	87	126	139
109	118	92	75
158	137	115	128
92	140	101	101
136	175	178	141
75	81	114	65
97	155	132	156
99	149	131	110
189	109	118	129
131	140	133	90
131	175	167	153
158	151	142	119
115	109	176	108
106	59	167	127
94	117	85	94
101	94	186	156
112	125	118	174

Table 4.3 Petersen Graph Population Parameters (table 4.1)

S.No.	Parameters	Value	Description/(Section 2.0)
1.	N	150	Population size
2.	n	40	Sample size
3.	$\bar{Z}$	116	Population Mean
4.	$\bar{U}$	115	Population Mean
5.	S_Z	29.4903	Populatin Mean Square
6.	S_U	30.2076	Population Mean Square
7.	C_Z	0.2542	Population Coefficient of Variation
8.	C_U	0.2626	Population Coefficient of Variation
9.	ρ_{UZ}	0.4217	Population Correlation Coefficient
10.	M	0.4082	Using Corollary 1

Table 4.4 Almost Unbiased Choice of q for given (M, g) [from eq (17)]

S.No.	M	g	Choice of q	Bias	MSE
1.	0.4082	0.2666	$q_1 = 1.0756$	0.1925	28.4753
2.	0.4082	0.2666	$q_2 = 1.9709$	0.0420	56.1204
3.	0.4082	0.2666	$q_3 = 2.9073$	-0.0323	14.7597
4.	0.4082	0.2666	$q_4 = --$	–	–
5.	0.4082	0.2666	$q_5 = --$	–	–

Table 4.5 Choice of q for Optimum MSE for given (M, g) [from eq (20)]

S.No.	M	g	Choice of q	MSE	Bias
1.	0.4082	0.2666	$q_{1(opt)} = 0.6335$	13.1075	-0.0007
2.	0.4082	0.2666	$q_{2(opt)} = 1.8270$	13.1075	-0.3053
3.	0.4082	0.2666	$q_{3(opt)} = 2.9830$	13.1075	-0.0236
4.	0.4082	0.2666	$q_{4(opt)} = --$	–	–
5.	0.4082	0.2666	$q_{5(opt)} = --$	–	–

Table 4.6 Special Cases At q = 1,2,3,4,5 for (g= 0.2666, M= 0.0773)

S.No.	q	A	B	C	D	Bias(theorem 2)	MSE (theorem 3)
1.	1	0	0	-6	0	0.0868	19.0684
2.	2	0	-2	0	0	0.0599	48.8626
3.	3	2	-2	0	0	-0.0217	13.1413
4.	4	6	0	0	0	0.0000	15.9440
5.	5	12	4	6	0	-0.0035	13.4878

Tables 4.7 Ready Reckoner for Choice of q Providing almost Unbiasedness for given (M,g) (Using corollary 1, eq. (17))[$Range\ 0.05 \le M \le 0.95;\ Range\ 0.3 \le g \le 0.9$]

S.No.	M	g	Choice of q	Bias	MSE	S.No.	M	g	Choice of q	Bias	MSE
1.	0.05	0.3	$q_1 = 1.0220$	0.0977	19.9902	31.	0.65	0.3	$q_1 = 1.1050$	0.2707	35.7252
2.	0.05	0.3	$q_2 = 1.8939$	0.0578	47.2832	32.	0.65	0.3	$q_2 = 1.9599$	0.0357	51.0935
3.	0.05	0.3	$q_3 = 2.9837$	-0.0277	13.1919	33.	0.65	0.3	$q_3 = 2.8847$	-0.0383	20.5160
4.	0.05	0.3	$q_4 = --$	–	–	34.	0.65	0.3	$q_4 = --$	–	–
5.	0.05	0.3	$q_5 = --$	–	–	35.	0.65	0.3	$q_5 = --$	–	–
6.	0.05	0.6	$q_1 = 1.0220$	0.1097	19.8885	36.	0.65	0.6	$q_1 = 1.1071$	0.2536	33.4299
7.	0.05	0.6	$q_2 = 1.9903$	0.5489	278.9120	37.	0.65	0.6	$q_2 = 1.9124$	0.0285	51.6161
8.	0.05	0.6	$q_3 = 2.9980$	-0.90664	62.8639	38.	0.65	0.6	$q_3 = 2.9365$	-0.1204	159.8629
9.	0.05	0.6	$q_4 = --$	–	–	39.	0.65	0.6	$q_4 = --$	–	–
10.	0.05	0.6	$q_5 = --$	–	–	40.	0.65	0.6	$q_5 = --$	–	–
11.	0.05	0.9	$q_1 = 1.0121$	0.0964	19.8288	41.	0.65	0.9	$q_1 = 1.1090$	0.2370	31.2449
12.	0.05	0.9	$q_2 = 1.9856$	0.0578	47.2592	42	0.65	0.9	$q_2 = 1.8745$	0.0271	51.1854
13.	0.05	0.9	$q_3 = 2.9886$	-0.5703	1526.2520	43.	0.65	0.9	$q_3 = 2.9742$	-0.6875	6756.3080
14.	0.05	0.9	$q_4 = --$	–	–	44.	0.65	0.9	$q_4 = --$	–	–
15.	0.05	0.9	$q_5 = --$	–	–	45.	0.65	0.9	$q_5 = --$	–	–
16.	0.35	0.3	$q_1 = 1.0773$	0.1743	26.7770	46.	0.95	0.3	$q_1 = 1.1050$	0.2707	35.7252
17.	0.35	0.3	$q_2 = 1.9714$	0.0448	49.6382	47.	0.95	0.3	$q_2 = 1.9599$	0.0357	51.0935
18.	0.35	0.3	$q_3 = 2.9198$	-0.0364	15.4016	48.	0.95	0.3	$q_3 = 2.8847$	-0.0383	20.5160
19.	0.35	0.3	$q_4 = --$	–	–	49.	0.95	0.3	$q_4 = --$	–	–
20.	0.35	0.3	$q_5 = --$	–	–	50.	0.95	0.3	$q_5 = --$	–	–
21.	0.35	0.6	$q_1 = 1.0682$	0.1676	25.8888	51.	0.95	0.6	$q_1 = 1.2370$	0.3544	42.4224
22.	0.35	0.6	$q_2 = 1.9445$	0.0444	49.5347	52.	0.95	0.6	$q_2 = 1.9000$	0.0106	53.5250
23.	0.35	0.6	$q_3 = 2.9553$	-0.1174	87.4308	53.	0.95	0.6	$q_3 = 2.9172$	-0.1041	258.4500
24.	0.35	0.6	$q_4 = --$	–	–	54.	0.95	0.6	$q_4 = --$	–	–
25.	0.35	0.6	$q_5 = --$	–	–	55.	0.95	0.6	$q_5 = --$	–	–
26.	0.35	0.9	$q_1 = 1.0690$	0.1609	25.0245	56.	0.95	0.9	$q_1 = 1.1397$	0.3251	38.4883
27.	0.35	0.9	$q_2 = 1.9950$	0.0439	49.4185	57.	0.95	0.9	$q_2 = 1.8420$	0.0080	52.5609
28.	0.35	0.0.9	$q_3 = 2.9970$	-0.6871	3462.7230	58.	095	0.9	$q_3 = 2.9819$	-0.6017	9656.5140
29.	0.35	0.9	$q_4 = --$	–	–	59.	0.95	0.9	$q_4 = --$	–	–
30.	0.35	0.9	$q_5 = --$	–	–	60.	0.95	0.9	$q_5 = --$	–	–

Tables 4.8 Ready Reckoner for Choice of q Providing Optimum MSE for given (M,g) (Using corollary 2, eq. (20)) [$Range\ 0.05 \le M \le 0.95;\ Range\ 0.3 \le g \le 0.9$]

S.No.	M	g	Choice of q	Bias	MSE	S.No.	M	g	Choice of q	Bias	MSE
1.	0.05	0.3	$q_1 = 1.8017$	-0.0320	15.3031	31.	0.65	0.3	$q_1 = 0.8497$	0.0219	14.1011
2.	0.05	0.3	$q_2 = 3.0989$	-0.0036	15.3001	32.	0.65	0.3	$q_2 = 1.8150$	-0.5915	14.1012
3.	0.05	0.3	$q_3 = 4.1930$	-0.0024	15.3008	33.	0.65	0.3	$q_3 = 2.9500$	-0.0336	14.1060
4.	0.05	0.3	$q_4 = --$	–	–	34.	0.65	0.3	$q_4 = --$	–	–
5.	0.05	0.3	$q_5 = --$	–	–	35.	0.65	0.3	$q_5 = --$	–	–
6.	0.05	0.6	$q_1 = 1.6810$	-0.2043	15.3018	36.	0.65	0.6	$q_1 = 0.8501$	0.0207	14.1020
7.	0.05	0.6	$q_2 = 3.6084$	-0.0036	15.3010	37.	0.65	0.6	$q_2 = 1.6905$	-2.7535	14.1026
8.	0.05	0.6	$q_3 = 4.3269$	-0.0019	15.3011	38.	0.65	0.6	$q_3 = 3.1127$	-0.0457	14.1052
9.	0.05	0.6	$q_4 = --$	–	–	39.	0.65	0.6	$q_4 = --$	–	–
10.	0.05	0.6	$q_5 = --$	–	–	40.	0.65	0.6	$q_5 = --$	–	–
11.	0.05	0.9	$q_1 = 1.5999$	0.0598	15.3012	41.	0.65	0.9	$q_1 = 0.2017$	0.0194	14.1020
12.	0.05	0.9	$q_2 = 3.7734$	-0.0034	15.3010	42.	0.65	0.9	$q_2 = 1.5951$	1.01378	14.1058
13.	0.05	0.9	$q_3 = 4.5800$	-0.0012	15.3010	43.	0.65	0.9	$q_3 = 3.2500$	-0.0502	14.1053
14.	0.05	0.9	$q_4 = --$	–	–	44.	0.65	0.9	$q_4 = --$	–	–
15.	0.05	0.9	$q_5 = --$	–	–	45.	0.65	0.9	$q_5 = --$	–	–
16.	0.35	0.3	$q_1 = 0.5400$	-0.0036	13.1652	46.	0.95	0.3	$q_1 = 0.9851$	0.0790	18.0995
17.	0.35	0.3	$q_2 = 1.8105$	-0.2723	13.1612	47.	0.95	0.3	$q_2 = 1.8176$	-0.9932	18.0961
18.	0.35	0.3	$q_3 = 3.0311$	-0.0219	13.1648	48.	0.95	0.3	$q_3 = 2.8670$	-0.0383	18.0973
19.	0.35	0.3	$q_4 = --$	–	–	49.	0.95	0.3	$q_4 = --$	–	–
20.	0.35	0.3	$q_5 = --$	–	–	50.	0.95	0.3	$q_5 = --$	–	–
21.	0.35	0.6	$q_1 = 0.5060$	-0.0043	13.1651	51.	0.95	0.6	$q_1 = 0.9840$	0.0748	18.0941
22.	0.35	0.6	$q_2 = 1.6875$	-1.2295	13.1600	52.	0.95	0.6	$q_2 = 1.6818$	-4.4638	18.0933
23.	0.35	0.6	$q_3 = 3.2101$	-0.0257	13.1651	53.	0.95	0.6	$q_3 = 3.0461$	-0.0635	18.0941
24.	0.35	0.6	$q_4 = --$	–	–	54.	0.95	0.6	$q_4 = --$	–	–
25.	0.35	0.6	$q_5 = --$	–	–	55.	0.95	0.6	$q_5 = --$	–	–
26.	0.35	0.9	$q_1 = 0.4718$	-0.0049	13.1653	56.	0.95	0.9	$q_1 = 0.9732$	0.0744	18.0995
27.	0.35	0.9	$q_2 = 1.5101$	0.4831	13.1601	57.	0.95	0.9	$q_2 = 1.5816$	1.6632	18.0925
28.	0.35	0.0.9	$q_3 = 3.3773$	-0.0267	13.1649	58.	095	0.9	$q_3 = 3.1741$	-0.0733	18.09847
29.	0.35	0.9	$q_4 = --$	–	–	59.	0.95	0.9	$q_4 = --$	–	–
30.	0.35	0.9	$q_5 = --$	–	–	60.	0.95	0.9	$q_5 = --$	–	–

5. Confidence Interval Estimation and Imputation

Consider the 10 random samples $A_1, A_2, A_3, ..., A_{10}$ each of size n=40 from population N=150 (from table 4.2). Description of samples is in table 5.1 given below:

Table 5.1 Ten Random Sample Selection

Sample No.	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}
(z_1, u_1)	(117,145)	(156,156)	(120,100)	(111,129)	(118,174)	(85,94)	(117,135)	(176,108)	(91,239)	(167,127)
(z_2, u_2)	(106,97)	(107,127)	(132,64)	(120,,100)	(85,94)	(142,119)	(106,97)	(142,119)	(114,63)	(133,100)
(z_3, u_3)	(132,64)	(142,119)	(113,64)	(113,64)	(101,101)	(133,90)	(113,64)	(131,110)	(167,127)	(178,141)
(z_4, u_4)	(61,117)	(131,110)	(61,117)	(128,137)	(170,123)	(131,110)	(128,137)	(178,141)	(167,153)	(126,139)
(z_5, u_5)	(81,126)	(178,141)	(137,149)	(127,101)	(116,99)	(178,141)	(128,81)	(92,75)	(132,156)	(160,177)
(z_6, u_6)	(95,106)	(115,128)	(81,126)	(119,105)	(111,116)	(115,128)	(103,114)	(161,139)	(178,141)	(143,152)
(z_7, u_7)	(128,81)	(160,117)	(126,77)	(141,147)	(140,156)	(134,129)	(138,109)	(122,94)	(92,75)	(111,116)
(z_8, u_8)	(103,114)	(143,152)	(119,105)	(159,95)	(47,110)	(174,109)	(84,119)	(111,116)	(170,123)	(101,137)
(z_9, u_9)	(138,109)	(49,92)	(128,81)	(99,77)	(134,129)	(66,84)	(159,95)	(102,126)	(111,116)	(122,121)
(z_{10}, u_{10})	(62,56)	(134,129)	(141,147)	(131,118)	(135,106)	(109,130)	(147,119)	(135,166)	(47,110)	(109,130)
(z_{11}, u_{11})	(78,69)	(123,131)	(138,109)	(97,85)	(123,131)	(121,89)	(164,137)	(143,141)	(134,129)	(111,99)
(z_{12}, u_{12})	(100,136)	(66,84)	(135,157)	(104,129)	(122,121)	(116,64)	(985)	(109,136)	(143,141)	(121,89)
(z_{13}, u_{13})	(99,77)	(122,121)	(78,69)	(112,65)	(101,129)	(131,95)	(100,117)	(116,64)	(109,130)	(116,64)
(z_{14}, u_{14})	(103,174)	(87,67)	(100,135)	(120,160)	(104,97)	(118,128)	(131,99)	(95,109)	(121,89)	(131,95)
(z_{15}, u_{15})	(131,118)	(111,99)	(84,119)	(139,162)	(140,105)	(140,105)	(97,85)	(106,136)	(95,109)	(118,128)
(z_{16}, u_{16})	(164,137)	(105,123)	(71,139)	(109,118)	(178,152)	(95,98)	(100,117)	(95,98)	(140,105)	(65,87)
(z_{17}, u_{17})	(111,163)	(178,152)	(97,85)	(99,149)	(117,92)	(117,92)	(131,99)	(164,172)	(104,72)	(140,105)
(z_{18}, u_{18})	(112,65)	(113,123)	(100,117)	(131,140)	(113,123)	(96,65)	(97,85)	(125,98)	(95,98)	(104,72)
(z_{19}, u_{19})	(120,166)	(125,98)	(111,136)	(158,151)	(106,136)	(137,152)	(104,129)	(115,109)	(117,92)	(178,152)
(z_{20}, u_{20})	(139,162)	(75,121)	(11,99)	(106,59)	(140,105)	(125,98)	(73,59)	(189,109)	(158,86)	(103,119)
(z_{21}, u_{21})	(109,118)	(112,125)	(84,113)	(94,117)	(118,128)	(112,125)	(99,107)	(75,81)	(121,142)	(96,65)
(z_{22}, u_{22})	(92,140)	(101,94)	(149,140)	(143,148)	(116,64)	(106,59)	(112,65)	(139,162)	(119,103)	(158,86)
(z_{23}, u_{23})	(75,81)	106,59)	(136,175)	(137,152)	(109,130)	(158,151)	(139,162)	(157,148)	(112,125)	(121,142)
(z_{24}, u_{24})	(99,149)	(158,151)	(97,155)	(164,172)	(101,129)	(131,140)	(158,137)	(100,117)	(131,140)	(103,119)
(z_{25}, u_{25})	(131,140)	(131,140)	(131,175)	(103,119)	(143,141)	(136,17)	(97,155)	(95,129)	(99,149)	(96,65)
(z_{26}, u_{26})	(158,151)	(99,149)	(106,59)	(113,123)	(122,121)	(140,140)	(132,117)	(147,119)	(92,140)	(158,86)
(z_{27}, u_{27})	(94,117)	(75,81)	(101,94)	(106,136)	(101,137)	(84,113)	(121,142)	(99,77)	(139,162)	(121,142)
(z_{28}, u_{28})	(141,98)	(92,140)	(141,98)	(105,94)	(123,131)	(91,120)	(164,172)	(100,136)	(157,148)	(119,103)
(z_{29}, u_{29})	(164,172)	(109,118)	(143,148)	(66,84)	(134,129)	(131,99)	(103,119)	(78,69)	(131,99)	(101,94)
(z_{30}, u_{30})	(96,65)	(56,87)	(121,142)	(174,109)	(47,110)	(70,65)	(140,105)	(135,157)	(73,59)	(131,157)
(z_{31}, u_{31})	(103,119)	(84,55)	(137,152)	(135,106)	(49,92)	(95,103)	(118,128)	(103,114)	(104,129)	(131,140)
(z_{32}, u_{32})	(178,152)	(157,148)	(103,119)	(124,99)	(132,117)	(164,137)	(111,99)	(128,81)	(95,103)	(92,140)
(z_{33}, u_{33})	(140,105)	(91,120)	(104,72)	(140,156)	(161,139)	(159,95)	(140,105)	(95,106)	(147,119)	(56,87)
(z_{34}, u_{34})	(118,128)	(73,59)	(140,105)	(132,117)	(92,75)	(78,69)	(118,128)	(127,101)	(109,117)	(139,162)
(z_{35}, u_{35})	(121,89)	(104,129)	(118,128)	(106,124)	(131,110)	(69.117)	(111,99)	(61,117)	(107,59)	(84,113)
(z_{36}, u_{36})	(123,131)	(95,103)	(122,121)	(161,139)	(167,153)	(141,147)	(66,84)	(106,117)	(127,101)	(120,166)
(z_{37}, u_{37})	(102,126)	(71,139)	(174,109)	(92,75)	(176,108)	(169,112)	(174,109)	(106,120)	(128,137)	(128,81)
(z_{38}, u_{38})	(106,124)	(103,174)	(135,106)	(178,141)	(75,81)	(105,133)	(135,106)	(117,135)	(113,64)	(104,129)
(z_{39}, u_{39})	(170,123)	(100,136)	(47,110)	(131,110)	(122,121)	(126,77)	(124,99)	(120,100)	(106,97)	(164,137)
(z_{40}, u_{40})	(85,94)	(115,139)	(170,123)	(176,108)	(160,117)	(61,117)	(102,126)	(81,126)	(120,100)	(115,139)

Table 5.2 Ten Sample Descriptive Statistic [eq. (1) to eq. (6)]

Sample No.	Mean ($\bar{z}$)	Mean($\bar{u}$)	s_z	s_u	c_z	c_u	ρ_{zu}
A_1	114.6250	119.6750	28.2822	33.2145	0.2467	0.2775	0.0165
A_2	112.6250	119.1500	33.2663	30.2329	0.2953	0.2537	0.0479
A_3	116.5500	116.6500	25.8069	31.0306	0.2214	0.2660	0.0274
A_4	125.0750	118.0000	25.4995	29.1196	0.2038	0.24167	0.0732
A_5	121.1500	119.8250	32.2566	22.2767	0.2662	0.1859	0.0437
A_6	119.7256	110.3500	30.3273	27.5249	0.2533	0.2494	0.0363
A_7	120.4000	113.6000	25.8852	28.5359	0.2149	0.2512	0.0303
A_8	118.3000	115.4250	30.7531	27.8815	0.2599	0.2415	0.1967
A_9	124.6250	113.6750	33.5954	28.4906	0.2695	0.2506	0.1885
A_{10}	118.7000	118.6500	28.5401	29.2991	0.2404	0.2469	0.1476

The table 5.2 presents the descriptive statistics of mean, variability and correlation of ten samples using eq. (1) to eq. (6).

5.1. Definition of Confidence Interval (CI)

Suppose $\bar{\theta_n}$ be an unbiased estimator of unknown θ based on random sample n from normal population $N[\theta, \sigma^2]$. Then 95% confidence interval is defined as:

$$P[\bar{\theta} - 1.96\sqrt{var(\bar{\theta})} < \theta < \bar{\theta} + 1.96\sqrt{var(\bar{\theta})}] = 0.95$$

where P[.] denotes the probability of event. The lower limit of confidence interval (CI) is LL= $[\bar{\theta} - 1.96\sqrt{var(\bar{\theta})}]$ and upper limit is UL=$[\bar{\theta} + 1.96\sqrt{var(\bar{\theta})}]$. As interpretation, there exists 95% chance that true but unknwon θ lies between lower limit and upper limit of confidence interval (CI). Deriving motivation from this, for biased estimator, two proposed limits are:

$$(LL)_{opt} = Lower\ Limit\ = [estimated\ mean\ - 1.96\sqrt{est(MSE)_{(q_{opt})}}] \tag{20}$$

$$(UL)_{opt} = Upper\ Limit\ = [estimated\ mean\ + 1.96\sqrt{est(MSE)_{(q_{opt})}}] \tag{21}$$

Table 5.3 Estimated Confidence Intervals Over 10 Samples at the q_{opt} Values [using (21) and (22)]

Sample No.	q_{opt}	E	est(MSE)	C.I. $[(LL)_{opt}, (UL)_{opt}]$	Length
A_1	$q_{1(opt)} = 0.6335$	112.7530	17.5324	[78.38,147.11]	68.7270
A_2	$q_{1(opt)} = 0.6335$	110.9900	22.1005	[67.67,154.30]	86.6338
A_3	$q_{1(opt)} = 0.6335$	115.8710	14.8170	[86.82,144.91]	58.0828
A_4	$q_{1(opt)} = 0.6335$	123.7570	13.9681	[96.37,151.13]	54.7549
A_5	$q_{1(opt)} = 0.6335$	119.1090	20.1490	[79.61,158.60]	78.9839
A_6	$q_{1(opt)} = 0.6335$	121.7330	19.0932	[84.31,159.15]	74.8452
A_7	$q_{1(opt)} = 0.6335$	121.0010	14.7219	[92.14,149.85]	57.7099
A_8	$q_{1(opt)} = 0.6335$	118.1220	17.2456	[84.32,151.92]	67.6072
A_9	$q_{1(opt)} = 0.6335$	125.2140	20.7113	[84.61,165.80]	81.1885
A_{10}	$q_{1(opt)} = 0.6335$	117.1810	15.7098	[86.39,147.97]	61.5823
Avg Length of CI				[84.06,153.07]	69.0115
A_1	$q_{2(opt)} = 1.8270$	112.2777	17.5233	[77.93,146.62]	68.6913
A_2	$q_{2(opt)} = 1.8270$	110.6306	22.0938	[67.32,153.93]	86.6080
A_3	$q_{2(opt)} = 1.8270$	115.8193	14.8085	[86.79,144.84]	58.0496
A_4	$q_{2(opt)} = 1.8270$	123.5596	13.9605	[96.19,150.92]	54.7252
A_5	$q_{2(opt)} = 1.8270$	118.5700	20.449	[79.08,158.05]	78.9682
A_6	$q_{2(opt)} = 1.8270$	121.3830	19.0856	[83.97,158.79]	74.8156
A_7	$q_{2(opt)} = 1.8270$	120.9650	14.7139	[92.19,149.80]	57.6786
A_8	$q_{2(opt)} = 1.8270$	118.1187	17.2418	[84.32,151.91]	67.5882
A_9	$q_{2(opt)} = 1.8270$	125.1801	20.7067	[84.59,165.76]	81.1705
A_{10}	$q_{2(opt)} = 1.8270$	116.8956	15.7045	[86.11,147.67]	61.5619
Avg Length of CI				[83.84,15282]	68.9857
A_1	$q_{3(opt)} = 2.9830$	112.7620	17.5268	[78.37,147.07]	68.0752
A_2	$q_{3(opt)} = 2.9830$	110.9680	22.0964	[67.65,154.27]	86.6180
A_3	$q_{3(opt)} = 2.9830$	115.8680	14.8119	[86.83,144.89]	58.0625
A_4	$q_{3(opt)} = 2.9830$	123.7450	13.9635	[96.37,151.11]	54.7300
A_5	$q_{3(opt)} = 2.9830$	119.0780	20.1465	[79.59,158.56]	78.9743
A_6	$q_{3(opt)} = 2.9830$	121.7000	19.0886	[84.28,159.11]	74.8271
A_7	$q_{3(opt)} = 2.9830$	120.9980	14.7170	[92.15,149.84]	57.6908
A_8	$q_{3(opt)} = 2.9830$	118.1220	17.2433	[84.32,151.91]	67.5938
A_9	$q_{3(opt)} = 2.9830$	125.2110	20.7085	[84.62,165.79]	81.1775
A_{10}	$q_{3(opt)} = 2.9830$	117.1640	15.7066	[86.37,147.94]	61.5698
Avg. Length of CI				[84.05,152.95]	68.9319

6. Application of Confidence Interval For Missing Value Imputation

6.1. Proposed CI- Imputation Procedure

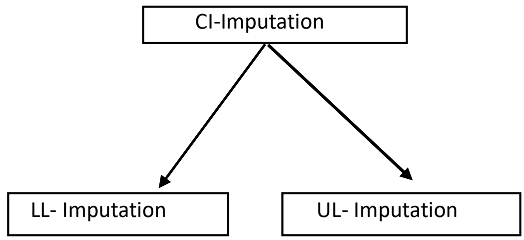


Fig 6.1 Type of CI-Imputation

Let random sample of size n drawn from N $(n < N)$ has only one missing value. The value imputed through pattern procedure is assumed available and missing one in sample is other than that. The proposed CI- imputation procedure is as under:

Step I: Find mean of sample of size (n-1) eliminating missing observation.

Step II: Calculate mean and MSE of sample data by suggested estimation method (eliminating missing value).

Step III: Calculate lower limit $(LL)_{opt} = [estimated\ mean - 1.96\sqrt{MSE(estimated\ mean)q_{opt}}]$ which is termed as LL-Imputation.

Moreover, calculate upper limit $(UL)_{opt} = [estimated\ mean + 1.96\sqrt{MSE(estimated\ mean)q_{opt}}]$ which is termed as UL-Imputation.

Step IV: Use $(LL)_{opt}$ or $(UL)_{opt}$ for imputing the missing value in sample.

Step V: Repeat the procedure over multiple random samples and average out the estimated mean value with imputation.

Note 6.1 CI-Imputaion seems logically better since it incorporates both mean and MSE in $(LL)_{opt}$ or $(UL)_{opt}$ while sample mean imputation of (n-1) observations does not incorporate variability information.

Table 6.1 Sample With one Missing Value

Sample No.	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	Remark
(z_1,u_1)	(117,145)	(156,156)	(120,100)	(111,129)	(118,174)	(85,94)	(117,135)	(176,108)	(91,239)	(167,127)	
(z_2,u_2)	(106,97)	(107,127)	(132,64)	(120,,100)	(85,94)	(142,119)	(106,97)	(142,119)	(114,63)	(133,100)	
(z_3,u_3)	(132,64)	(142,119)	(113,64)	(113,64)	(101,101)	(133,90)	(113,64)	(131,110)	(167,127)	(178,141)	
(z_4,u_4)	(61,117)	(131,110)	(61,117)	(128,137)	(170,123)	(131,110)	(128,137)	(178,141)	(167,153)	(126,139)	
(z_5,u_5)	(81,126)	(178,141)	(137,149)	(127,101)	(116,99)	(178,141)	(128,81)	(92,75)	(132,156)	(160,177)	
(z_6,u_6)	(95,106)	(115,128)	(81,126)	(119,105)	(111,116)	(115,128)	(103,114)	(161,139)	(178,141)	(143,152)	
(z_7,u_7)	(128,81)	(160,117)	(126,77)	(141,147)	(140,156)	(134,129)	(138,109)	(122,94)	(92,75)	(111,116)	
(z_8,u_8)	(103,114)	(143,152)	(119,105)	(159,95)	(47,110)	(174,109)	(84,119)	(111,116)	(170,123)	(101,137)	
(z_9,u_9)	(138,109)	(49,92)	(128,81)	(99,77)	(134,129)	(66,84)	(159,95)	(102,126)	(111,116)	(122,121)	
(z_{10},u_{10})	(62,56)	(134,129)	(141,147)	(131,118)	(135,106)	(109,130)	(147,119)	(135,166)	(47,110)	(109,130)	
(z_{11},u_{11})	(78,69)	(123,131)	(138,109)	(97,85)	(123,131)	(121,89)	(164,137)	(143,141)	(134,129)	(111,99)	
(z_{12},u_{12})	(100,136)	(66,84)	(135,157)	(104,129)	(122,121)	(116,64)	(985)	(109,136)	(143,141)	(121,89)	
(z_{13},u_{13})	(99,77)	(122,121)	(78,69)	(112,65)	(101,129)	(131,95)	(100,117)	(116,64)	(109,130)	(116,64)	
(z_{14},u_{14})	(103,174)	(87,67)	(100,135)	(120,160)	(104,97)	(118,128)	(131,99)	(95,109)	(121,89)	(131,95)	
(z_{15},u_{15})	(131,118)	(111,99)	(84,119)	(139,162)	(140,105)	(140,105)	(97,85)	(106,136)	(95,109)	(118,128)	
$\rightarrow (z_{16},u_{16})$	(–,137)	(–,123)	(–,139)	(–,118)	(–,152)	(–,98)	(–,117)	(–,98)	(–,105)	(–,87))	Missing Values
(z_{17},u_{17})	(111,163)	(178,152)	(97,85)	(99,149)	(117,92)	(117,92)	(131,99)	(164,172)	(104,72)	(140,105)	
(z_{18},u_{18})	(112,65)	(113,123)	(100,117)	(131,140)	(113,123)	(96,65)	(97,85)	(125,98)	(95,98)	(104,72)	
(z_{19},u_{19})	(120,166)	(125,98)	(111,136)	(158,151)	(106,136)	(137,152)	(104,129)	(115,109)	(117,92)	(178,152)	
(z_{20},u_{20})	(139,162)	(75,121)	(11,99)	(106,59)	(140,105)	(125,98)	(73,59)	(189,109)	(158,86)	(103,119)	
(z_{21},u_{21})	(109,118)	(112,125)	(84,113)	(94,117)	(118,128)	(112,125)	(99,107)	(75,81)	(121,142)	(96,65)	
(z_{22},u_{22})	(92,140)	(101,94)	(149,140)	(143,148)	(116,64)	(106,59)	(112,65)	(139,162)	(119,103)	(158,86)	
(z_{23},u_{23})	(75,81)	106,59)	(136,175)	(137,152)	(109,130)	(158,151)	(139,162)	(157,148)	(112,125)	(121,142)	
(z_{24},u_{24})	(99,149)	(158,151)	(97,155)	(164,172)	(101,129)	(131,140)	(158,137)	(100,117)	(131,140)	(103,119)	
(z_{25},u_{25})	(131,140)	(131,140)	(131,175)	(103,119)	(143,141)	(136,17)	(97,155)	(95,129)	(99,149)	(96,65)	
(z_{26},u_{26})	(158,151)	(99,149)	(106,59)	(113,123)	(122,121)	(140,140)	(132,117)	(147,119)	(92,140)	(158,86)	
(z_{27},u_{27})	(94,117)	(75,81)	(101,94)	(106,136)	(101,137)	(84,113)	(121,142)	(99,77)	(139,162)	(121,142)	
(z_{28},u_{28})	(141,98)	(92,140)	(141,98)	(105,94)	(123,131)	(91,120)	(164,172)	(100,136)	(157,148)	(119,103)	
(z_{29},u_{29})	(164,172)	(109,118)	(143,148)	(66,84)	(134,129)	(131,99)	(103,119)	(78,69)	(131,99)	(101,94)	
(z_{30},u_{30})	(96,65)	(56,87)	(121,142)	(174,109)	(47,110)	(70,65)	(140,105)	(135,157)	(73,59)	(131,157)	
(z_{31},u_{31})	(103,119)	(84,55)	(137,152)	(135,106)	(49,92)	(95,103)	(118,128)	(103,114)	(104,129)	(131,140)	
(z_{32},u_{32})	(178,152)	(157,148)	(103,119)	(124,99)	(132,117)	(164,137)	(111,99)	(128,81)	(95,103)	(92,140)	
(z_{33},u_{33})	(140,105)	(91,120)	(104,72)	(140,156)	(161,139)	(159,95)	(140,105)	(95,106)	(147,119)	(56,87)	
(z_{34},u_{34})	(118,128)	(73,59)	(140,105)	(132,117)	(92,75)	(78,69)	(118,128)	(127,101)	(109,117)	(139,162)	
(z_{35},u_{35})	(121,89)	(104,129)	(118,128)	(106,124)	(131,110)	(69.117)	(111,99)	(61,117)	(107,59)	(84,113)	
(z_{36},u_{36})	(123,131)	(95,103)	(122,121)	(161,139)	(167,153)	(141,147)	(66,84)	(106,117)	(127,101)	(120,166)	
(z_{37},u_{37})	(102,126)	(71,139)	(174,109)	(92,75)	(176,108)	(169,112)	(174,109)	(106,120)	(128,137)	(128,81)	
(z_{38},u_{38})	(106,124)	(103,174)	(135,106)	(178,141)	(75,81)	(105,133)	(135,106)	(117,135)	(113,64)	(104,129)	
(z_{39},u_{39})	(170,123)	(100,136)	(47,110)	(131,110)	(122,121)	(126,77)	(124,99)	(120,100)	(106,97)	(164,137)	
(z_{40},u_{40})	(85,94)	(115,139)	(170,123)	(176,108)	(160,117)	(61,117)	(102,126)	(81,126)	(120,100)	(115,139)	

Table 6.1 represents the ten samples as in table 5.1 but 16^{th} value is assumed missing in each sample.

Table 6.2 Sample Statistic Excluding Missing Value (table 6.1) for (n-1) observations

Sample No.	Mean ($\bar{z}$)	Mean($\bar{u}$)	s_z	s_u	c_z	c_u	ρ_{zu}
A_1	113.3590	119.6750	32.6380	33.2145	0.2879	0.2775	0.1992
A_2	112.0510	119.1500	36.5645	30.2329	0.3263	0.2537	0.5064
A_3	116.0150	116.6500	32.5492	31.0306	0.2805	0.2660	0.0971
A_4	125.0020	118.0000	32.3158	29.1196	0.2585	0.24167	0.3446
A_5	119.1500	119.8250	36.6034	22.2767	0.3072	0.1859	0.1487
A_6	120.3590	110.3500	35.7087	27.5249	0.2966	0.2494	0.3820
A_7	118.2500	113.6000	32.0423	28.5359	0.2709	0.2512	0.3115
A_8	118.1030	115.4250	35.9131	27.8815	0.3040	0.2415	0.4420
A_9	124.2310	113.6750	39.9801	28.4906	0.3218	0.2506	0.3954
A_{10}	117.3590	118.6500	28.8045	29.2991	0.2198	0.2469	0.0474

The tabele 6.2 reveals decriptive statistic of ten samples in terms of mean, variability and sample correlation when one value is missing.

Table 6.3 Estimated Confidence Intervals Over 10 Samples at the q_{opt} Excluding Missing Value (for (n-1) observations)

Sample No.	q_{opt}	E	est(MSE)	C.I.	Length
A_1	$q_{1(opt)} = 0.6335$	111.508	19.4906	[78.30,149.40]	76.403
A_2	$q_{1(opt)} = 0.6335$	110.4240	19.0999	[72.98,147.86]	74.8718
A_3	$q_{1(opt)} = 0.6335$	115.34000	25.9601	[74.42,156.25]	81.8199
A_4	$q_{1(opt)} = 0.6335$	123.6840	16.9108	[90.53,156.82]	66.2905
A_5	$q_{1(opt)} = 0.6335$	117.1430	24.2574	[69.59,164.68]	95.0888
A_6	$q_{1(opt)} = 0.6335$	122.3780	20.0005	[83.17,161.57]	78.4021
A_7	$q_{1(opt)} = 0.6335$	118.8400	17.0798	[85.36,152.31]	66.9500
A_8	$q_{1(opt)} = 0.6335$	117.9250	19.3529	[79.99,155.85]	75.8633
A_9	$q_{1(opt)} = 0.6335$	124.8180	24.8983	[76.01,173.61]	97.6012
A_{10}	$q_{1(opt)} = 0.6335$	115.8930	14.2437	[87.97,143.81]	55.8355
			Average	[79.83,156.21]	76.816
A_1	$q_{2(opt)} = 1.8270$	111.0377	19.4860	[72.84,149.23]	76.3852
A_2	$q_{2(opt)} = 1.8270$	110.0667	19.1044	[72.62,147.51]	74.8894
A_3	$q_{2(opt)} = 1.8270$	115.2877	20.8657	[74.39,156.18]	81.7938
A_4	$q_{2(opt)} = 1.8270$	123.4870	16.9098	[90.34,156.63]	66.2865
A_5	$q_{2(opt)} = 1.8270$	116.6126	24.2555	[69.07,164.15]	95.0817
A_6	$q_{2(opt)} = 1.8270$	122.0257	20.0014	[82.82,161.22]	78.4058
A_7	$q_{2(opt)} = 1.8270$	118.8049	17.0783	[85.33,152.27]	66.9471
A_8	$q_{2(opt)} = 1.8270$	117.9215	19.3556	[79.98,155.85]	75.8741
A_9	$q_{2(opt)} = 1.8270$	124.7844	24.9004	[75.97,173.58]	97.6098
A_{10}	$q_{2(opt)} = 1.8270$	115.6105	14.2367	[87.70,143.51]	55.8079
			Average	[79.10,156.03]	76.8681
A_1	$q_{3(opt)} = 2.9830$	111.4810	19.4878	[73.28,149.67]	76.3922
A_2	$q_{3(opt)} = 2.9830$	110.4030	19.1027	[72.96,147.84]	74.8820
A_3	$q_{3(opt)} = 2.9830$	115.3360	20.8684	[74.43,156.23]	81.8040
A_4	$q_{3(opt)} = 2.9830$	123.6720	16.9102	[90.52,156.81]	66.2880
A_5	$q_{3(opt)} = 2.9830$	117.1130	24.2562	[69.57,164.65]	95.0845
A_6	$q_{3(opt)} = 2.9830$	122.3450	20.0011	[83.14,161.54]	78.4044
A_7	$q_{3(opt)} = 2.9830$	118.8370	17.0789	[85.36,152.31]	66.9494
A_8	$q_{3(opt)} = 2.9830$	117.9250	19.3546	[79.98,155.85]	75.8699
A_9	$q_{3(opt)} = 2.9830$	124.8150	24.8996	[76.01,173.79]	97.6065
A_{10}	$q_{3(opt)} = 2.9830$	115.8760	14.2394	[87.96,143.78]	55.8186
			Average	[79.31,156.24]	76.9099

Table 6.3 displays the three optimum choices of q over 10 samples along with sample estimates (at opt q) opt MSE and optimum length of CI. Different q_{opt} showing the similar length of CI.

6.2. CI-Imputation Using Lower Limit (LL-Imputation)

In tables 6.4, 6.5 and 6.6, the CI-Imputation with $(LL)_{opt}$ is attempted against missing value as sample in table 6.1

Table 6.4 Sample Where Missing Value Replaced by LL-Imptation (table 6.1)

Sample No.	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	Remark
$\rightarrow (z_{16}, u_{16})$	(–,137)	(–,123)	(–,139)	(–,118)	(–,152)	(–,98)	(–,117)	(–,98)	(–,105)	(–,87))	16^{th} Missing values in table 6.1
$\rightarrow (z_{16}, u_{16})$	(78.30,137)	(72.98,123)	(74.42,139)	(90.53,118)	(69.59,152)	(83.17,98)	(85.36,117)	(79.99,98)	(76.01,105)	(87.97,87)	imputed 16^{th} values by LL in table 6.1

Table 6.4 shows value under LL-Imputation against 16^{th} missing value of table 6.1.

Table 6.5 Sample Statistic When Missing Value replaced by LL-imputation in 10 Samples of table 6.4

Sample No.	Mean ($\bar{z}$)	Mean($\bar{u}$)	s_z	s_u	c_z	c_u	ρ_{zu}
A_1	112.3575	119.6750	27.5043	33.2145	0.2447	0.2775	0.2653
A_2	111.0510	119.1500	32.0446	30.2329	0.2884	0.2537	0.5707
A_3	116.0100	116.6500	27.1908	31.0306	0.2331	0.2660	0.1626
A_4	124.6133	118.0000	25.6345	29.1196	0.2057	0.24167	0.4235
A_5	118.4398	119.8250	31.5086	22.2767	0.2660	0.1859	0.2481487
A_6	119.4293	110.3500	30.2454	27.5249	0.2532	0.2494	0.4089
A_7	118.2400	113.6000	26.0275	28.5359	0.2181	0.2512	0.3837
A_8	117.9248	115.4250	30.7422	27.8815	0.2606	0.2415	0.4628
A_9	123.0253	113.6750	33.9271	28.4906	0.2757	0.2506	0.4273
A_{10}	117.3180	118.6500	26.0493	29.2991	0.2098	0.2469	0.0991

The table 6.5 is obtained by using data of table 6.4 after LL-imputation.

Table 6.6 Estimated Confidence Intervals Over 10 Samples at the q_{opt} When Missing Value Replaced by LL-Imputation (at optimum q=0.6335)

Sample No.	q_{opt}	E	est(MSE)	C.I.	Length
A_1	$q_{1(opt)} = 0.6335$	110.5280	13.4334	[84.19,136.85]	52.6591
A_2	$q_{1(opt)} = 0.6335$	109.4615	13.5378	[82.92,135.99]	53.0687
A_3	$q_{1(opt)} = 0.6335$	115.9363	14.4416	[87.65,144.26]	56.6113
A_4	$q_{1(opt)} = 0.6335$	123.30010	9.9386	[103.82,142.77]	38.9593
A_5	$q_{1(opt)} = 0.6335$	116.4447	17.1051	[82.91,149.97]	67.0522
A_6	$q_{1(opt)} = 0.6335$	121.4324	13.9675	[94.05,148.80]	54.7527
A_7	$q_{1(opt)} = 0.6335$	118.8400	17.0798	[85.36,152.31]	66.9500
A_8	$q_{1(opt)} = 0.6335$	117.7472	13.7387	[90.81,144.67]	53.8570
A_9	$q_{1(opt)} = 0.6335$	123.6065	17.3161	[89.66,157.54]	67.8793
A_{10}	$q_{1(opt)} = 0.6335$	116.3164	13.9007	[89.07,143.56]	54.4908
			Average	[88.94,144.87]	56.6280

Table 6.6 provides optimum length of confidence intervals under LL-Imputation.

6.3. CI-Imputation by Upper Limit (UL-Imputation)

In tables 6.7, 6.8 and 6.9 the CI-Imputation with $(UL)_{opt}$ is taken into account against missing values related to table 6.1

Table 6.7 Sample in Which Missing Value Replaced by UL-Imputation (table 6.1)

Sample No.	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	Remark
$\rightarrow (z_{16}, u_{16})$	(–,137)	(–,123)	(–,139)	(–,118)	(–,152)	(–,98)	(–,117)	(–,98)	(–,105)	(–,87))	16^{th} Missing values in table 6.1
$\rightarrow (z_{16}, u_{16})$	(149.40,137)	(147.86,123)	(156.25,139)	(156.82,118)	(164.68,152)	(161.57,98)	(152.31,117)	(155.85,98)	(173.61,105)	(143.81,87)	imputed 16^{th} values by UL in table 6.1

Table 6.7 displays the value under UL-Imputation against 16^{th} missing value of table 6.1. The fifth column of table 6.3 showing upper limit is used to replace the missing.

Table 6.8 Sample Statistic After When Missing Value is Replaced by UL-Imputation in 10 Samples (obtained by table 6.7)

Sample No.	Mean ($\bar{z}$)	Mean($\bar{u}$)	s_z	s_u	c_z	c_u	ρ_{zu}
A_1	114.26000	119.6750	27.3685	33.2145	0.2395	0.2775	0.3028
A_2	112.9470	119.1500	31.9515	30.2329	0.2828	0.2537	0.5798
A_3	118.6810	116.6500	27.0155	31.0306	0.2276	0.2660	0.02182
A_4	126.2710	118.0000	25.5201	29.1196	0.2021	0.24167	0.4254
A_5	120.8170	119.8250	31.3200	22.2767	0.2592	0.1859	0.3593
A_6	121.3890	110.3500	30.3722	27.5249	0.2502	0.2494	0.3782
A_7	120.9580	113.6000	25.9448	28.5359	0.2114	0.2512	0.3926
A_8	119.8210	115.4250	30.6834	27.8815	0.2560	0.2415	0.4251
A_9	125.4650	113.6750	33.9678	28.4906	0.2707	0.2506	0.4049
A_{10}	120.3590	118.6500	25.7112	29.2991	0.2135	0.2469	0.0864

The table 6.8 is obtained by using data of table 6.7 after UL-Imputation against 16^{th} value of table 6.1.

Table 6.9 Estimated Confidence Intervals Over 10 Samples at the q_{opt} When Missing Values Replaced by UL-Imputation (at optimum q=0.6335)

Sample No.	q_{opt}	E	est(MSE)	C.I.	Length
A_1	$q_{1(opt)} = 0.6335$	112.3940	12.8696	[87.16,137.61]	50.4488
A_2	$q_{1(opt)} = 0.6335$	111.3070	13.2788	[85.28,137.33]	52.0530
A_3	$q_{1(opt)} = 0.6335$	117.9900	13.6392	[91.25,144.72]	53.4657
A_4	$q_{1(opt)} = 0.6335$	124.9400	9.8421	[105.65,144.23]	38.5811
A_5	$q_{1(opt)} = 0.6335$	118.7820	15.7420	[87.92,149.63]	61.7085
A_6	$q_{1(opt)} = 0.6335$	123.4250	14.5061	[94.99,151.85]	56.8641
A_7	$q_{1(opt)} = 0.6335$	121.5620	10.5279	[100.92,142.19]	41.2695
A_8	$q_{1(opt)} = 0.6335$	119.6410	14.1687	[91.86,147.41]	55.5414
A_9	$q_{1(opt)} = 0.6335$	126.0580	17.7000	[91.36,160.75]	69.3841
A_{10}	$q_{1(opt)} = 0.6335$	118.8550	13.8306	[91.74,145.96]	54.2161
			Average	[92.78,146.16]	53.3532

Table 6.9 reveals optimum length of confidence intervals after UL-Imputation for 16^{th} values of table 6.1.

7. Comparison

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum. One can compare the

LL-Imputation and UL-Impuation mutually using following formula over 10 samples.

$$Percentage\ Relative\ Gain(PRG) = \frac{[Length\ of\ CI\ Under\ (LL)_{opt} - Imputation]}{[Length\ of\ CI\ Under\ (UL)_{opt} - Imputation]} X100$$

Table 7.1 Percentage Relative Gain (PRG)

Sample No.	PRG(%)
A_1	104.3012%
A_2	101.9512%
A_3	105.8833%
A_4	100.9802%
A_5	108.6595%
A_6	106.2809%
A_7	102.2263%
A_8	106.9673%
A_9	104.2353%
A_{10}	100.2826%
Average	104.1767%

$$Percentage\ Relative\ Efficiency(PRE) = \left[\frac{MSE(E)_q - MSE(E)q_{opt}}{MSE(E)_q}\right] X100$$

where q=1, 2, 3, 4, 5

Table 7.2 Percentage Relative Efficiency (PRE) of Proposed Strategy E

S.No.	q	PRE(%)
1	q=1	31.6200%
2	q=2	73.1700%
3	q=3	00.2500%
4	q=4	17.7900%
5	q=5	2.8100%

8. Discussion

The proposed estimation strategy E has constants A, B, C, D who are linked with another single constant $q > 0$. For given data in table 4.2 and population parameters in table 4.3, the most suitable choices of q are in table 4.4 and table 4.5. For given population (M=0.4082, g=0.2666), the proposed estimation strategy is almost unbiased when $q_1 = 1.0756, q_2 = 1.9709$ and $q_3 = 2.9073$. The best is $q = q_3 = 2.9073$ because it reduces MSE also as shown in table 4.4. Likewise, in table 4.5 the choices of q are $q_1 = 0.6335, q_2 = 1.8270, q_3 = 2.9830$ on which the MSE is optimum (minimum). Best option is $q = q_1 = 0.6335$ having the least bias. Overall, for given data in table 4.2, the most suitable q is $q \in (0.6, 3.0)$ producing optimum MSE with least bias.

The general Ready-Reckoner table 4.7 and table 4.8 reveal for any given data where M ranging $M \in (0.05, 0.95)$, g ranging $g \in (0.3, 0.9)$, the best q ranging $q \in (0.6, 4.55)$ for which MSE and bias both are at the lowest level, whatever be positive $M < 1$ and $g < 1$ using the proposed estimation strategy. The simulation results of confidence interval (CI) over 10 samples, each of size n = 40, are in table 5.1 who estimate sample statistic of the proposed over 10 samples as in table 5.2. The calculation of 95% confidence intervals are in table 5.3. All the CI are catching the true mean value of internal+secret economic bond $\bar{Z} = 116$. The length of confidence intervals have extremely minor variations among them.

The proposed strategy E is efficient at q_{opt} compare to other q-values (table 8.2). The table 5.1 presents ten samples each of size fourty and using equations (21) and (22) the confidence intervals limits $(LL)_{opt}$ and $(UL)_{opt}$ are calculated in table 5.3. The critical observation is that confidence intervals are catching the true value $\bar{Z} = 116$ and they are robust for different values of q_{opt}. Predicting confidence intervals using (21) and (22) makes the result independent to the selection of best q_{opt}. Table 6.1 has 16^{th} value missing in terms of Z but U is available. Even after eliminating the missing, remaining (n-1) observations produce confidence intervals containing unknown mean ($\bar{Z} = 116$) but their lengths have variations. A new CI-Imputation is proposed in section 6 having two types strategies as LL-Impuation and UL-Imputation. Both are compared in the convert text. Overall in the ten samples, the UL-Imputation found better (more efficient) than LL-Imputation. Confidence intervals after imputation are close to the before imputation (table 6.6 and table 6.9).

9. Conclusion

On recapitulation, the problem opted is to estimate the average economic inner+secret bond existing between supporting countries involved in the war. Their connectivity is modeled like a generalised Petersen graph, sampling and imputation technique are used as methodological tools. As new methods of imputation, named after "Pattern Imputation and CI- Imputation" are proposed in the content in order to maintain the completeness in the symmetry in view of the sampling strategy implementation. Pattern imputation is found efficient and useful for filling the missing data. An estimation strategy is proposed whose expressions of bias and mean squared error are derived. It has four constants A, B, C and D who are linked with another single constant q having expressions in terms of power five. This has led to the best selection of q for making the proposed estimation strategy optimum with least bias. The most plausible selection of q is $q \in (0.6, 3.0)$ for given M = 0.4082, g = 0.2666. Two Ready-Reckoner tables provide general range of most suitable q as $q \in (0.6, 4.55)$ whatsoever be the positive most frequent value of M and g characterizing the population. As a part of secondary verification of performance of proposed estimation strategy, which is sample based with CI-Imputation, the method of confidence interval (CI) is used as a tool. It is found that all the estimated confidence intervals are catching the true unknown mean value of the internal+secret economic bond levele of interest which is strength of the proposed. The proposed estimation strategy E found robust in terms of different q_{opt} values as the predicted range of confidence intervals are almost same over varying q_{opt}. The UL-Imputation method is better than the LL-Imputation in comparision. The content of this paper has use of double imputations use like Pattern imputation and CI-Imputation together and both are effectively implemented.

In the war-weapon current situation of Ukraine-Russia this methodology can be used to evaluate the average amount of Economic bond (specially secret support) existing among countries assisting or involved indirectly to the war of Ukraine and Russia on either side.

Acknowledgement

Authors are thankful to both the refrees for valuable comments and useful suggestion which has improved quality of manuscript.

References

Coxeter, H., (1950). Self-dual configurations and regular graphs. *Bulletin of American Math. Soc*, 56, pp. 413–455.

Cochran, W. G., (2005). Sampling Techniques. *Jhon Willy and Sons*.

Donga, J., Bhojak, P., Patel, K. and Shah, S., (2021). An analysis of different computer science algorithms with the graph theory of Mathematics. *Reliability: Theory and Methods*, vol. 16, pp. 376–383.

Deo, N., (2001). Graph theory with application to engineering and computer science.*Eastern Economy Edn. New Delhi, India* .

Krnc, M., Pisanshi, T., (2018). Characterization of generalized Petersen graphs that are kronecker covers. arxiv:1802.07134v1[Math.Co].

Rajoriya, D. and Shukla, D., (2020). Optimal estimation strategies of parameters in Hamiltonian circuit type population. *Journal of Critical Reviews*, 7, pp. 3561–3573.

Rajoriya, D. and Shukla, D., (2021). An optimum resource score estimation method using Bipartite graph model and single node systematic sampling. *Reliability: Theory and Methods*, vol. 16, pp. 322–338.

Shukla, D., (2002). F-T estimator under two-phase sampling. *Metron*, 60, pp. 97–106.

Shukla, D., Rajput, Y. S. and Thakur, N. S., (2009). Estimation of spanning tree mean-edge using node sampling. *Model Assisted Statistics and Applications*, 4, pp. 23–37.

Shukla, D., Rajput, Y. S. and Thakur, N. S., (2010). Edge estimation in population of planer graph using sampling. *Journal of Reliability and Statistical Studies*, 3, pp. 13–29.

Shukla, D., Rajput, Y. S. and Thakur, N. S., (2014). Edge estimation in the population of a binary tree using node-sampling. *Communication in Statistics Theory and Methods*, 43, pp. 2815–2829.

Shukla, D., Pathak, S. and Trivedi, S. K., (2016). Graph sampling by Spanning tree under stratified setup. *Research and Review: Journal of Statistics*, 5, pp. 11–31.

Singh, S., (2003). Advanced Sampling Theory with Applications.*Kluwer Academic Publisher, Netherland, Springer, Dordrecht*, htttps://doi.org/10.1007/978-94-007-0789-4.

Shukla, D.,Trivedi, S. K., Pathak, S. and Rajoriya, D., (2018). Mean edge estimation in population of Binary tree using stratified sampling procedure. *Research and Reviews:Journal of Statistics*, 7, pp. 60–73.

Shukla, D., Trivedi, S. K., Pathak, S. and Rajoriya, D., (2019). Mean edge estimation in population of Planar graph using stratified sampling procedure. *Research and Reviews: Journal of Statistics*, 8, pp. 11–31.

Watkins, M. E., (1969). A theorem on tait colorings with an application to the generalized Petersen graphs *Journal of Combination Theory*, 6(2), pp. 152–164.

Russia-Ukraine war update: https://www.hindustantimes.com/world-news/russia-ukraine-war-live-updates-may-4-wednesday-day-70-of-war-101651621959042-amp.html

China support to Ukraine, *https://www.nytimes.com/2022/03/13/us/politics/russia-china-ukraine.html*

Nine big question answered by Russia, *https://www.google.com/search?channel=nrow5 client=firefox-b-d q=9+big+questions+about+Russia%E2%80%99s+war+in+Ukraine %2C+answered+-+Vox.html*

Russia- Ukraine history and conflict, *https://www.csis.org/analysis/russias-war-ukraine-identity-history-and-conflict*

The independent resource on global security, *https://sipri.org/media/press-release/2021/global-nuclear-arsenals-grow-states-continue-modernize-new-sipri-yearbook-out-now*

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 321–330

About the Authors

Bondaruk Taisiia is a Full Professor of Finance, head of the Department of Finance, Banking and Insurance of the National Academy of Statistics, Accounting and Auditing. Field of scientific interests: problems of the financial and budget system, public debt, regulatory policy, local finances and local self-government, economic security of the state. Author of more than 200 scientific publications, training manuals, monographs. Professor Bondaruk is one of the founders of the scientific school about research of the modern concept of reforming inter-budgetary relations. She is a member of two editorial boards: the scientific journal "Economic Horizons" and the Scientific Bulletin of the National Academy of Statistics, Accounting and Auditing.

Bulatova Olena is a Doctor of Economics, professor, First Vice-Rector, Mariupol State University. Her main areas of interest include investigation of theoretical and practical issues of the development of world economy, international economic relations, focusing on topical issues of integration processes and regionalization of world economy, development of international and regional competitiveness, interregional cooperation, European integration line of Ukrainian policy in terms of globalization. Prof. O. Bulatova has published more than 200 research papers in international/national journals and conferences. She is a member of the Scientific Council of the Ministry of Education and Science of Ukraine (Section of Economics). Currently, she is a member of the editorial boards of Bulletin of Mariupol State University, Series: Economics (Ukraine), Journal of European Economy (Ukraine), Research Papers in Economics and Finance (Poland).

Chala Tetyana is an Associate Professor of Statistics, Accounting and Auditing Department of V.N. Karazin Kharkiv National University; Associate Professor of Management and Administration Department of Educational and Research Institute "Karazin Business School", Kharkiv, Ukraine. Her research interests are statistics, international statistics, international classifications, market conditions analysis. She is a member of the Union of Competitive Intelligence Analysts and Professionals.

Chernenko Daryna is an Associate Professor of Statistics, Accounting and Auditing Department of V.N. Karazin Kharkiv National University; Associate professor of Management and Administration Department of Educational and Research Institute "Karazin Business School", Kharkiv, Ukraine. Her research interests are statistics, time management, market conditions analysis, business communications. She is a member of the Union of Competitive Intelligence Analysts and Professionals.

Gladun Oleksandr is a Doctor of Economics, Corresponding Member of the National Academy of Sciences of Ukraine, Deputy Director for Research at the Ptoukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine. He has worked at the Research Institute of Statistics and the Ministry of Statistics of Ukraine (now the State Statistics Service of Ukraine). Research interests: demography, demographic statistics, demographic losses due to social catastrophes, censuses, population registers, sample surveys. He is the author of more than 100 scientific papers and two individual monographs.

Holubova Halyna – PhD in Economics, Associate Professor of Statistics and Mathematical Methods in Economics Department of the National Academy of Statistics, Accounting and Audit. H. Holubova has published more than 50 research papers in the national journals and conferences. Her scientific interests are: multidimensional statistical methods of analysis, financial statistics and statistics of foreign economic activity. Currently, she is a member of StatTeam GraphPad Prism of the Dotmatics Company. For many years she has held the position of responsible secretary of the Organizing Committee of the All-Ukrainian Student Olympiad in the discipline “Statistics”.

Horobets Olena – PhD in Economic Sciences, Associate Professor of the Department of Statistics and Mathematical Methods in Economics, Deputy head of the editorial and publishing department of the National Academy of Statistics, Accounting and Audit. Responsible editor of the Scientific Bulletin of the National Academy of Statistics, Accounting and Audit, member of the editorial board of the Journal of Modern Economy and Management (Hong Kong). The author of about 30 scientific papers devoted to the issues of big data, the development and dissemination of digitalization processes and information technologies.

Hrynchak Nataliia – PhD in Economic Sciences, Associate Professor of the Department of Economics and Management of Foreign Economic Activity of the National Academy of Statistics, Accounting and Audit. Her current research interests concern the problems of research, evaluation and forecasting of the transformation of world and national markets for goods and services. Author of more than 50 scientific papers. Scientific Secretary of the National Academy of Statistics, Accounting and Audit.

Ivashchenko Oksana is a PhD in Economics, Docent. Head of Department of Economics and Management of Foreign Economic Activity. In 2007, she defended her thesis for the degree of PhD in Economics on the topic: «Financial and Institutional Instruments of Risk Management in Foreign Economic Activity». Docent Ivashchenko has extensive experience in research and practical work in the field of global economy, green economy, sustainable development, global value chains. Author of more than 250 scientific articles indexed in international scientific databases and three monographs.

Kobylynska Tetiana – Doctor of Economics, Deputy Head of the Main Department of Statistics in Zhytomyr Region, Professor of the Department of National Security, Public Administration and Administration of the State University "Zhytomyr Polytechnic". Permanent member of the International Statistical Institute (ISI), member of the editorial board of the scientific information journal "Statistics of Ukraine", chief developer of "Methodological Regulations for the preparation of auxiliary (satellite) account of environmental expenditures", approved by the State Statistics Service of Ukraine from 29.12. 2018 № 297. Her current research interests are in environmental and economic accounting, environmental statistics, climate change statistics, green economy. Author of more than 70 scientific papers.

Korepanov Georgii is an Associate Professor of Statistics, Accounting and Auditing Department of V.N. Karazin Kharkiv National University; Associate Professor of Management and Administration Department of Educational and Research Institute "Karazin Business School", Kharkiv, Ukraine. His research interests are statistics, information and communication technologies, statistical databases. He is a member of the Union of Competitive Intelligence Analysts and Professionals.

Korepanov Oleksiy is a Professor of Statistics, Accounting and Auditing Department of V.N. Karazin Kharkiv National University; Professor of Management and Administration Department of Educational and Research Institute "Karazin Business School", Kharkiv, Ukraine. His research interests are statistics, statistical modelling and forecasting, methods of data classification, managing the development of smart cities. He is a member of Specialized Academic Council of the National Academy of Statistics, Accounting and Audit, a member of the jury of the committee of the All-Ukrainian competition of student scientific research in the specialty "Economic analytics and statistics" at Taras Shevchenko National University of Kyiv. He is also a member of the Union of Competitive Intelligence Analysts and Professionals.

Krekhivskyi Oleg is a PhD in Physics and Mathematics. His research interests are creation of statistical and analytical tools for monitoring and evaluating the effectiveness of industrial and innovation policies for making adequate management decisions. Author of more than 30 papers in peer-reviewed publications. Currently, he is a Senior Desk Officer, Secretariat of Cabinet of Ministers of Ukraine and conducts analysis of normative documents that regulate industries.

Kuzmenko Olha is a Doctor of Science, Professor, Head of the Economic Cybernetics Department, of Academic and Research Institute of Business, Economics and

Management, Head of the Research and Training Centre for Business Analysts, Sumy State University. She is a Chief Editor of the International Scientific Journal "SocioEconomic Challenges", a reviewer of scientific journals "Global Economics Science (Singapore), "Financial and credit activity: problems of theory and practice" (WoS index), a member of the European Marketing and Management Association (EUMMAS), the Specialized Academic Council at Sumy State University and at Simon Kuznets Kharkiv National University of Economics. Her main research interests: modelling of economic processes, health system modelling, econometric models, digitalization of the public sector, social and economic relations, stability and sustainability of socio-economic systems, economic security, insurance and reinsurance markets.

Lazebnyk Iuliia is a Professor of Statistics, Accounting and Auditing Department of V.N. Karazin Kharkiv National University; Professor of Management and Administration Department of Educational and Research Institute "Karazin Business School", Kharkiv, Ukraine. Her research interests are statistics, business research methods, statistical modelling, optimization of management decisions, data visualisation. She is a member of the Union of Competitive Intelligence Analysts and Professionals.

Legan Iryna – Doctor of Law, PhD in Economic Sciences, Associate Professor, Head of the Department of Psychology and Social Security of State University "Zhytomyr Polytechnic". She has completed postgraduate studies at the Institute of Demography and Social Research named after M. Ptoukha of the National Academy of Sciences of Ukraine. In 2021, at the University of the State Fiscal Service of Ukraine, she defended her doctoral dissertation on the topic: "Theoretical and legal foundations of international cooperation in preventing and combating transnational crime" in the specialty 12.00.08 "Criminal Law and Criminology; criminal executive law". Research interests – transnational crime, international law enforcement organizations, legal psychology, criminology, social problems of the labour market. Author of more than 50 scientific papers.

Libanova Ella – Academician of National Academy of Sciences of Ukraine, Doctor of Science (Economics), Professor, Director of the Ptoukha Institute of Demography and Social Studies of the National Academy of Sciences of Ukraine. Academician of the International Union for the Scientific Study of Population and Member of State Commission on the Economic and Social Development. Research experience in the area of demography, social economy, social policy, labor market, human development and migration. Author of more than 400 publications.

Lubenchenko Olga is a Full Professor of Accounting and Audit Department. Her scientific interests include: accounting, audit, quality of audit services. Professor Olga

Lubenchenko has published more than 120 research papers in international/national journals and conferences. She has also published eight books/monographs. Professor Lubenchenko is an active member of many scientific professional bodies.

Momotiuk Liudmyla is a Full Professor of Finance. Her research interests: statistics of finance, public finance, financial stability. Professor Momotiuk has published more than 120 research papers in international/national journals and conferences. She is deputy editor-in-chief of the journal "Statistics of Ukraine", a member of the editorial boards of the journal "Scientific Bulletin of the National Academy of Statistics, Accounting and Audit".

Motuzka Olena – PhD in Economic Sciences, Associate Professor of the Department of economics and management of foreign economic activity, head of the editorial and publishing department of the National Academy of Statistics, Accounting and Audit. Responsible editor of the scientific and information journal "Statistics of Ukraine". The author of about 50 scientific papers devoted to issues of waste management, water resources management, environmental accounting, management of international investment policy.

Ogay Maryna, PhD in Economics, Senior Researcher Fellow, Leading Research Fellow of the Department of Socio-Economic Processes and Structures Modeling in the Ptoukha Institute of Demography and Social Studies of the National Academy of Sciences of Ukraine (Kyiv). She has more than 50 scientific publications on statistics and economics. Her field of scientific interests is the methodology of conducting statistical and monitoring studies in the socio-demographic and economic spheres (information provision, research tools, statistical analysis using modern data processing methods), tools for evaluating the effectiveness of social policy.

Osaulenko Oleksandr H. – Rector of the National Academy of Statistics, Accounting and Audit, Doctor of Public Administration, Professor, Corresponding Member of the National Academy of Science of Ukraine. During 1996–2014, he headed the national statistical office of Ukraine. Author of more than 200 scientific papers, including 20 monographs and hand books on the problems of statistics, public administration and information security. Head of the Specialized Scientific Council of the National Academy of Statistics, Accounting and Audit. Editor-in-chief of two editorial boards: Statistics of Ukraine, Scientific Bulletin of National Academy of Statistics, Accounting and Audit. His research results became the theoretical and methodological basis for the preparation of long-term programs for development of national statistics and improvement of statistical support management at national and regional levels, and received practical implementation in a number of legislative and other normative legal acts.

Pavlova Halyna is a Doctor of Economics, Professor of the Department of Accounting, Taxation and Management of Financial and Economic Security at Dnipro State Agrarian and Economic University, and simultaneously holds the position of Dean of the Faculty of Accounting and Finance at the same university. Her main areas of interest include: accounting and financial accounting, corporate taxation, auditing. Professor Pavlova has published more than 150 scientific papers in international/national journals and conferences, and is a member of the editorial board of two scientific professional publications "Agrosvit" and the electronic journal "Effective Economy", which are included in the list of scientific professional publications of Ukraine on economics.

Perkhun Larysa is a PhD, Associate Professor, Data Engineer of the Big Data Expertise Centre of the Information Technology Department of the Agro-Industrial Holding MHP. She has 25 years of experience in teaching information technologies and mathematical methods in economics in higher education institutions. Her research interests are mathematical models of economic processes at micro and macro levels, information technologies in economics and business, distance learning, business analytics and data scientists. She has published four tutorial/monographs and more than 80 research papers in international/national journals and conferences.

Pozniak Oleksii – PhD (Economics), Head of the Migration Studies Department at Ptoukha Institute of Demography and Social Studies of National Academy of Sciences of Ukraine. He has been studying population migrations since 1993. His research interests are population migration, labour migration, migration impact on social and economic development, migration policy, irregular migration, situation of IDPs and refugees, migrants' integration, migration forecasting. Author of more than 200 publications.

Ptashchenko Olena – Doctor of Economics, Professor, Professor of the Department of International Economic Relations, Dean of the Faculty of International Relations and Journalism of the Simon Kuznets Kharkiv National University of Economics. Promotes the development of youth initiatives and the involvement of young people in international grants and projects in cooperation with the Charitable Foundation "Ukrainian Charitable Alliance" and the NGO "Elite of the Nation". Her current research interests are in the world economy (migration flows; global problems of international economic relations, introduction of international marketing tools for the high-tech sphere). Author of more than 260 scientific publications.

Puhachova Maryna – Doctor of Economics on Statistics, PhD in Technical Sciences on Information Technologies, Chief Researcher of the Ptoukha Institute for Demography and Social Studies of the National Academy of Sciences of Ukraine. She was a part-time professor of the Kyiv National Economic University named after Vadym Hetman, of the National University of Kyiv-Mohyla Academy and of Taras Shevchenko National University of Kyiv. She was a researcher at the Department of Information Systems of Kyiv Polytechnic Institute, director of the Institute of Statistical Technologies, Director General of the Scientific and Technical Complex of Statistical Research of the State Statistics Service of Ukraine. Her current research interests are in the use of electronic registers and databases to study demographic and social processes. She has more than 180 scientific papers on statistics and decision-making theory.

Rajoriya Deepika recently submitted her PhD thesis in Mathematics under the supervision of Prof (Dr.) Diwakar Shukla in the department of Mathematics and Statistics, Dr. Harisingh Gour Central University, Sagar (MP) India. She has published four articles during her PhD. She has MSc (Mathematics). During her PhD, she was a recipient of Research Fellowship of UGC. Her research interests are sampling theory, graph theory and graph sampling.

Reznikova Nataliia is a Doctor of Economic Sciences, Full Professor of World Economy and International Economic Relations. Professor Reznikova has extensive experience in research and practical work in the field of international economic policy, sustainable development, green economy, international trade, analysis of strategies of economic development of countries, the development of mechanisms, incentive tools and instruments for the promotion of investments and innovations. An expert of the project "Internal Displacement in Ukraine: Scale and Challenges of Managing Under Uncertainty" (2022). Member of working group and researcher on the development of draft laws for the Committee of the Parliament of Ukraine on Industrial Policy and Entrepreneurship. Author of more than 400 scientific articles indexed in international scientific databases and three monographs.

Salikhova Olena is a DSc in Economics, Leading Researcher, State Organization "Institute for Economics and Forecasting of NAS of Ukraine". Research interests: creation of scientific-methodical and information-statistical tools for evaluating high-tech manufacturing, including pharmaceutical industry, foreign trade in high-tech goods, monitoring and analysis of high-tech enterprises; author of more than 100 papers in peer-reviewed publications.

Sarioglo Volodymyr, Doctor of Science (Statistics), Professor of Economics, Head of Department of Socio-Economic Processes and Structures Modeling in the Ptoukha Institute of Demography and Social Studies of the National Academy of Sciences of Ukraine (Kyiv). Has more than 90 scientific publications on statistics and economics, including 10 monographs. Fields of scientific interest: information support of economic and socio-demographic studies, methodology of sample surveys of the population and enterprises; quality of statistical data and information; development and implementation of modern methods on using microdata of sample surveys and registers, in particular microsimulation methods; gathering, processing and statistical use of big data; monitoring and evaluation of social policy. Member of the International Institute of Statistics (ISI), the International Association of Survey Statisticians (IASS), the International Microsimulation Association (IMA).

Shukla Diwakar is presently working as a Professor in Statistics, Department of Mathematics and Statistics, H.S. Gour University, Sagar, India. He has MSc (Statistics), PhD (Statistics), MTech (Computer Science) and with more than 30 years of teaching experience. During his PhD he was a recipient of Research Fellowship of CSIR. Until now he has published more than 150 research papers in journals and supervised 20 PhD theses. He was recipient of MPCOST Young Scientist Award, ISAS Young Scientist Medal and UGC Career Award. He has authored 8 books and is a member of 11 learned bodies. His research interest are sampling theory, graph theory, stochastic modelling, big data, data mining, computer network, scheduling and operating system.

Shylga Svetlana is a Full Professor of Accounting and Audit Department. Her scientific interests include: accounting, audit, quality of audit services. Professor Shylga has published more than 100 research papers in international/national journals and conferences. She has also published five books/monographs. Professor Shylga is an active member of many scientific professional bodies.

Vasyechko Olga received her Doctor in Economics from Kyiv National Economy University, and an Honorary Professor from National Academy of Statistics, Accounting and Audit, Kyiv. Her professional experience includes being a Director of Business Statistics Department in the Ukraine State Statistics Committee, a General Director of Scientific and Technical Complex for statistical research, and a professor at Kyiv National Economy University, at the University Paris 1 Panthéon-Sorbonne, an Associate Professor at the Paris-Panthéon-Assas, as well as an international expert in official statistics. Her research interests are business statistics, statistical indices, FDI, outlier detection, and economic risk management.

Yarovenko Hanna is a Doctor of Science, Associate Professor, Visiting Professor of the Computer Sciences and Engineering Department of Carlos III University of Madrid, Associate Professor of the Economic Cybernetics Department of Sumy State University, member of the European Marketing and Management Association, Head of the scientific-research work "National security through the convergence of financial monitoring systems and cyber security: intelligent modelling of financial market regulation mechanisms" financed by the Ministry of Science and Education in Ukraine (No. 0121U109559, 2021–2023). Her scientific interests are information systems and mathematic modelling in economics, cyber security of business entities, digitization of socio-economic relations, financial monitoring, and legalization of criminal proceeds. She is the author of more than 100 scientific publications in international and national journals.

Yatsenko Olha – Doctor of Sciences (Economics), Professor of the Department of International Trade and Marketing at Kyiv National Economic University named after Vadym Hetmana (KNEU), at Kyiv, Ukrainian. She has also taught at the Hennadii Udovenko Diplomatic Academy of Ukraine at the Ministry of Foreign Affairs of Ukraine (DAU) and Polissya National University. She has diverse research interests, mainly analysis of trade policy, economic diplomacy, globalization of agriculture. She has published more than 250 research papers in international/national journals and conferences. Since 2004, she has more than 55 full research articles published in different internationally well reputed journals. Prof. Dr. Olha Yatsenko is a member of editorial boards for many journals, including statistical and international journals. She is a chief editor of scientific works collection KNEU Scientific Notes. She has also served as a peer reviewer to many international journals. She is a member of many professional bodies, in particular she was a Chief of Scientific Committee from Ukraine (economy direction) of Global non-governmental organization Apimondia (2008–2013). She won a research grant for an internship and study in the United States, including scientific and educational internship for the Faculty Exchange Program in the USDA and University of Missouri (2016); she also won a Modules of the British Council's Researcher Connect Course (2017).

Zaichko Iryna – Associate Professor, Associate Professor of the Department of Finance, Banking and Insurance of the National Academy of Statistics, Accounting and Audit. Her scientific interests are research in the field of budgetary policy and financial security of the state, financial management of enterprises and investment activities in the market of financial services. She is the author of more than 120 scientific and educational-methodical works.

Zvarych Iryna is a Doctor of Economic Sciences, Professor, Head of the Department of International Economics of the Educational and Scientific Institute of International Relations named after B.D. Havrylyshyn of the West Ukrainian National University. Performer of fundamental state-budget scientific research on the topic "National concept of eco-security of society and inclusion of the circular economy in the conditions of the pandemic" (2020–2023). Scholarship holder of the Cabinet of Ministers of Ukraine for young scientists 2020–2022. The sphere of her current scientific interests is the global inclusive circular economy; global economy; global (international) chains of added value, international production networks. Author of more than 100 scientific publications.

Zvarych Roman is a Doctor of Economic Sciences, Professor, Head of the Department of International Economic Relations of the Educational and Scientific Institute of International Relations named after B. D. Havrylyshyn of the West Ukrainian National University. The responsible executor of the fundamental state-budget scientific research on the topic “National concept of eco-security of society and the inclusion of the circular economy in the conditions of the pandemic” (2020–2023). Scholarship holder of the Cabinet of Ministers of Ukraine for young scientists 2020–2022. His research interests are related to the global economy, global conflicts, alter-globalization, economic convergence, ecological growth, circular economy, corporate social responsibility. Author of more than 110 scientific publications.

STATISTICS IN TRANSITION new series and *STATISTICS OF UKRAINE, February 2023*
A New Role for Statistics: Joint Special Issue
Vol. 24, No. 1, pp. 331–332

Acknowledgements to Reviewers

The Editor and Editorial Board of the Statistics in Transition new series wish to thank the following persons who served as peer-reviewers of manuscripts for the *Statistics in Transition new series* and *Statistics of Ukraine* – Volume 24, Number 1 the authors` work has benefited from their feedback.

Aleksandrova Petrova Simeonka, Department of Commerce, Tsenov Academy of Economics, Bulgaria

Brown James J., Department of Statistics, School of Mathematical & Physical Sciences, University of Technology Sydney, Australia

Cyrek Magdalena, Department of Economics and Finance, University of Rzeszow, Poland

Dubey Vyas, Department of Statistics, Ravi Shankar Shukla University, India

Fesenko Valeriia, Department of Accounting, Auditing, Analysis, University of Customs and Finance in Dnipro, Ukraine

Ganushchak-Efimenko Lyudmyla, Department of Computer engineering, Kyiv National University of Technologies and Design, Ukraine

Gonchar Igor, Department of Statistics, Information, Analytical systems and Demography, Taras Shevchenko National University of Kyiv, Ukraine

Grmanová Eva, Department of Social and Economic Relations, Alexander Dubcek University of Trencin, Slovakia

Grun-Rehomme Michel, Department of Statistics, Paris Panthéon-Assas University, France

Honcharenko Iryna, Department of Economics, Cherkasy State University, Ukraine

Ion Negru, Department of Management and Entrepreneurship, Academy of Economic Studies of Moldova, Moldova

Jasilionis Domantas, Max Planck Institute for Demographic Research**,** Germany, Vytautas Magnus University, Lithuania

Kaczmarczyk Paweł, Centre of Migration Research, Poland

Khan Bhola, Department of Economics, Yobe State University, Nigeria

Kordoš Marcel, Department of Public Administration and Regional Economics, Alexander Dubček University in Trenčín, Slovakia

Kovtun Natalia, Department of Statistics, Information Analytical Systems and Demography, Taras Shevchenko National University of Kyiv, Ukraine

Ladiray Dominique, INSEE (Institut National de la Statistique et des Études Économiques), France

Lunina Inna, Department of Public Finances, National Academy of Sciences of Ukraine, Ukraine

Malynovska Olena, National Institute for Strategic Studies, Ukraine

Mazurenko Olga, Department of Statistics, Information and Analytical Systems and Demography, Kyiv National University, Ukraine

Nifatova Olena, Department of Economics and Business, University of Barcelona, Spain

Orekhova Tatiana, Department of International Economic Relations, Donetsk National University, Ukraine

Rębilas Rafał, WSB University, Poland

Riashchenko Viktoriia, Department of Statistics, ISMA University of Applied Sciences (ISMA), Latvia

Rohr Margarita, Department of Applied Economics, University of Valencia, Spain

Sadova Ulyana, Dolishniy Institute of Regional Research of the National Academy of Sciences of Ukraine, Ukraine

Sarioglo Volodymyr, Ptoukha Institute for Demography and Social Studies of the National Academy of Science of Ukraine, Ukraine

Tsal-Tsalko Yuzef S., Department of Accounting, Taxation and Audit, Polissia National University, Ukraine

Valinkevich Natalya, Department of Economics, Business and Tourism, Polissia National University, Ukraine

Vasa László, Department of Management, Széchenyi István University, Hungary

Wosiek Małgorzata, Department of Economics and Finance, University of Rzeszow, Poland

GUIDELINES FOR AUTHORS

We will consider only original work for publication in the Journal, i.e. a submitted paper must not have been published before or be under consideration for publication elsewhere. Authors should consistently follow all specifications below when preparing their manuscripts.

Manuscript preparation and formatting

The Authors are asked to use *A Simple Manuscript Template (Word or LaTeX) for the Statistics in Transition Journal (published on our web page:* https://sit.stat.gov.pl/ForAuthors.

- ***Title and Author(s)***. The title should appear at the beginning of the paper, followed by each author's name, institutional affiliation and email address. Centre the title in **BOLD CAPITALS**. Centre the author(s)'s name(s). The authors' affiliation(s) and email address(es) should be given in a footnote.
- ***Abstract.*** After the authors' details, leave a blank line and centre the word **Abstract** (in bold), leave a blank line and include an abstract (i.e. a summary of the paper) of no more than 1,600 characters (including spaces). It is advisable to make the abstract informative, accurate, non-evaluative, and coherent, as most researchers read the abstract either in their search for the main result or as a basis for deciding whether or not to read the paper itself. The abstract should be self-contained, i.e. bibliographic citations and mathematical expressions should be avoided.
- ***Key words***. After the abstract, Key words (in bold) should be followed by three to four key words or brief phrases, preferably other than used in the title of the paper**.**
- ***Sectioning***. The paper should be divided into sections, and into subsections and smaller divisions as needed. Section titles should be in bold and left-justified, and numbered with **1.**, **2.**, **3.**, etc.
- ***Figures and tables***. In general, use only tables or figures (charts, graphs) that are essential. Tables and figures should be included within the body of the paper, not at the end. Among other things, this style dictates that the title for a table is placed above the table, while the title for a figure is placed below the graph or chart. If you do use tables, charts or graphs, choose a format that is economical in space. If needed, modify charts and graphs so that they use colours and patterns that are contrasting or distinct enough to be discernible in shades of grey when printed without colour.
- ***References.*** Each listed reference item should be cited in the text, and each text citation should be listed in the References**.** Referencing should be formatted after the Harvard Chicago System – see http://www.libweb.anglia.ac.uk/referencing/harvard.htm. When creating the list of bibliographic items, list all items in alphabetical order. References in the text should be cited with authors' name and the year of publication. If part of a reference is cited, indicate this after the reference, e.g. (Novak, 2003, p.125).