IDENTIFICATION OF DETERMINANTS OF CORRUPTION IN GOVERNMENT: A MAR-SPLINE APPROACH

**Purpose.** Development of a scientific and methodological approach to the identification of the most impactful determinants of corruption on the basis of multivariate adaptive regression splines.

**Methodology.** Methodological tools of the research methods are comparison, grouping, bibliometric analysis, and multivariate adaptive regression splines in the form of piecewise linear functions.

**Findings.** Systematization of the literary sources and approaches for factors influencing corruption indicates that most empirical studies are based on using panel data. Panel data allows you to insert general patterns, but does not consider the patterns of the national economy. For the study on corruption in Ukraine, 15 influencing factors were selected, characterizing the institutional, economic, and social environment. Based on the constructed MAR Spline models, three regression equations were obtained that describe the linear dependence of the level of corruption in governance on the selected factors. The paper found that the relevant factors influencing corruption in Ukraine are: tax burden, general government final consumption expenditure, average monthly wage in Governance and rule of law.

**Originality.** The proposed approach makes it possible to determine the dynamics of the degree of factor influence on the level of corruption in the country. The paper defines the threshold values of statistically significant indicators at which the maximum degree of correlation with the corruption perception index is achieved.

**Practical value.** The regularities between the level of corruption and economic, institutional, and social factors revealed by the research results can be used in the development of tools to fight corruption in Ukraine. The formation of an effective anti-corruption system will strengthen financial stability in the country and increase the level of public trust in society.

**Keywords:** effective institutions, corruption, tax burden, regression splines, tax burden, living wage

**Introduction.** Corruption is a complex and multifaceted phenomenon that permeates various spheres of business. For many decades, the issue of corruption has attracted the attention of sociologists, political scientists, and economists, and the governments and international organizations have allocated significant financial resources to combat it. Despite this, the forms of corruption offenses are constantly being improved, and therefore corruption is still widespread in many countries of the world.

Financial losses from corruption, bribery, tax evasion and other illegal financial transactions for developing countries amount to $1.26 trillion annually. The member states of the European Union lose 132 billion dollars due to corruption every year [1].

This shows that the fight against corruption is not a problem of any individual countries or regions, it has become a global challenge for the international community. Corruption creates destructive mechanisms for stable economic development, weakens democratic institutions and the principle of rule of law, exacerbates the problem of uneven distribution of income and access to medical, educational and other services [2]. The high level of corruption risk in the country affects the inflow of foreign investments into the country [3].

**Literature review.** The issue of corruption constantly causes sharp controversy in both political and scientific circles. The discussion is mainly about determining the causes of the appearance of unscrupulous behavior of public administration bodies and methods for combating corruption. A bibliometric analysis was carried out to study trend patterns of publishing activity on corruption. One of the most authoritative scientometric and reference databases, Scopus, was chosen for the analysis. In the search query, scientific publications containing the concept of corruption in the title of the work and its keywords were selected.

In the period of 1991–2000, 738 scientific works were published, which were indexed by the scientometric database Scopus, in 2001–2010 – 3624 works, and in 2011–2020 – 10,951 works. That is, the number of publications has increased almost 10 times over 20 years, and almost 3 times over the last 10 years.

The largest number of articles was published in the USA (27 %), Great Britain (12 %), Germany (6 %), Australia (5 %), China (5 %). Analyzing the fields of research in which the articles were published, it is possible to trace a clear advantage of publications of an economic direction, in particular, these are social sciences, Economics, Econometrics, and Finance and Business.

In the scientific literature, there are quite a lot of works that analyze the determinants of influence on corruption. In particular, Elbansawy and Revier [4] have divided the determinants of influence on corruption into 2 groups; those that vary over time and those that are time-invariant. Relevant determinants of influence are the rule of law, the quality of law enforcement, freedom of speech, and accountability.

Main determinants of corruption in Visegrad countries are economic, political, and socio-cultural: phase of economic development, openness to trade, size of the public sector, degree of urbanization, percentage of females in the labor force [5].

Cariole [6] has suggested multi-level framework for estimation of corruption that considers the economic and human development levels, the size of governments, trade openness, and democracy.
In paper by Shabbir and Anwar [7], the determinants of corruption are divided into economic (economic freedom, globalization, education level, income distribution and average income level) and non-economic (press freedom, democracy and share of the population associated with a certain religion). Shabbir and Anwar [7] proved that all economic determinants are negatively related to the perceived level of corruption, except for income distribution, and non-economic determinants do not significantly explain the variation in the level of corruption.

Consequently, there are numerous empirical studies in the scientific literature that analyze various variables to explain corruption by identifying significant determinants. However, the basis of most empirical studies aimed at determining the factors of the spread of corruption is the use of panel data. The use of panel data allows us to insert general patterns, but does not take into account the specifics of the development of the national economy. This necessitated the collection of statistical data across Ukraine and the use of mathematical tools that will adequately assess the impact of various determinants on the level of corruption.

**Purpose** is to identify which determinants are the most impactful on corruption in Ukraine using MAR spline.

**Methods.** Multivariate Adaptive Regression Splines (MAR-spline) is an algorithm for complex non-linear regression problems. The algorithm involves finding a set of simple linear functions that in aggregate result in the best predictive performance [8, 9].

The MAR-spline toolkit was used to identify the points of increasing corruption risk in the country caused by external economic shocks. Multivariate adaptive regression MAR splines are a nonparametric procedure for formalizing dependence on a set of basic functions and coefficients that are completely determined by the input data set. This procedure is based on an approach according to which the set of values of input variables (regressors) is divided into areas with their own specific equations of regression and classification. This approach involves the construction of adaptive models that allow obtaining reliable forecasts and are used in cases of turning points and formalization of the non-monotonic nature of the relationship between effects and responses, which are difficult to approximate parametric models.

The basic functions of multivariate adaptive regressive MAR splines before and after the regression switch point are described as follows

\[ (x-t)_+ = \begin{cases} -t, & \text{if } x < t \\ 0, & \text{if } x \geq t \end{cases} \]

\[ (x-t)_- = \begin{cases} t-x, & \text{if } x < t \\ 0, & \text{if } x \geq t \end{cases} \]

where \( t \) is the inflection point of the piecewise function.

Basic functions in MARSplines (Multivariate Adaptive Regression Splines) in STATISTICS software are usually formalized as the following mathematical relations

\[ (x-t)_+ = \max(0; x-t); \]

\[ (x-t)_- = \max(0; t-x); \]

In the case of formalizing a multidimensional dependence, for each component of the regressor vector, basic functions (1 and 2) are constructed, which determine the set of basic functions built based on a set of input data

\[ B = \{(x_1-t)_+, (t-x_1)_+, \ldots, (x_n-t)_+, (t-x_n)_+\}_{i=1}^{m} \]

The general equation of multivariate adaptive regressive MAR splines for the \( m \) nonzero constituent members is written as a combination of the weighted sum of basic functions and their products

\[ y = f(X) = \alpha_0 + \sum_{j=1}^{m} \alpha_j \cdot B_j(X), \]

where \( \alpha_0 \) is a constant, free term; \( \alpha_j \) is constant, parameter of the multivariate adaptive regressive equation; \( m \) is the total number of basic functions; \( X \) is vectors of input regressors; \( B_j(X) \) is the \( j^{th} \) basic function from set \( B \) or product of two or more such functions.

The basic principle of constructing multidimensional adaptive regressive MAR splines provides for the definition of both basic functions and terms that determine the number of different combinations of basic functions, considering requests to each of the relevant regressor factors.

The Corruption Perceptions Index (CPI), calculated by Transparency International, was chosen to characterize corruption in the country. In this study we divided the determinants of corruption into three parts: economic factors, institutional factors and social factors.

**Economic Factors. Tax Burden score (TAX).** An excessive tax burden can be a reason for tax evasion or a reduction of the tax base due to bribery of a tax administration employee [10]. However, under the condition of an effectively functioning system of tax administration and supervision, even an excessive tax burden will not serve as a driver of the growth of corruption risks in the country.

**General government final consumption expenditure, % of GDP (GFCE).** The high level of corruption in the country leads to irrational and untargeted use of budget funds. In order to gain an undue advantage, government officials may approve budgets and other financial plans with “inflated” funding [11].

**Monetary freedom score (MON).** The transparent monetary policy of the central bank eliminates manifestations of corruption in the country. However, in the case of bribery of high-ranking officials, short-term and/or artificial obstacles may be created in the money market for financial gain by a separate group of persons.

**Trade freedom score (TRD).** The introduction of quotas for the import/export of relevant products is a tool of corruption. Quotas will leave no space for free competition [12].

**Number of state-owned enterprises (SOE).** State-owned enterprises were closely linked to political processes, acting as “black boxes” for the political financing of incumbent governments.

**Institutional Factors. Number of political parties (PP).** The excessive number of political parties leads to an increase in the manifestations of political corruption: illegal financing of parties, actions related to deputy ethics, ethics of high-ranking officials, selling seats in the potentially passable part of the party’s electoral list; bribery of voters, etc. [13].

**Press Freedom Index (PFI).** Freedom of expression, free access to public information, and freedom of mass media are fundamental principles of a democratic society, the observance of which allows reducing the level of corruption in the country. However, the lobbying of interests by individual political forces due to the ownership of mass media and TV channels can hide manifestations of corruption in society [14].

**Political Stability and Absence of Violence/Terrorism (PS).** The unstable political environment provokes government officials to act more opportunistically, thereby increasing the risk of corruption in the country. At the same time, a long period of stay of a civil servant in one position can serve to establish stable relations with potential suppliers [15].

**Regulatory Quality (RQ).** An effectively functioning regulatory mechanism, which provides for the existence of appropriate standards regarding the necessary preparation and decision-making procedures at all levels of power, allows reducing information asymmetry and manifestations of corruption [16].

**Rule of Law (RL).** Improving the principles of the rule of law in society increases the probability of identifying and pun-
ishing the initiators of unscrupulous behavior, and also reduces the level of corruption risk in the country [17].

**Social Factors. Income share held by lowest 10 % (INC_L).** Uneven conditions for employment and a low level of remuneration can lead to excessive differentiation of the population’s incomes, which causes the growth of domestic corruption, increasing social tension in society [18].

**Income share held by highest 10 % (INC).** Illegal obtaining of illegal benefits in particularly large amounts, illegal appropriation of budget funds and oligarchization of the economy are systemic results of the existence of corruption in the country.

**Average monthly wage in Governance (WG_G).** A decent salary is one of the basic elements for combating corruption in the sphere of public administration. However, a high level of wages cannot be compared with the level of receiving corruption rents, therefore combating corruption is a synosis of financial and ethical norms.

**Unemployment (% of total labor force) (UNM).** In the presence of a high level of unemployment, the supply of labor on the labor market increases, therefore mechanisms of unscrupulous behavior may arise in the selection of candidates [19].

**Number of registered crimes per 100 thousand population (CRM).** The criminogenic situation in the country can serve as a driver for the growth of corruption offenses. However, the lack of an effective system of accountability for the commission of crimes can also act as a motive for the activities of corrupt officials [20].

Thus, 15 indicators were chosen to determine the factors influencing corruption in Ukraine. The source of primary data was the data of the World Bank, the State Statistics Service of Ukraine. 1998–2021 was chosen as the research period.

**Results.** In this part, we implemented MARS models to uncover the possible determinants of corruption in Ukraine.

The first step in constructing MAR-splines is to determine the relevance of the influence of groups of economic, social, and institutional factors on the corruption promotion index. Parameters for MAR-splines are presented in Table 1.

So, the number of independent variables is 5, the number of dependent variables is 1, the number of terms is 3, 2 and 2 in terms of economic, social, and institutional factors, respectively; the number of basic functions – 2, 1, 1, respectively; the order of interaction (the number of components of the product of basic functions) is 1, as well as the number of references to regressor factors: economic – 1 to TAX, GFCE, social – 1 to WG_G, institutional – 1 to RL.

On the basis of the parameters presented in Table 1, models of the influence of economic, social and institutional factors on the Corruption Perceptions Index were built in the form of multidimensional adaptive regression

\[
\text{CPI} = 1.84971910593605e+001 + \\
+ 3.65181328112318e-001 \cdot \text{max}(0; \\
\text{TAX}-6.23000000000000e+001) + \\
+ 1.25475176258030e+000 \cdot \text{max}(0; \\
\text{GFCE}-1.68907828666189e+001).
\]

Thus, among the considered 5 economic factors, only 2 factors were found to be relevant in the study on the influence on the Corruption Perceptions Index: Tax Burden score (TAX) and General government final consumption expenditure (GFCE).

Analyzing (5), we conclude that an increase in the TAX indicator by 1 point leads to an increase in the Corruption Perceptions Index by 0.365 points, in case if TAX is more than 62.30. Otherwise, TAX will not have an impact on the performance indicator.

Turning to the indicator of general government final consumption expenditure, we note the following: an increase in the GFCE indicator by 1 point leads to an increase in the Corruption Perceptions Index by 1.25 points (only if GFCE is greater than 16.89, otherwise, there will be no GFCE separate-ly have an impact on the performance indicator).

Among the 5 social factors considered, only 1 factor was found to be relevant in the study on the influence on the Corruption Perceptions Index — average monthly wage in Governance (WG_G).

\[
\text{CPI} = 2.35731588615035e+001 + \\
+ 5.51053846161786e-004 \cdot \text{max}(0; \\
\text{WG_G}-2.50000000000000e+002).
\]

An increase in WG_G by 1 point leads to an increase in the Corruption Perceptions Index by 0.00055 points, so that WG_G is more than 25. In other cases there will be no statistically significant relationship between wages in the field of public administration and the level of corruption in the country.

Among the 5 indicators characterizing the institutional environment in the country, only 1 factor turned out to be relevant — Rule of Law (RL).

---

**Table 1**

Parameters of spline modeling of the influence of groups of economic, social, and institutional factors on the Corruption Perceptions Index

<table>
<thead>
<tr>
<th>Model specifications</th>
<th>Economic factors on the CPI</th>
<th>Social factors on the CPI</th>
<th>Institutional factors on the CPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independents</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Dependents</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Number of terms</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Number of basic functions</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Order of interactions</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Penalty</td>
<td>2.000000</td>
<td>2.000000</td>
<td>2.000000</td>
</tr>
<tr>
<td>Threshold</td>
<td>0.000500</td>
<td>0.000500</td>
<td>0.000500</td>
</tr>
<tr>
<td>GCV error</td>
<td>9.243926</td>
<td>9.012103</td>
<td>12.57490</td>
</tr>
<tr>
<td>Prune</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

---

**Table 2**

Regression statistics of MAR-spline models of the dependence of the Corruption Perceptions Index on groups of economic, social and institutional factors

<table>
<thead>
<tr>
<th>Regression statistics</th>
<th>Economic factors on the CPI</th>
<th>Social factors on the CPI</th>
<th>Institutional factors on the CPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (observed)</td>
<td>26.00000</td>
<td>26.00000</td>
<td>26.00000</td>
</tr>
<tr>
<td>Standard deviation (observed)</td>
<td>4.05398</td>
<td>4.05398</td>
<td>4.05398</td>
</tr>
<tr>
<td>Mean (predicted)</td>
<td>26.00000</td>
<td>26.00000</td>
<td>26.00000</td>
</tr>
<tr>
<td>Standard deviation (predicted)</td>
<td>3.22326</td>
<td>3.03890</td>
<td>2.52755</td>
</tr>
<tr>
<td>Mean (residual)</td>
<td>0.00000</td>
<td>-0.00000</td>
<td>-0.00000</td>
</tr>
<tr>
<td>Standard deviation (residual)</td>
<td>2.45874</td>
<td>2.68326</td>
<td>3.16958</td>
</tr>
<tr>
<td>R-square</td>
<td>0.63216</td>
<td>0.76191</td>
<td>0.68872</td>
</tr>
<tr>
<td>R-square adjusted</td>
<td>0.57698</td>
<td>0.52019</td>
<td>0.33050</td>
</tr>
</tbody>
</table>
CPI = 1.97021343376212e+001 + + 2.0205096354147e+000 · max(0;
(7)
RL + 1.10880517959595e+000).

An increase in RL by 1 point causes an increase in the corruption promotion index by 20.21 point in case, if this RL is more than −1.11. Otherwise, RL will not have an impact on the performance indicator.

To prove the adequacy of the models (5–7) presented above, consider the indicators listed in Table 2.

Analysis of Table 2 allows us to state that 63.22 % of the variation in the corruption promotion index is explained by the variation in economic factors, 76.19 % is explained by the variation in social factors, and 68.87 % is explained by the variation in institutional factors. In addition, the reliability and accuracy of the models are confirmed by: the minimum value of the general model quality criterion — the generalized moving average error (GCV error), which takes values from 9.01 to 12.57 (Table 1); insignificant deviation of actual and forecast values.

Conclusions. Corruption is one of the main threats to national security, as it undermines the financial system, public trust, and the ability of civil servants to protect national interests through the use of office for personal enrichment.

The fight against corruption is not a sectoral reform. It will have an impact on all spheres of society’s life, but it requires systemic institutional changes, which should ultimately change public consciousness. Only zero tolerance for corruption will allow the introduction of new standards of behavior for civil servants and will restore trust in state institutions.

In this study, we tried to investigate various determinants for perceived level of corruption in Ukraine. We considered the economic, social and institutional determinants of corruption. The empirical findings show that statistically significant determinants of corruption in Ukraine are tax burden score, general government final consumption expenditure, average monthly wage in governance and rule of law. The estimated multivariate adaptive regression splines indicate fluctuation monthly wage in governance and rule of law. The estimated general government final consumption expenditure, average monthly wage in governance and rule of law. The estimated multivariate adaptive regression splines indicate fluctuation monthly wage in governance and rule of law.

References.
Методика. У процесі дослідження використовувалися методи порівняння, групування, бібліометричного аналізу, а також побудовані багатовимірні адаптивні регресійні сплайни у вигляді кусково-лінійних функцій.

Результати. Проведений аналіз наукової літератури з питань дослідження факторів впливу на корупцію засвідчив, що в основі більшості емпіричних досліджень є використання панельних даних. Використання панельних даних дозволяє вставити загальні закономірності, проте не враховує специфіку розвитку саме національної економіки. Для дослідження корупції в Україні відібрано 15 факторів впливу, що характеризують інституційне, економічне й соціальне середовище. За побудованими MAR Spline-моделями отримані три регресійні рівняння, що описують лінійну залежність рівня корупції у країні від обраних факторів. За результарами дослідження встановлено, що релевантними факторами впливу на корупцію в Україні є: рівень податкового навантаження, рівень видатків загального державного управління на кінцеве споживання до ВВП, обсяг середньомісячної заробітної плати в державному управлінні та рівень верховенства права.

Наукова новизна. Запропонований підхід дозволяє визначити динаміку ступеня факторного впливу на рівень корупції у країні. У роботі визначені граничні значення статистично значимих індикаторів, за яких досягається максимальний ступінь взаємозв'язку з індексом сприйняття корупції.

Практична значимість. Виявлені за результатами дослідження закономірності між рівнем корупції та економічними, інституційними й соціальними факторами можуть бути використані при розробці інструментів боротьби з корупцією в Україні. Формування ефективної системи протидії корупції дозволить посилити фінансову стабільність у країні, підвищити рівень громадської довіри в суспільстві.

Ключові слова: ефективні інститути, корупція, податкове навантаження, регресійні сплайни, прожитковий мінімум

The manuscript was submitted 20.06.22.