mathematics

Article

Mathematical Modeling and Nonlinear Optimization in Determining the Minimum Risk of Legalization of Income from Criminal Activities in the Context of EU Member Countries

Alena Vagaská [1,*], Miroslav Gombár [2] and Antonín Korauš [3]

[1] Department of Natural Sciences and Humanities, Faculty of Manufacturing Technologies with a Seat in Prešov, The Technical University of Košice, 080 01 Presov, Slovakia
[2] Department of Management, Faculty of Management and Business, University of Prešov, 080 01 Presov, Slovakia
[3] Department of Information Science and Management, Academy of the Police Force in Bratislava, 835 17 Bratislava, Slovakia
* Correspondence: alena.vagaska@tuke.sk

Abstract: Legalization of the proceeds of crime represents a worldwide problem with serious economic and social consequences. Information technologies in conjunction with advanced computer techniques are important tools in the fight against money laundering (ML), financial crime (FC) and terrorism financing (TF). Nowadays, the applied literature on ML/FC/TF uses much more mathematical modelling as a solving strategy to estimate illicit money flows. However, we perceive that there is preference of linear models of economical dependences and sometimes lack of acceptance of nonlinearity of such investigated economic systems. To characterize the risk of legalization of crime proceeds in a certain country, the scientific researchers use the Basel anti-money laundering (AML) index. To better understand how the global indicators (WCI, CPI, EFI, GII, SEDA, DBI, GSCI, HDI, VAT_{GAP}, GDP per capita) affect the level of risk of ML/TF in the countries of EU, the authors use a unique data set of 24 destination countries of EU over the period 2012–2019. The article deals with two main research goals: to develop a nonlinear model and optimize the ML/TF risk by implementation of nonlinear optimization methods. The authors contribute: (i) providing the cross-country statistical analysis; (ii) creating the new nonlinear mathematical-statistical computational model (MSCM); and (iii) describing the observed dependent variable (Basel AML index). This study deepens previous knowledge in this research field and, in addition to the panel regression analysis, also applies nonlinear regression analysis to model the behavior of the investigated system (with nonlinearity). Our results point out the differences between the estimates of the investigated system behavior when using panel and nonlinear regression analysis. Based on the developed MSC model, the optimization procedure is conducted by applying an interior point method and MATLAB toolboxes and the second goal is achieved: the statement that such values of input variables at which the risk of legalization of income from criminal activity will be minimal.

Keywords: panel regression analysis; nonlinear regression analysis; mathematical optimization; interior point method; money laundering; risk of legalization of financial crime proceeds; Basel AML index

MSC: 62J02; 62J05; 62J20; 90-10; 90-08

check for updates

Citation: Vagaská, A.; Gombár, M.; Korauš, A. Mathematical Modeling and Nonlinear Optimization in Determining the Minimum Risk of Legalization of Income from Criminal Activities in the Context of EU Member Countries. *Mathematics* **2022**, *10*, 4681. https://doi.org/10.3390/math10244681

Academic Editors: Svajonė Bekešienė and Ieva Meidutė-Kavaliauskienė

Received: 13 November 2022
Accepted: 3 December 2022
Published: 9 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The prevention of money laundering (ML), financial crime (FC) and terrorism financing (TF) has been subject to an increasing amount of attention recently. For example, Canadian authors Hilal et al. [1] reviewed research works on the anomaly detection techniques indicating financial fraud and stated that acts of financial frauds have become much more prevalent these days than ever. The economic growth and the continued development of technology in modern society results in the evolution of fraudsters' approaches to exploit the vulnerabilities of the current preventative measures, on the other hand it evokes the need for improvement of prevention and detection of financial frauds. In recent years, efforts to suppress tax evasion and illegal economic activity led to the promotion of the idea of eliminating cash [2], as cash (perceived as "anonymous money"), can play the role of a facilitator of such activities. The socio-economic system is currently very dynamic; the trend of dominance of cashless payment systems and convergence to cryptocurrencies, bitcoins, etc. is evident [3,4]. All this, in addition to the positives, also has its negatives, which causes heated debate among experts [5,6]. Cohen et al. [2] even point to the unwanted consequences of eliminating cash. In their study they construct a simple general equilibrium model in order to demonstrate how elimination of cash-paying can lead to a misallocation of resources in a naturally segmented economy.

In 2020, the United States enacted The Anti-Money Laundering Act, expanding the FinCEN's power [7]. The increasing level of this attention is also evident in Europe, where the requirement to strengthen the anti-money laundering (AML) institutional framework was evoked due to a series of banking scandals in several countries (Cyprus, Estonia, Latvia, Denmark, Finland, the Netherlands and the United Kingdom) [8].

From the point of view of the institutions that investigate the fight against money laundering and terrorism financing, financial and state institutions were mainly active in the past [9,10]. Financial institutions still spend a significant amount of their resources on automated information systems aimed at tracking illicit transactions [1,10]. Efforts to identify money laundering and sophisticated detection of non-standard financial flows (especially through data mining) have accelerated the development of automated information systems and special software [1,10–12], which are associated with huge financial costs. Some of these systems are implemented quickly, so-called "out of the box" to satisfy regulators, and are only later calibrated to detect serious suspicious activity [13]. However, experience shows that institutions are still dissatisfied with the current automated tracking efforts, i.e., they are still looking for software that can reduce the burden on regulators and compliance. After the tightening of legislation and the intensification of the fight against money laundering and terrorism in recent years, universities are gradually taking the initiative through research projects.

Nowadays, the applied literature on ML/FC/TF uses much more mathematical modelling as a solving strategy to estimate the risk of illicit money flows or the differences in ML-related risk-perception across countries [1,14–16]. Overall, mathematical modelling is often used for risk and uncertainty management and modelling [17–27]. As the modus operandi becomes more sophisticated for the financial fraudsters and money launderers, regulators and financial institutions have to fight back by applying innovative countermeasures. To the most effective weapons currently available belong advanced risk-rating models, using statistical analysis and machine learning to provide better quality data analysis [10,23–27]. However, we perceive that in this research field (ML/FC/TF modelling) there is preference for linear models of economical dependences [9,28–31] and sometimes lack of acceptance of nonlinearity of such investigated economic systems, so this situation has supported and stimulated our interest in this research topic.

Contemporary approaches to the risk assessment of ML and TF are related to the Basel AML index, by which scholars characterize the risk of legalization of proceeds from criminal activities [32,33]. In accordance with the authors' effort to contribute to the solution of this issue and optimize (reduce) the risk of legalization of illicit income in the EU countries, the main (the first) research goal of this study is set to create a new computational nonlinear mathematical model that will describe the Basel AML Index behavior. The AML index represents the investigated dependent variable, which is affected by the values of the independent variables, in this case with the global indices: WCI, CPI, EFI, GII, SEDA, DBI, GSCI, HDI, VAT_{GAP} and GDP per capita; all indices defined in EU countries. To create the models, the authors worked with a bank of real data obtained between 2012 and 2019 within the EU member states. According to available sources, until now researchers have mainly focused on the use of panel regression analysis when creating such models [28–31,34–38]. The authors of the article deepen previous knowledge and, in addition to panel regression analysis, apply nonlinear regression analysis to model the behavior of the investigated system (showing nonlinearity). As part of solving the first research task, the authors show how the new model behaves, pointing out the differences in the description of the behavior of the investigated system when using panel regression analysis and nonlinear regression analysis. Addressing the second research question, i.e., using the implementation of nonlinear optimization methods to find values of input variables at which the risk of legalization of income from criminal activity will be minimal, brings interesting results that are interpreted and discussed in this study.

The remainder of the study is organized as follows. In Section 2, we provide contextual setting and the empirical methodology. The empirical findings and linear model (obtained by panel regression) with corresponding descriptive statistics, analysis and interpretation are discussed in Section 3. Section 4 deals with development of the nonlinear computational mathematical model and its estimation results; and with comparing both of these compiled and suitable models. In addition, the optimization process of determining the minimum level of the ML/TF risk (symbolized by AML index) is conducted and its results reported at the end of Section 4. Finally, the paper is concluded in Section 5.

2. Materials and Methods

2.1. Contextual Setting

These days, in 2022, while we face an energy and economic crisis around the world (implied by the COVID-19 and the war conflict in Ukraine), money laundering and terrorism financing represent an unhealthy economic activity with undesirable effects on the countries' economies. Therefore, there is a need to prevent this unfavorable phenomenon (ML/FC/TF) and simultaneously the need to investigate this problem. The present research study deals with mathematical modelling of the risk of legalization of criminal incomes and terrorist financing. The study primarily focuses on investigating the Basel AML index behavior, as the main indicator of the ML/TF risk level in a certain country [32,33], in dependence on the defined global indices over the time period from 2012 to 2019.

The Basel AML index, defined by the Basel Institute for Governance, was developed in purpose to identify and assess if a certain country is at risk for money laundering or terrorism financing through use of the country's financial institutions. The Basel AML index does not actually measure the volume of money laundering (the actual money amount) or terrorist financing [32,33], but is aimed at assessing the risk of occurrence of such activity (ML/TF) in a given country. ML/TF risk in this sense covers a broad risk area that includes not only a country's vulnerability to ML/TF, but also the country's ability to counter this. To measure the AML index, a 10-point scale is used; where 0 stands for the minimum value of ML/TF risk (0 indicates the absence of risk of ML/TF in a given country), 10 stands for the worst, i.e., the maximum value of ML/TF risk. The rating value of AML index assigned to the country (the resulting score) is determined on the basis of a complex methodology, which takes into account the share of five basic domains, which are specified using 18 indicators. Specifically: Domain 1 (65%)—the quality of anti-money laundering and counter financing of terrorism (quality of AML/CFT framework); domain 2 (10%)—corruption and bribery risk; domain 3 (10%)—financial transparency and standards; domain 4 (5%)—transparency and accountability; and domain 5 (10%)—political and legal risk. More detailed information on the calculation of the resulting AML index (methodology, percentage representation of individual domains and indicators, scaling and standardization of individual indicators, determining the weight of each variable, aggregation of all scores into one score, etc.) are presented and updated (annually) on the website [32].

Finally, the Basel AML index aggregates a wide scale of all 18 indicators, and as a comprehensive and integrated indicator ranks countries based on their overall score, captures the complex global nature of ML/TF risks and provides useful data for comparative purposes. This index serves as a good background for examining progress over time. The primary aim of AML determination is to not only rank countries, but also provide an overall picture of the different risk levels of countries. Obviously, it is difficult to describe and capture all the factors contributing to the high level of the ML/TF risk in a country, but the main areas of risk factors can be simplified as in Figure 1.

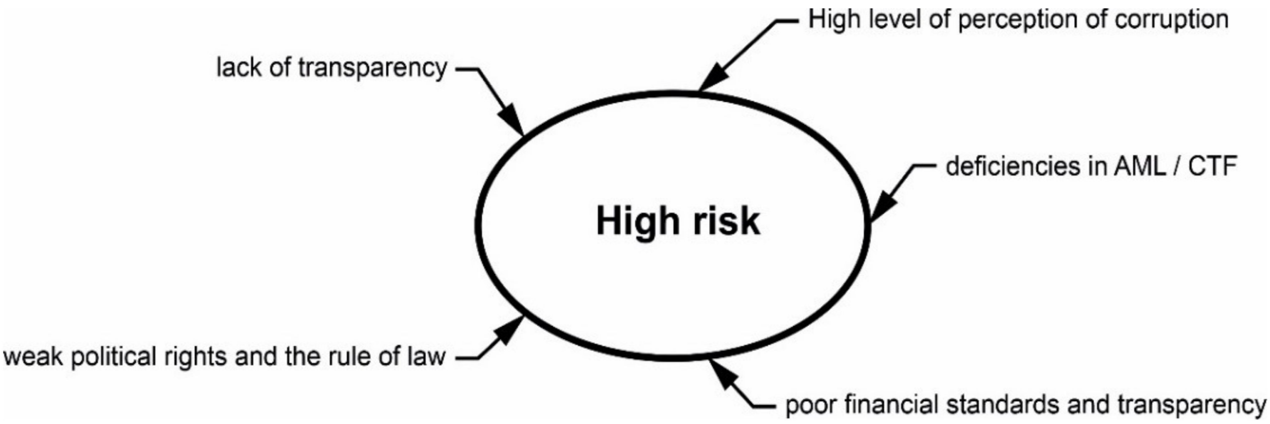

Figure 1. Factors contributing to the high ML/TF risk score. Source: Compiled by the authors.

It is also worth mentioning that a complete set of 18 indicators is not always available for each country. In this case, the total score for the country is calculated only on the basis of available data and missing values are not replaced. A country is included in the public release of the AML index only if data are available for at least two key indicators from domain 1. Data for the individual indicators (components) needed to determine the Basel AML index are drawn from various publicly available sources, such as the Financial Action Task Force (FATF), Transparency International, the World Bank and the World Economic Forum.

2.2. Data

In the study we consider the problem of modelling a response variable y (the Basel AML index) as a function of k input variables $x_1, x_2, \ldots x_k$ based upon a data set consisting of n sets of observed values ($y_i, x_{1i}, x_{2i}, \ldots, x_{ki}$), where $i = 1,2, \ldots, n$ (n—number of destination countries, $n = 24$). Ten input variables ($k = 10$) are under consideration through previous experience and understanding for the factors, namely, the Value Added Tax gap (VAT_{GAP}), the Gross Domestic Product per capita (GDP per capita) and 8 global indices, i.e., World Competitiveness Index (WCI), Corruption Perceptions Index (CPI), Index of Economic Freedom (EFI), Global Innovation Index (GII), Sustainable Economic Development Assessment (SEDA), Doing Business Index (DBI), Global Sustainable Competitiveness Index (GSCI) and Human Development Index (HDI). The study sample covers data with annual frequency for the observed period of 2012 to 2019. The data for this purpose were obtained from the official world indicators' websites, namely: the Basel AML from [39], WCI [40], CPI [41], EFI [42], GII [43], SEDA [44], DBI [45], GSCI [46], HDI [47], VAT_{GAP} [48] and GDP per capita from the official Eurostat website [49]. We accessed all of these indexes during August 2022, and since the VAT gap has available the last data for 2019 in the last report for 2021, this limited us to research in the range of the mentioned years from 2012 to 2019. However, on the websites of world indicators the required data are not available for each EU member country in the monitored period of 2012–2019. Some data (indices) are missing for Cyprus, Malta, Lithuania and Latvia. Therefore, these EU member countries were excluded from the final research sample. It is worth mentioning that these indicators are already aggregated and determined by the appropriate methodology of a given world institution, which publishes statistical reports about these indicators.

Due to the fact that the analytical part of this study concerns the Basel AML index, as the main indicator of the level of risk of legalization of income from criminal activities and terrorist financing, Figure 2 provides an overview of the development of this index in countries of the European Union. The EU member countries under investigation are presented in Figure 2.

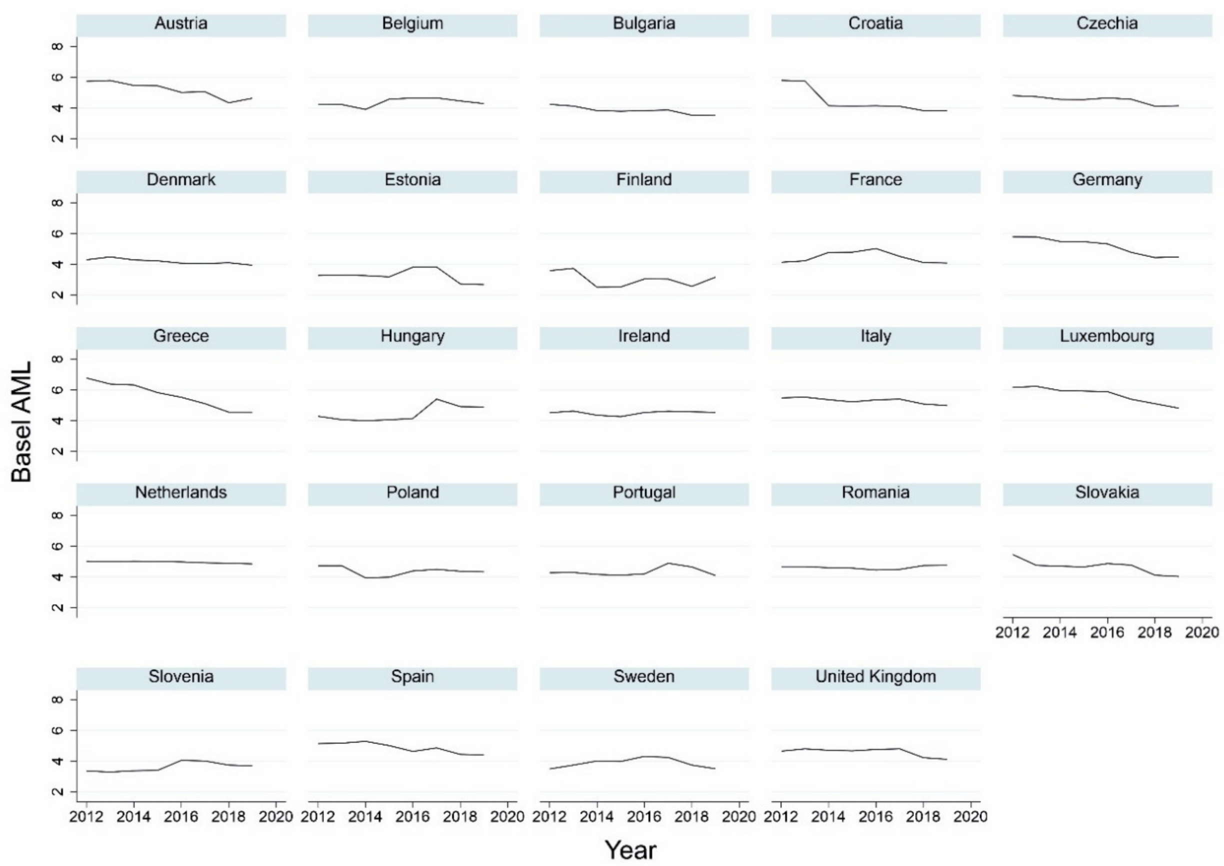

Figure 2. Development of the Basel AML index in 24 EU countries during the observed period of 2012–2019, except Cyprus, Malta, Lithuania and Latvia due to missing data. Source: Compiled by the authors.

We have prepared a panel graph of the development of the Basel AML index within given EU countries based on the data published in [39–50] during the monitored years of 2012–2019. The average value of Basel AML index is 4.441 $\pm$ 0.092 within the EU countries in the monitored period of 2012 to 2019. The standard deviation of the average value of the AML index within EU countries is 0.746. Basel AML index reached its minimum value of 2.36 in Estonia in 2019, and on the contrary, the maximum value of the Basel AML index of 6.78 was achieved in Greece in 2012. The lower quartile of the Basel AML index within EU countries during the period 2012–2019 represents the value 3.985 and the upper quartile represents the value 4.900.

2.3. Methodology

The present study focuses on performing an optimization procedure in order to determine the minimum risk of legalization of income from criminal activities in the context of EU member countries. In other words, to find out the appropriate combination of input factors affecting the Basel AML index that produces its minimum value. When solving such a complex problem, it is required to establish the mathematical formulation of the optimization problem based on the developed mathematical model of the controlled system according to the theory of optimal decision-making [51–56]. During the optimization process aiming to reach the best result, the sequence of individual steps [57] must be followed: (a) mathematical model, (b) objective function, (c) constraint conditions, (d) selection of optimization method, (e) software for processing and (f) result verification.

The first stage in this study of the influence of the global indicators, VAT_{GAP} and GDP per capita on the AML index is to develop a multifactorial regression model using panel data analysis. Primarily, a descriptive statistical approach and secondarily inferential statistical approach are applied to explore the collected econometrical data through a year-by-year and cross-sectional method. In [58], the use of panel data is suggested as more suitable for analysis and estimation model regression coefficients in comparison to other data forms—time series and cross-sectional data. Because of a larger number of observations, which improves the normality and reduces the multicollinearity and serial correlation issues, the panel data occur to be more reliable. However, the basic econometric model established by usage of panel data analysis is linear. In order to also take into account nonlinearities and interactions between input variables (the global indexes, VAT_{GAP} and GDP per capita as factors) the second stage is the application of multidimensional statistical analysis tools to create a nonlinear multivariate regression model of the impact of repressors. In the third stage, based on the nonlinear regression model created, the objective function and boundaries are defined and the nonlinear optimization process is elaborated. Empirical analyses are carried out using advanced software tools, specifically Matlab 2019b, Statistica 13.5 and JMP 11.

2.4. Panel Regression Analysis

Within the framework of the presented study, the set of data mentioned above will be subjected to a panel regression analysis in the first stage. In panel data, there are observations at different points in time and for various investigated subjects (i.e., different countries in this case). The basic framework considered for the multiple linear regression model of panel data including fixed and time effects is model (1) of the generic form (more detailed in [54]):

$$y_{it} = \beta_1 x_{it1} + \beta_2 x_{it2} + \ldots + \beta_k x_{itk} + \alpha_1 z_{i1} + \alpha_2 z_{i2} + \ldots + \alpha_q z_{iq} + \varepsilon_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{z}_i\boldsymbol{\alpha} + \varepsilon_{it} \quad (1)$$

where y_{it} stands for the value of the response (explained variable) y; index i denotes the cross-sectional dimension $i = 1, \ldots, n$, $n = 24$ of destination countries; index t denotes the time dimension $t = 1, \ldots, T$ (over the period 2012–2019) and $\boldsymbol{\beta}$ labels a vector of regression parameters. In this setting y is called the regressand. The variables $\mathbf{x}_1$ to $\mathbf{x}_k$ are input (explanatory) variables (so-called regressors or covariates) not including the unit vector, and $\mathbf{z}_1$ to $\mathbf{z}_q$ represent individual effects characterizing individual countries (a possible vector of units is included here) [54]. Individual effects do not change over time. The observed value of y_{it} is the sum of a deterministic part and the random part ε_{it}—a random disturbance [54]. Model (1) specifies a linear relationship between response and regressors. The panel data estimations must be subjected to several data analysis procedures: comparative descriptive statistics, comparative unit root testing and panel regression estimations, such as fixed, random, and pooled regression estimations. As in simple linear regression, the unknown parameters in multiple linear regression are estimated using the method of least squares, which produces the maximum likelihood estimates of the parameters (the parameter estimates are the solutions of the normal equations).

Based on the abovementioned, there are a variety of different models for panel data, which can be arranged as follows [59–61]:

1. Pooled regression model (PRM)—if the individual effect $\mathbf{z}_i$ is only a vector of units (contains only a constant term), this means that only parameter α is a common constant (omitted effects)

$$y_{it} = \alpha + \beta_1 x_{it1} + \beta_2 x_{it2} + \ldots + \beta_k x_{itk} + \varepsilon_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \alpha + \varepsilon_{it}; \quad (2)$$

   PRM in its general form is the simplest case and is based on the assumption that the absolute term and all parameters of the explanatory variables are the same for all cross-sectional units;

2. Fixed effects model (FEM)—if individual effects z_1 to z_q are unobservable, but correlated with $\mathbf{x}_{it}$, i.e., explanatory variables, then the solution is to include all effects in an estimable conditional mean using the relationship $\alpha_i = \alpha_1 z_{i1} + \alpha_2 z_{i2} + \ldots + \alpha_q z_{iq}$ and the FEM model may be formulated as

$$y_{it} = \alpha_i + \beta_1 x_{it1} + \beta_2 x_{it2} + \ldots + \beta_k x_{itk} + \varepsilon_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it}; \tag{3}$$

where the fixed effect $\alpha_i = \mathbf{z}_i\boldsymbol{\alpha}$ stands for a specific constant for each cross-sectional unit;
3. Random effects model (REM)—can be viewed as a regression model with a random constant term, i.e., if individual effects z_1 to z_q are unobserved and uncorrelated with explanatory variables, then the solution is to use a composite random component, which, in addition to the original one, also assumes a specific random component for each cross-sectional unit. Then the REM model might appear as

$$y_{it} = \beta_1 x_{it1} + \beta_2 x_{it2} + \ldots + \beta_k x_{itk} + (\alpha + u_i) + \varepsilon_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \alpha + u_i + \varepsilon_{it} \tag{4}$$

The basic panel data model (PRM) within the study can be expressed as follows

$$AML = \alpha + \beta_1 WCI + \beta_2 CPI + \beta_3 EFI + \beta_4 GII + \beta_5 SEDA + \beta_6 DBI + \beta_7 GSCI + \beta_8 HDI + \beta_9 VAT_{GAP} + \beta_{10} GDP_{per\_cap} \tag{5}$$

To decide between FEM and REM, the Hausman specification test [62] needs to be implemented, based on its significance, estimations are made using the fixed effects model. When FEM is confirmed, other procedures are required (testing some assumptions, autocorrelation, heteroscedasticity, etc.), which are more detailed in [63–65].

2.5. Nonlinear Regression Model Development

There are many applications where linear regression modelling is very useful, but sometimes it is necessary to fit a nonlinear model to a data set, above all if the data set indicates that there is some curvature or interaction in the relationship between response and input variables. To achieve our research goal of developing the nonlinear regression (mathematical) model of complex (socio-econometric) data with explanatory variables and the error term, the estimated model can be written in general form [59,66,67]:

$$Y_t = f(X_{1,t}, X_{2,t}, \cdots, X_{p,t}) + \varepsilon_t \tag{6}$$

where $p = 1, 2, \cdots, p$ and $t = 1, 2, \cdots, T$ represents the range of indexes and ε_t denotes the error term. The set $(X_{1t}, X_{2t}, \cdots, X_{pt}) \in \mathbb{R}$ is the set of explanatory variables and $Y = (Y_t) \in \mathbb{R}$ symbolizes the response variable. Practically, scholars need to develop the relatively precise mathematical model of socio-econometric data (or other data). To illustrate this, let us build the function $f : \{X\} \to \{Y\}$ and consider fitting the correct shape of $f(\circ)$. Since it is difficult to fit the correct form of $f(\circ)$ [66,68], the wrong model form is fitted sometimes. The mistake of a wrong model shape can be expressed by δ_t in a modified version of Equation (7):

$$Y_t = f_m(X_{1,t}, X_{2,t}, \cdots, X_{p,t}) + \delta_t + \varepsilon_t \tag{7}$$

To achieve the relatively precise mathematical model of observed data, δ_t should be small enough so that $f_m(\circ)$ then approximates $f(\circ)$. In regression model (6), f is an arbitrary continuous function. If $f(\circ)$ is a non-parametric function—we obtain a non-parametric model, if $f(\circ)$ is a parametric function—we have a parametric model, which includes classic linear and nonlinear models. Searching the regression function is conditioned by relation (8), where $\beta_j, j = 0, 1, \cdots, k$ are the regression coefficients (unknown parameters).

$$Y = g(X, \beta_0, \beta_1, \cdots, \beta_p) = E(Y|\, X) \tag{8}$$

Based on formula (8), where $E(Y|\, X)$ expresses the dependence of the conditional mean values of the random variable Y on the values of the random variable X, in other words, the expected value of the response value of the response at $\mathbf{x} = (x_1, \ldots, x_k)$ will be $E(Y|_x) = \beta_0 + \beta_1 x_1 + \cdots \beta_k x_k$ in the case of a multiple linear regression model. It is clear that there are several options to create a regression model (e.g., several types: linear, quadratic, polynomial, hyperbolic, logarithmic, etc.) in accordance with (6) by usage of certain mathematical formulas. The most suitable method for estimating the unknown parameters $\beta_0, \beta_1, \ldots, \beta_k$ of the regression curve is the method of least squares, which uses the assumption that the sum of squares of the residual errors of the model will be as small as possible when approximating the regression function. Within the framework of the article, to a nonlinear regression model, the estimate of the investigated dependent variable is generally described by a polynomial equation (of order n) [59,69]:

$$\widehat{y} = \beta_0 + \sum_{j=1}^{n} \beta_j x_j + \sum_{\substack{u,j=1 \\ u \neq j}}^{n} \beta_{uj} x_u x_j + \sum_{\substack{u,j=1 \\ u \neq j}}^{n} \beta_{uj} x_u^2 x_j + \cdots + \sum_{\substack{u,j=1 \\ u \neq j}}^{n} \beta_{uj} x_u x_j^n + \sum_{j=1}^{n} \beta_{jj} x_j^n \tag{9}$$

When deciding on the "best", "the most correct" regression model, some essential facts must be taken into account [70–72]. Mathematically, the best way to deal with minimizing the corresponding sum of least squares when estimating unknown parameters of the regression model (labelled as Q and also known as the normal equations) and any subsequent inferences concerning the regression model is to use matrix algebra. In order to fit the "most accurate" multiple linear or nonlinear regression model and determine which input variables are needed in the regression model (or can be "dropped" from the model), the analysis of variance to test the hypotheses $H_0 : \beta_i = 0$ versus the alternative $H_A : \beta_i \neq 0$ (at least one of this β_i is nonzero) can be performed by using computer software packages. It is necessary to calculate p-values, F-statistic, R^2 (coefficient of multiple determination), Rsquare Adj, RMSE, AIC and other coefficients and information criterions. In detail:

- The model must be adequate in terms of the Fisher–Snedecor test criterion. The significance of the entire regression model is assessed based on the F-test for hypothesis (9), using test-statistics (10) and the critical region of the F-test (11).

$$\begin{array}{l} H_0 : \beta_0 = k,\ \ k \neq 0,\ \ \beta_1 = \beta_2 = \cdots \beta_j = 0 \\ H_A : \beta_j \neq 0 \ \text{ at least for one } j = 1, 2, \cdots p \end{array} \tag{10}$$

The test criterion is statistics:

$$F = \frac{\frac{S_T(y)}{c-1}}{\frac{S_e(y)}{n-c}} \sim F(c-1, n-c) \tag{11}$$

$$W_\alpha : \left| t_j \right| \geq t_{1-\frac{\alpha}{2}}(n-c) \tag{12}$$

where $S_T(y) = \widehat{\mathbf{b}}'\mathbf{X}'\mathbf{Y} - n\overline{y}^2$ represents the part of the total sum of squared deviations that is explained by regression model (9) and $S_e = \mathbf{Y}'\mathbf{Y} - \widehat{\mathbf{b}}'\mathbf{X}'\mathbf{Y}$ represents that part of the total sum of squares of deviations that is not explained by regression model (2). However, it is necessary to consider some fundamental facts [70–73]:

- Rsquare (also labeled as R^2)—index of determination, defined as the proportion of the variability that regression model (8) is able to describe to the total variability of the explained variable. It takes values from 0 to 1, a larger value of the index indicates a more appropriate model [72].
- Rsquare Adj—adjusted index of determination adjusts the value of Rsquare in order to make it independent of the number of used measurements [71,72].
- Root mean square error (RMSE)—an estimate of the standard deviation of the random error; the lowest value of $RMSE$ indices a more appropriate model [70–72].
- AICc—corrected Akaike's information criterion $AIC = 2k - 2\ln(L)$; the best model has the smallest value as shown more detailed in [73].
- BIC—the Bayesian information criterion (BIC) is defined as $BIC = -2\log\text{likelihood} + N\ln(n)$; where N is the number of estimated model parameters and n is the number of observations in the data set. From the set of possible models, we choose the one that achieves the lowest BIC value;
- HQC—Hannan–Quinn's information criterion; represents an alternative to Akaike's and Bayes's information criterion. $HQC = -2L_{\max} + 2k\ln(\ln(n))$, where $L_{\max}$ is the log-likelihood, k is the number of parameters and n is the number of observations [71,73].

Therefore, if we consider a set of possible regression models that can be created by using relation (9), we choose the best one that meets the mentioned criteria and then we use the model selected in this way to describe the investigated dependence. We performed the model fitting procedure by applying Statistica 13.5, Matlab2019b and JPM11, which provide the opportunity to implement a stepwise procedure that combines the backward elimination and forward selection procedures, inferences on the response variable, residual analysis and other techniques. The estimation of significant regression parameters is conducted (at chosen 95% confidence interval) and the nonlinear regression model is developed. The obtained empirical results are discussed in the Section 4.

2.6. Nonlinear Mathematical Optimization

Within the optimization procedure, our effort will be to minimize the level of the ML/TF risk symbolized by the Basel AML index. When minimizing an objective function $f_0(x)$ in a feasible region K, the mathematical optimization problem (OP) can be formulated in the following form [57]:

$$Min\{f_0(x) \ | x \in K \subseteq \mathbb{R}^n\} \quad \text{where} \quad f_0 : X_0 \subseteq \mathbb{R}^n \to \mathbb{R}. \tag{13}$$

In case that $K = \{x \in X | f_i(x) \leq 0, i = 1, 2, \ldots, p\} \neq 0$, where $f_i : X_i \to \mathbb{R},\ i = 0, 1, 2, \ldots, p$ and $\neq X_i \subseteq \mathbb{R}^n,\ X = \bigcap_{i=0}^{p} X_i$, the OP in the narrower sense can be written as (14):

$$Min\{f_0(x) \ | x \in X, f_i(x) \leq 0, i = 1, 2, \ldots, p\}. \tag{14}$$

If $x \in X \subset \mathbb{R}^n$, we discuss constrained optimization, for $x \in \mathbb{R}^n$ regarding unconstrained optimization. Presence of the equality and inequality constraints in OP can be expressed in the following way

$$Min\Big\{f_0(x) \ | x \in X, f_i(x) \leq 0,\ i \in I,\ h_j(x) = 0,\ j \in J\Big\}, \tag{15}$$

which is usually called a mathematical programming problem (MP) in the broader sense [52]. In addition to the p inequality constraints $f_i(x) \leq 0,\ i \in I = \{1, 2, \ldots, p\}$ r constraints in the form of equations $h_j(x) = 0,\ j \in J = \{p+1, p+2, \ldots, p+r\}$ are also considered, where I, J are index sets. If at least one of the functions $f_0,\ f_i,\ i \in I,\ h_j,\ j \in J$ in (15) is nonlinear, we discuss the nonlinear programming problem (NLP).

To find the optimal solution to the problem, (15) is used to find such a vector, $x^* \in X$, which assumes the smallest value among all the vectors in the feasible region, i.e., applies the formula $\forall x \in X : f_0(x^*) \leq f_0(x)$. It is obvious that vector $x = (x_1, x_2, \ldots, x_n)^T$ is the n-dimensional vector of design variables (vector of independent variables, also known as design vector). The optimization problem involves one or more objectives and contains a finite number of equality and inequality constraints defining a feasible region X, in dependency on application of mathematical optimization.

Due to development of a nonlinear regression model describing the behavior of the AML index, the objective function is expected be also be nonlinear. Hence, nonlinear programming is implemented to perform the optimization procedure by usage of advanced tools of the Matlab 2019b software system. It should be mentioned that the "fmincon ()" solver for constrained nonlinear minimization and the interior point method (IPM) algorithm are used within the framework of this article. We formulate the mathematical optimization problem, the objective function and constraints for the minimizing ML/FT (represented by AMK index) in Section 4, and rewrite OP in a format suitable for the Matlab2019b software environment.

3. Empirical Analysis and Results

3.1. Panel Regression Analysis

The research requires exploring the relationship between the Basel AML (response) and the input variables for the target population for the considered period of study. For this purpose, the study examines the impact of the following global indices on the change in the AML value, namely: the World Competitiveness Index (WCI); Corruption Perceptions Index (CPI); Index of Economic Freedom (EFI), Global Innovation Index (GII); Sustainable Economic Development Assessment (SEDA); Doing Business Index (DBI); Global Sustainable Competitiveness Index (GSCI) and Human Development Index (HDI). The Value Added Tax Gap (VAT_{GAP}) and Gross Domestic Product per capita (GDP per capita) are also considered as input independent parameters when examining the response within target EU member countries (as seen in Figure 2).

3.1.1. Pooled Regression Model (PRM)

Statistical analysis of collected data was performed. Table 1 reports the main results of descriptive statistics useful for developing the pooled regression model.

Table 1. The regression coefficients estimation of the pooled regression model for Basel AML index within EU destination countries.

Variable	Coefficient	±95% CI for Coefficient		Std. Error	*t*-Ratio	*p*-Value
Const.	7.287540	3.084830	11.490200	2.129940	3.421000	0.000800 *
WCI	0.018079	0.002133	0.034026	0.008082	2.237000	0.026500 *
CPI	−0.0409183	−0.060025	−0.021812	0.009683	−4.226000	0.000038 *
EFI	−0.0292790	−0.056029	−0.002529	0.013557	−2.160000	0.032100 *
GII	0.030716	0.002313	0.059119	0.014395	2.134000	0.034200 *
SEDA	0.045715	0.001853	0.089576	0.022229	2.057000	0.041200 *
DBI	−0.0354838	−0.069959	−0.001009	0.017472	−2.031000	0.043700 *
GSCI	−0.0426950	−0.072656	−0.012734	0.015184	−2.812000	0.005500 *
HDI	−0.00154235	−0.063128	0.060044	0.031212	−0.049420	0.960600
VAT_{GAP}	20.987300	9.219740	32.754800	5.963810	3.519000	0.000500 *
GDP per Capita	0.000016	0.000006	0.000026	0.000005	3.290000	0.001200 *

*—factor is significant at level of α = 0.05. Source: calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

Having used the estimated regression coefficients, presented in Table 1, the PRM—pooled regression model—has the form (16):

$$\begin{aligned} AML = 7.2875 + 0.018 \cdot WCI - 0.0409 \cdot CPI - 0.0293 \cdot EFI + 0.0307 \cdot GII + \\ +0.0457 \cdot SEDA - 0.0355 \cdot DBI - 0.0427 \cdot GSCI + 20.9873 \cdot VAT_{GAP} + 1.609 \cdot 10^{-5} \cdot GDP_{per\_cap} \end{aligned} \tag{16}$$

The average value of the analyzed Basel AML parameter within the data set of EU member countries mentioned above, i.e., excluding Cyprus, Malta, Lithuania and Latvia due to missing data and in the observed period of 2012–2019, is 4.5058 with a standard deviation of 0.7585. Based on the achieved level of significance of the Fisher–Snedecor test criterion $p = 7.14 \times 10^{-18}$ at the value of the Fisher test statistic $F(10.181) = 13.74507$, it is possible to say that model (16) is adequate, so there is at least one regression coefficient of the used factor, which is different from zero. For the adjusted index of determination, we obtained the value Rsquare Adj = 40.0221%. The individual values of the information criteria are as follows: Bayes–Schwarz information criterion is 387.5760, Akaiko's information criterion is 351.7436 and Hannan–Quinn's information criterion reaches the value of 366.2560.

Table 1 shows that when increasing the values of the variables WCI, GII, SEDA, VAT_{GAP} and GDP, the conditional value of the Basel AML index also increases, which indicates the increasing level of the ML/TF risk. The summarized findings from descriptive statistics indicate, that the value of the Corruption Perception Index (CPI) has the greatest effect on the change of the Basel AML index value, this impact represents 15.127%. The second most significant impact on the change in the value of the dependent variable Basel AML is the change in the value of the VAT_{GAP}, while this impact represents 12.596%. The third most significant factor that affects the change in the value of the Basel AML index is the value of GDP per capita, this factor has an 11.777% impact. At this point, it is necessary to mention that the influence of the intercept of the model (16) represents 12.246% share, what can be interpreted as the effect of influences that we did not consider when building the pooled regression model.

3.1.2. Fixed Effects Model (FEM)

The Fixed Effects Model unlike the PRM, assumes a diversity of cross-sectional units in the absolute terms. The basic fixed effects model (FEM) has the form:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} i & 0 & \cdots & 0 \\ 0 & i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & i \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \boldsymbol{\beta} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \mathbf{D}\boldsymbol{\alpha} + \mathbf{x}\boldsymbol{\beta} + \mathbf{u}. \tag{17}$$

As seen in the model (16) the columns of matrix **D** represent dummy variables D_1—D_n, which take the value $d_{it} = 1$ for the $i - th$ cross-sectional unit, and the value $d_{it} = 0$ for all other cross-sectional units. Table 2 reports the panel estimation results for regression coefficients of FEM model for Basel AML index within EU destination countries. The significant factors (significant for changing the value of investigated parameter Basel AML) at a 95% confidence level are labeled by *.

Table 2. The regression coefficients estimation of Fixed Effects Model for Basel AML index within EU destination countries.

Variable	Coefficient	±95% CI for Coefficient		Std. Error	*t*-Ratio	*p*-Value
Const.	24.381700	12.447600	36.315900	6.042310	4.035000	0.000085 *
WCI	0.011358	−0.003031	0.025747	0.007285	1.559000	0.121000
CPI	−0.042407	−0.068935	−0.015878	0.013432	−3.157000	0.001900 *
EFI	−0.013366	−0.060349	0.033617	0.023788	−0.561900	0.575000
GII	0.048148	−0.002876	0.099172	0.025834	1.864000	0.064200
SEDA	0.031625	−0.045133	0.108383	0.038863	0.813800	0.417000
DBI	−0.028584	−0.072078	0.014910	0.022021	−1.298000	0.196200
GSCI	−0.000797	−0.028133	0.026539	0.013840	−0.057570	0.954200
HDI	−0.213211	−0.356425	−0.069997	0.072510	−2.940000	0.003800 *
VAT_{GAP}	−35.616700	−54.147800	−17.085600	9.382390	−3.796000	0.000200 *
GDP per capita	−0.000003	−0.000030	0.000024	0.000014	−0.214900	0.830100

*—significant at level of α = 0.05. Source: calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

Having used the regression coefficients estimated, presented in Table 2, the FEM—Fixed Effects Model (18) can be written as follows:

$$AML = 24.3817 - 0.0424 \cdot CPI - 0.2132 \cdot HDI - 35.6167 \cdot VAT_{GAP} \quad (18)$$

The average value of the analyzed *Basel AML* parameter in the data set, i.e., EU countries excluding Cyprus, Malta, Lithuania and Latvia due to missing data and in the observed period of 2012—2019, is 4.5058 with a standard deviation of 0.7585. Based on the achieved level of significance of the Fisher–Snedecor test criterion $p = 3.37 \times 10^{-43}$ at the value of the Fisher test statistic F (33.158) = 21.55307, it can be concluded that model (18) is adequate, so there exists at least one non-zero regression coefficient of the used factor. Adjusted index of determination (Adjustet R^2) reaches the value Rsquare Adj = 30.125%. The individual values of the information criteria are as follows: Bayes–Schwarz information criterion is 289.6056, Akaiko's information criterion is 178.8508 and Hannan–Quinn's information criterion reaches the value of 223.7073.

According to the results of descriptive statistics shown in Table 2, we can state that the value of VAT_{GAP} has the greatest influence on the change in the value of the Basel AML index, where this influence represents 18.702%. The second most significant influence on the change in the value of Basel AML as the dependent variable is the change in the value of the Corruption Perceptions Index (CPI), where this share represents 15.554%. The third most important factor that affects the change in the value of the Basel AML index is the value of the Human Development Index (HDI), where this factor has a 14.485% impact. At this point, it is necessary to mention that the influence of the intercept of model (9) represents 19.880%, which can be interpreted as the effect of influences that we did not consider when building the model.

The difference between individual cross-sectional units is tested using an *F*-test comparing both models, FEM and PRM. The test statistic has the form of Equation (19):

$$F = \frac{\frac{RSS_{POL} - RSS_{FEM}}{n-1}}{\frac{RSS_{FEM}}{nT-k-n}} \quad (19)$$

and we compare it with the table value $F(n-1, nT-k-n)$ at the chosen level of significance $\alpha = 0.05$. If the value of the statistic is greater than the table value, we reject the null hypothesis stating that the cross-sectional units have equal absolute terms. Based on the value of the achieved level of significance $p = 2.1007 \times 10^{-28}$ at the value of the Fisher test statistic F (23.158) = 14.6114, we can reject the null hypothesis which states that the cross-sectional units have the same absolute terms.

Because of the presence of dummy variables in the FEM, this model is also called the least squares dummy variable model (LSDV). We can then estimate the model as a regression model without a constant, or one cross-sectional unit is chosen form the main group (for example, the third) whose value will be represented by the absolute term in the LSDV model, so we will use only (n−1) dummy variables. In such a case, the basic fixed effects model expressed by Equation (17) will be modified into the form $\mathbf{y} = \alpha_1 + \mathbf{D}_1\boldsymbol{\alpha}^* + \mathbf{x}\boldsymbol{\beta} + \mathbf{u}$; where the submatrix $\mathbf{D}_1$ is created from the original matrix $\mathbf{D}$ by omitting the first column and $\boldsymbol{\alpha}^*$ is a vector of dimension (n−1); $\boldsymbol{\alpha}^*$ is a vector of differentiating absolute terms relative to the absolute term of the fundamental group.

3.1.3. Random Effects Model (REM)

This model is defined by Equation (4). When combining the random component of a particular observation in the cross-sectional unit u_{it} and the random component specific to the cross-sectional unit ε_i, we obtain the composite random component v_{it}. The absolute term α in REM model (4) represents the average of the cross-sectional absolute terms in the model, and the random component specific to the cross-sectional unit is the random deviation from this average.

Having used the regression coefficients estimated, presented in Table 3, the REM—random effects model (20) can be written by formula:

$$AML = 14.2353 - 0.0396 \cdot CPI + 0.0641GII - 0.0467 \cdot DBI - 0.2132 \cdot HDI \quad (20)$$

Table 3. The regression coefficients estimation of random effects model for the Basel AML index within EU destination countries.

Variable	Coefficient	±95% CI for Coefficient		Std. Error	z	p-Value
Const.	14.235300	7.528540	20.942000	3.398980	4.188000	0.000028 *
WCI	0.012070	−0.002055	0.026196	0.007159	1.686000	0.091800
CPI	−0.039648	−0.063514	−0.015782	0.012095	−3.278000	0.001000 *
EFI	−0.034927	−0.070278	0.000425	0.017916	−1.949000	0.051200
GII	0.064064	0.029054	0.099073	0.017743	3.611000	0.000300 *
SEDA	0.046734	−0.011836	0.105304	0.029683	1.574000	0.115400
DBI	−0.046677	−0.084545	−0.008808	0.019192	−2.432000	0.015000 *
GSCI	−0.001176	−0.027667	0.025315	0.013426	−0.087570	0.930200
HDI	−0.096662	−0.191158	−0.002166	0.047891	−2.018000	0.043600 *
VAT_{GAP}	−13.138400	−29.516900	3.240090	8.300650	−1.583000	0.113500
GDP per capita	0.000005	−0.000012	0.000021	0.000008	0.555100	0.578800

*—significant at level of α = 0.05. Source: calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

In accordance with the reported results in Table 3, it can be said that the value of the Global Innovation Index (GII) has the greatest impact on the change in the value of Basel AML index, this influence represents 15.730%. The change in the CPI value shows the second most significant effect on the change of the Basel AML value, this impact represents 14.280%. The third most important factor is the DBI value with its impact of 10.590%. The influence of the model's intercept, labeled in Table 3 as *Const.*, represents an 18.240% impact, which can be interpreted as being caused by non-considered influences when building the model. The values of the individual information criteria, namely, the Bayes–Schwarz, Akaiko's and Hannan–Quinn's information criteria are 458.8840, 423.0515 and 437.5639, respectively.

To decide between the FEM and REM models, the Hausman specification test is used, the test statistic of which is displayed below the model estimate at the end of the listing. The null hypothesis assumes that the parameter estimates of the generalized least squares method in the REM model and the least squares method in the FEM model are consistent, and thus the least squares estimation is not efficient. In the alternative hypothesis, only the method of least squares is consistent. If $H > \chi_c^2$ is valid, then we can reject the null hypothesis about the consistency of both estimators and the FEM model is more suitable. If inequality $H < \chi_c^2$ for the statistical value is valid, then we cannot reject the null hypothesis and the REM model is recommended. For our data set, the value of the achieved level of significance is $p = 0.00002321$, and thus the pooled regression model (16) or FEM model (18) is more suitable for describing the change in the value of the Basel AML index.

Within the PRM model (16), which can be considered as a model appropriately describing the change in the Basel AML index value, the CPI index (Corruption Perception Index) is an independent variable that significantly describes the change in the Basel AML index. The total CPI score ranges from 0 to 100. Each of the sources included in the CPI is standardized in such way to allow aggregation into a CPI score. Standardization converts all data points to a scale of 0–100, where 0 represents the highest level of perceived corruption and 100 the lowest level of perceived corruption. Regarding the CPI values, it is necessary to mention that the method of calculating the CPI index was changed in 2012; therefore, according to Transparency International, the values from 2012 are not comparable with the values from 2011 and earlier. However, they basically measure equally—the perception of corruption on the same, or convertible scale (until 2011 inclusive on a scale of 0–10 to two decimal places, from 2012 on a scale of 0–100). The Corruption Perceptions Index (CPI) focuses on the perception of the existence of corruption among public administration officials and politicians and defines corruption as the abuse of public authority for personal gain. Due to the definition of the independent variable CPI, a close link between this index and the Basel AML index is assumed. From the nature of the indexes themselves, this dependence appears to be natural. By increasing CPI value, the index declares a lower value for the perception of corruption, and conversely, lower values of the AML index declare a lower risk of money laundering and terrorist financing. This logical relationship also declares a negative sign for the regression coefficient of model (9). The influence of the CPI index on the change in the value of the dependent variable of the Basel AML index represents a value of 15.127%.

Another important factor that affects the change in the value of the risk of money laundering and terrorist financing is the size of the tax gap (VAT_{GAP}) as an indicator of tax crimes in the form of tax evasion. VAT_{GAP} is the difference between the potential revenue from a particular activity and the current tax collection. The source of the gap is non-compliance with tax regulations (compliance gap) and a legislative gap (policy gap) in the form of tax advantages or exemptions, compared to a hypothetical tax structure in the form of the widest possible tax base. The effect of VAT_{GAP} amount on the change in the value of the monitored Basel AML index is 12.596%. From the point of view of influence, it is clear that with the increasing value of the size of the tax gap as tax evasion, the conditional value of the Basel AML index also increases, i.e., the risk of money laundering naturally increases. Likewise, the value of GDP per capita is also a significant factor that implies value changes of the studied variable (AML). The impact of this factor is 11.770% and it can be stated that with increasing the value of GDP per capita, the risk of money laundering and terrorist financing also increases. The World Competitiveness Index (WCI) affects the AML value, its effect symbolizes a level of 8.007%. The WCI is determined on a scale from 0 to 100 and contains four basic assessed dimensions. Economic performance, government efficiency, business efficiency and infrastructure. The Economic Freedom Index (EFI) is another significant factor that affects the change in the value of the studied Basel AML index. EFI is measured based on 12 quantitative and qualitative factors aggregated into four broad categories (or pillars) of economic freedom:

- Rule of law (property rights, government integrity, and court effectiveness);
- Size of government (government spending, tax burden, fiscal health);
- Effectiveness of legal regulations (freedom of business, labor freedom, monetary freedom);
- Open markets (freedom of trade, freedom of investment, financial freedom).

Each of the twelve economic freedoms (property rights, judicial effectiveness, government integrity, tax burden, government spending, fiscal health, freedom of business, freedom of labor, monetary freedom, freedom of trade, freedom of investment, financial freedom) within these categories is rated on a scale from 0 to 100. A country's total score is derived from the average of these twelve economic freedoms, each of which is given equal weight. The Economic Freedom Index (EFI) affects the value change of the Basel AML index by 7.732%.

The Global Innovation Index (GII) influences the AML by a share of 7.639%. The GII is scaled from 0 to 100 as an arithmetic average of 80 indicators. The Sustainable Economic Development Assessment (SEDA) affects the change in the Basel *AML* index by a share of 7.363% and is determined on a scale from 0 to 100. It consists of three categories (economy, investments and sustainability) which are divided into 10 dimensions and consist of 40 indicators. The Doing Business index (DBI) influences the AML index at the level of share of 7.270%. DBI is assessed on a scale from 0 to 100 and contains 10 assessment dimensions in total. The methodology for quantifying the score is set in such a way that the percentile, in which the economy is located, is calculated for individual indicators firstly. The arithmetic mean is then calculated from the results, which gives us information about the average percentile for each monitored dimension of the business environment. The resulting ranking of the countries is determined by re-averaging these average percentiles of the monitored dimensions for each economy and mathematically ordering them in ascending order of percentiles. Despite the obtained value of Adj R^2 of considered multiple linear regression model (16) not being very satisfactory for us, it can provide us with basic information about the influence of individual regressors. However, in order to better understand the influence of predictors on the response, we will extend model (16) by nonlinear regression analysis with mutual interactions of its individual members.

4. Nonlinear Modelling and Optimization—Results and Discussion

Nonlinear Regression Model Development

Firstly, it is worth introducing some reasons for the nonlinear modelling approach to investigate the relationship between the ML/TF risk (symbolized by the AML index) and the mentioned global indices. Firstly, panel data analysis provides only linear dependencies of the investigated variables according to the type of model used. However, real economic situations are often nonlinear [74–76].

The era of the industrial age (characterized by vertical integration, economies of mass production, hierarchical organization based on command and control) is on the wane. It is being replaced by cooperation with external suppliers, minimization of seriality, profit centers and network structures that are related to the globalization of markets. For the management of the industrial age, it was typical that modelling of economical relationships was based on the assumptions of linearity, equilibrium and a high degree of quantification, it was like a dominant paradigm. On the contrary, nowadays the wealth creation system of the third wave is dominant (characterized by hyper-competition, a series of technological revolutions, social displacements and conflicts), which results in dominated of considerable unpredictability and the nonlinear nature of economies. These are other reasons why we proceeded to create a nonlinear model for estimating the value of the studied Basel AML index depending on the previously mentioned regressors. The already known regression models (PRM, FEM and REM) give very unsatisfactory results. It was our intention to highlight this. Nevertheless, they are popularly applied in econometrics. Although multiple linear regression modelling provides useful models for many applications, sometimes it is necessary to apply nonlinear regression techniques to develop models that cannot be transformed into a linear format. The assumption of linearity between related variables when modelling economic phenomena often leads to a distorted explanation of the real relationships within applied models. In this study we provide an example of such a situation.

It was our previous analyses that led us to the assumed regression model (polynomial multivariate model with cross product or interaction). Skills at modelling are developed primarily through experience, because every problem has its own unique characteristics due to the types of variables considered and the relationships between them. It is generally known, that when developing multiple linear regression or nonlinear regression models, the key task is to identify which of the input variables are needed to provide the best fit (which subset of the k input variables is required to model the dependent variable y in the best and most useful manner). The mathematics of the model fitting procedure has been performed (we have also applied computer packages with implementation of a stepwise procedure combining the backwards elimination modelling procedure and forward selection procedures). Based on this, the predictors were determined: eight global indices, the amount of the tax gap and the gross domestic product per capita value.

Table 4 presents a summary of the results of the statistical analysis focused on the creation of an appropriate statistical-mathematical model expressing the influence of individual factors on the change in the value of the Basel AML index.

Table 4. The summary statistics of nonlinear model of the Basel AML index.

Parameter	Value
RSquare	0.702151
RSquare Adj	0.661374
Root Mean Square Error	0.441947
Mean of Response	4.505885
Observations (or Sum Wgts)	192

Source: Compiled and calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

Table 4 reports that the proportion of the variability of the explained variable Basel AML that the regression model (21) is able to describe to the total variability of AML (RSquare) is 70.2151%. The adjusted index of determination (RSquare Adj) conditioning the degree of explanation of data variability by the given model reaches a value of 66.1374%. At this point, it should be said that due to the relatively high value of the modified determination index, it can be stated that there is a strong mutual relationship between the explanatory variables and the explained parameter (the Basel AML index). The model average error is 0.4419 and the average value of the Basel AML index is 4.5058. The analysis of variance (ANOVA) of the observed data is reported in Table 5. Based on Table 5, it can be concluded that the variability caused by random errors is significantly smaller than the variability of the measured values explained by the model, and the value of the level of significance reached (Prob > F) indicates the adequacy of the used model, based on Fisher–Snedecor's test criterion. The reason is the very nature of the test. The null statistical hypothesis that is being tested states that none of the effects (members) in the model have an impact on the value of the variable under investigation. Since we are working with a significance level of 5% and the achieved value of Prob > F is smaller than the significance level, we can therefore say that there is at least one non-zero term in the model that has an impact on the value of the investigated variable.

Table 5. ANOVA for nonlinear model of the Basel AML index.

Source	DF	Sum of Squares	Mean Square	*F* Ratio	Prob > *F*
Model	23	77.3682	3.36383	17.2297	0.0001 *
Error	168	32.79945	0.19523		
C. Total	191	110.1677			

*—significant at level of $\alpha = 0.05$. Source: Compiled and calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

After fulfilling the basic requirement when creating the model, i.e., the model is adequate, we can create a table of estimates of model parameters, which is presented in Table 6. For the sake of clarity of data and subsequent optimization, we will use the following labels for the input independent variables in the article: x_1—WCI, x_2—CPI, x_3—EFI, x_4—GII, x_5—SEDA, x_6—DBI, x_7—GSCI, x_8—HDI, x_9—VAT_{GAP} and x_{10}—GDP per capita. In order to eliminate the influence of the scale of the factors, especially for the size of the tax gap (VAT_{GAP}) and gross domestic product per capita (GDP per capita), we standardized the input data using the arithmetic mean and standard deviation.

Table 6. The regression coefficients of the developed nonlinear model of the Basel AML index.

Term	Estimate	Std. Error	*t* Ratio	Prob > |*t*|	Lower 95%	Upper 95%
Intercept (x_0)	4.4855	0.0797	56.2800	0.0001 *	4.3281	4.6429
x_6	0.4000	0.0997	4.0100	0.0001 *	0.2031	0.5970
x_{10}	0.3962	0.1393	2.8400	0.0050 *	0.1210	0.6713
x_2	−1.5416	0.1773	−8.6900	0.0001 *	−1.8918	−1.1914
x_8	−0.6234	0.1677	−3.7200	0.0003 *	−0.9546	−0.2922
x_9	0.4421	0.0932	4.7400	0.0001 *	0.2579	0.6262
x_7	−0.2518	0.0745	−3.3800	0.0009 *	−0.3990	−0.1047
x_5	1.5835	0.2241	7.0600	0.0001 *	1.1409	2.0262
x_4	0.4333	0.0994	4.3600	0.0001 *	0.2371	0.6296
x_1	0.3476	0.0948	3.6700	0.0003 *	0.1603	0.5348
x_3	−0.1583	0.0779	−2.0300	0.0436 *	−0.3122	−0.0045
$x_6 \cdot x_6$	0.2665	0.0579	4.6000	0.0001 *	0.1522	0.3808
$x_6 \cdot x_{10}$	−0.3844	0.1177	−3.2700	0.0013 *	−0.6169	−0.1519
$x_2 \cdot x_2$	−0.1911	0.0761	−2.5100	0.0130 *	−0.3414	−0.0408
$x_6 \cdot x_9$	0.4171	0.0853	4.8900	0.0001 *	0.2486	0.5857
$x_9 \cdot x_9$	0.2167	0.0814	2.6600	0.0085 *	0.0560	0.3774
$x_6 \cdot x_5$	0.3455	0.1425	2.4300	0.0164 *	0.0641	0.6268
$x_{10} \cdot x_2 \cdot x_2$	0.6183	0.2411	2.5600	0.0112 *	0.1422	1.0944
$x_2 \cdot x_8 \cdot x_8$	0.4389	0.0961	4.5700	0.0001 *	0.2492	0.6286
$x_6 \cdot x_8 \cdot x_9$	0.0954	0.0557	1.7100	0.0882	−0.0145	0.2054
$x_9 \cdot x_9 \cdot x_9$	−0.0860	0.0260	−3.3100	0.0011 *	−0.1373	−0.0347
$x_7 \cdot x_7 \cdot x_7$	0.0372	0.0184	2.0300	0.0444 *	0.0009	0.0735
$x_{10} \cdot x_2 \cdot x_5$	−1.1528	0.2973	−3.8800	0.0002 *	−1.7400	−0.5657
$x_6 \cdot x_5 \cdot x_4$	−0.4846	0.0854	−5.6800	0.0001 *	−0.6532	−0.3160

*—significant at level of $\alpha = 0.05$. Source: calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11. x_0—the model's absolute term (Intercept), x_1—WCI, x_2—CPI, x_3—EFI, x_4—GII, x_5—SEDA, x_6—DBI, x_7—GSCI, x_8—HDI, x_9—VAT_{GAP}, x_{10}—GDP per capita, Prob > |t|—achieved level of significance, Estimate—regression coefficient estimation, Std Error—standard error of regression coefficient estimation, *t* Ratio—the calculated value of the Student's test statistic. Source: Compiled and calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

Within the very selection of a specific form of the regression model, the several variants of its form have been analyzed based on general Equation (9). The main criteria for choosing a specific form of the model describing the investigated dependence of the Basel AML index on the selected predictors are the adequacy of the regression model, the maximum achieved value of the coefficient of determination and the adjusted value of the coefficient of determination, the minimum value of RMSE, AICc and BIC. Based on these criteria, the "most suitable" model has been subsequently decided upon. The estimates of regression coefficients, estimates of the standard error of the estimate of regression coefficients, their confidence intervals and their statistical significance are shown in Table 6 and the final form of the regression model is expressed by Equation (21). Although the determination index of model (21) reaches the value of 0.702151, according to the conclusions of Meloun et al. [77], it is acceptable for economic and humanitarian oriented research. At the same time, it must be said that within the analyzed variants of the models, the developed model (21) reached the highest value precisely for this suitability indicator. It is a natural question to think about additional regressors that would increase the accuracy of the model, but this represents the next stage of planned research in the subject area.

Based on the obtained results of the performed analysis, we can proclaim the prediction model (21) as statistically and numerically correct. According to the results shown in Table 6, it is possible to compile a nonlinear regression model (21) in a standardized form expressing the dependence of the studied Basel AML index on the considered independent variables. Specifically, we consider as input factors, global indices, the size of the tax gap (VAT_{GAP}) and the value of gross domestic product per capita (GDP per capita) within destination EU countries in the evaluated period from 2012 to 2019.

$$\begin{aligned} Basel\ AML = \ & 4.4855 + 0.3476 \cdot x_1 - 1.5416 \cdot x_2 - 0.1583 \cdot x_3 + 0.4333 \cdot x_4 + 1.5835 \cdot x_5 + \\ & + 0.4001 \cdot x_6 - 0.2518 \cdot x_7 - 0.6234 \cdot x_8 + 0.4421 \cdot x_9 + 0.33962 \cdot x_{10} + 0.2665 \cdot x_6^2 + \\ & - 0.3844 \cdot x_6 \cdot x_{10} - 0.1911 \cdot x_2^2 + 0.4171 \cdot x_6 \cdot x_9 + 0.2167 \cdot x_9^2 + 0.3455 \cdot x_6 \cdot x_5 + \\ & + 0.6183 \cdot x_{10} \cdot x_2^2 + 0.4388 \cdot x_2 \cdot x_8^2 + 0.0954 \cdot x_6 \cdot x_8 \cdot x_9 - 0.0859 \cdot x_9^3 + 0.0372 \cdot x_7^2 + \\ & - 1.1528 \cdot x_{10} \cdot x_2 \cdot x_5 - 0.4846 \cdot x_6 \cdot x_5 \cdot x_4 \end{aligned} \tag{21}$$

However, the relatively complicated model (21) yields several interesting findings compared to the analyzed model (16) obtained through panel regression analysis. It is primarily the nonlinear effect of some predictors as well as the existence of significant interactions of independent variables.

First of all, it is possible to conclude from model (21) that the model's absolute term is statistically significant at the selected level of significance $\alpha = 0.05$ ($p = 0.0001$) with a total of 38.846% share. Next, we will consider the influence of individual members of the model without considering the influence of the intercept. In this case, the most significant member of model (21) with a 9.808% share is the regressor x_2 (CPI), whose effect in the model, unlike model (16), is nonlinear. However, the quadratic action of the analyzed index occurs in interaction with the predictor x_{10} (GDP per capita) with influence value of 2.889% on the investigated variable. The predictor x_2 influence trend is similar to that in model (16), that is, as the value of the CPI index decreases, the conditional value of the dependent variable (Basel AML index) also decreases.

The first significant regressor with a nonlinear effect on the change in the Basel *AML* value is the DBI index (x_6) with a 5.192% share. This trend has a parabolic character. From the analysis of the predictor x_6 impact it is clear that when the value of the *DBI* index increases from the minimum value of 71.37, the conditional value of the studied *Basel AML* index decreases to the value of 4.335. After exceeding this value and its further increase (DBI), the conditional value of the Basel AML index begins to increase parabolically.

The second significantly nonlinear member of model (21) is the GSCI index (x_7). In the range of values of the GSCI index, from the minimum achieved value to the value of approximately 45.33, there is a nonlinear growth, where the value of the studied Basel AML index reaches its maximum value over the entire studied interval, namely, 4.586. When the value of the GSCI index increases further in the interval from 45.33 to the value of 57.48, there is a nonlinear decrease in the dependent variable (Basel AML), while at this limit value the Basel AML index reaches the value of 4.085. After exceeding this value (GSCI = 57.48), the value of the Basel AML index begins to rise again.

The effect of the size of the VAT_{GAP} tax gap (x_9) is interesting. Its influence in the model represents 5.369% as a separate linear effect, 3.002% as a quadratic effect and 3.736% as a cubic effect. By increasing the value of the amount of tax evasion from the minimum value to the value of 0.9308%, the conditional value of the examined Basel AML index decreases and reaches the value of 4.069 at the border point of the first interval. On the interval of values of the amount of tax evasion expressed as VAT_{GAP} 0.9308% to 3.306%, there is a conditional growth of the Basel AML value to a maximum value of 4.514. After exceeding the value of the amount of tax evasion above the value of 3.306%, there is a sharp decrease in the value of Basel AML to the value of 3.037.

Considering the interactions' points of view, omitted in model (16), the most significant interaction occurs for DBI (x_6), SEDA (x_5) and GII (x_4) with a 6.411% share in the change in the value of the studied variable Basel AML. It is followed by the interaction of the *DBI* index (x_6) and the value of the VAT_{GAP} (x_9) as a measure of tax evasion with 5.430% and by the interaction of the CPI index (x_2) and the square of the HDI index (x_8) with a 5.519% share in model (21). If we return to model (16) as an acceptable result of the panel regression analysis, predictor x_8 (HDI index) does not have a significant impact on the change in the Basel AML value. However, if we look at the nonlinear model (13), the same predictor (x_8), as a separate linear effect, has a significant influence on the change in the value of the Basel AML index (p = 0.0003) with a 4.197% share and it also appears as a quadratic term in a significant interaction with the CPI index (x_2).

Although linear models can provide relatively satisfactory results and answers in some cases, their application to real-world problems and questions leads to considerable distortion. On the other hand, nonlinear models, obtained by regression analysis, are more complicated but allow a better and more detailed understanding of the interrelationships of the investigated processes and systems expressed in the form of regression equations. They make it possible to find even subtle variations and differences, which, however, can be crucial for understanding the investigated phenomena. As an example, we present in Figure 3 the dependence of the Basel AML index value change on the change of two significantly nonlinear regressors of model (21), namely, input variable $\mathrm{VAT_{GAP}}$ (x_9) and the DBI index (x_6). The other considered predictors of model (10) are observed by their mean values, which are expressed by the arithmetic mean.

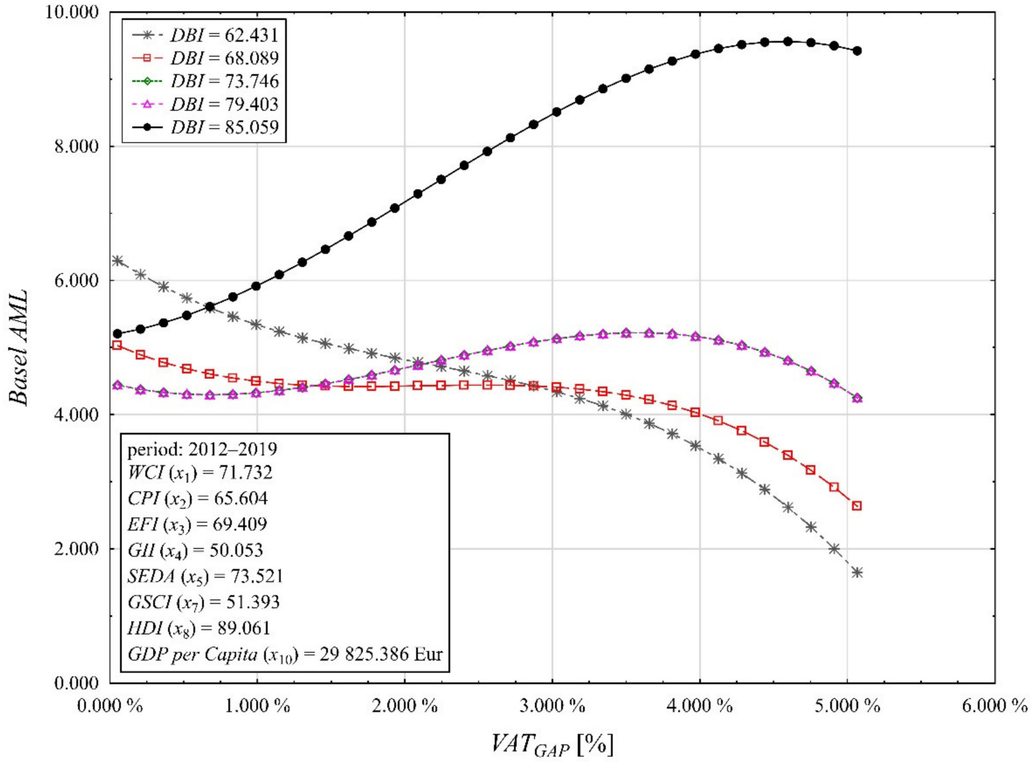

Figure 3. Dependence of the change in the value of the Basel AML index on the change in the value of $\mathrm{VAT_{GAP}}$ and the DBI index based on model (21). Source: Compiled and calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.4.2. Comparison of Developed Models (linear and nonlinear).

Figure 3 shows the significantly nonlinear behavior of the examined dependence. It is interesting to note that with the minimum value (62.431) of the Doing Business Index (DBI), and the minimum value of the $\mathrm{VAT_{GAP}}$ at the level of 0.051%, the studied Basel AML index reaches its maximum value (6.295). At the minimum modelled level of the $\mathrm{VAT_{GAP}}$ size, by increasing Doing Business Index to the value of 68.089, the conditional value of the investigated variable will drop to the level of 5.028. With a further increase in the value of DBI to the values of 73.746 and 79.043, a further decrease in the conditional value of the Basel AML index to the level of 4.442 is observed. However, by increasing DBI to the its maximum modelled value of 85.059, there is an increase in the conditional value of the Basel AML index to the value of 5.025. We observe these principally interesting changes within the entire interval of values of the size of the tax gap ($\mathrm{VAT_{GAP}}$) for individual modelled values of the DBI business performance index.

In the next part of the presented study, the focus will remain on the mutual comparison of the two suitable models, namely, model (16), as the result of panel regression analysis, and model (21), as the result of nonlinear regression analysis. Due to the common European space and common legislation that is adopted in individual EU member states at approximately the same time, we will analyze the "accuracy" of models (16) and (21) within individual years. For clarity, in Figure 4 a graphical display of the relative residuals of the two analyzed models is presented for the selected years 2012, 2015, 2016 and 2019, respectively. The main analysis of the relative residuals for the entire research sample set (all EU countries under analysis between 2012 and 2019) points to the fact that model (16) created by panel regression analysis reaches an average value of −4.430%, while the maximum negative value is −69.684% and the maximum positive value is 28.279%. The range therefore represents a value of 97.964% and an interquartile range of 15.509%. Model (21), created by nonlinear regression analysis, achieves an average value for relative residuals of −1.029% with a maximum negative value at the level of −58.222% and a maximum positive value of 25.196%. The range of relative residuals of model (21) therefore represents 83.419% and the interquartile range is 10.700%.

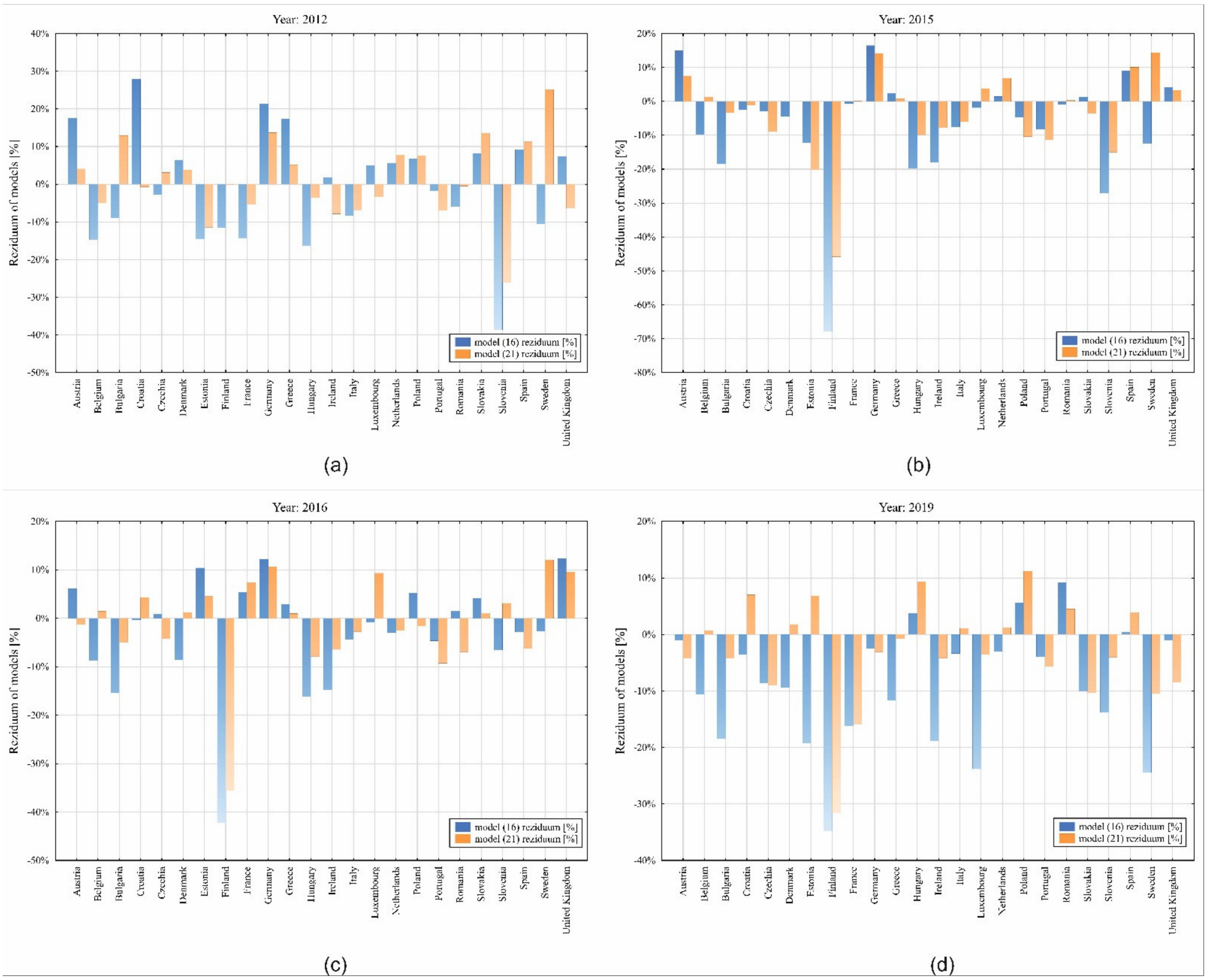

Figure 4. Relative value of residuals of analyzed models (16) and (21) according to monitored years and analyzed EU member countries. (**a**) 2012, (**b**) 2015, (**c**) 2016, (**d**) 2019. Source: Compiled and calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

In the first analyzed year, 2012, the average relative value of the difference between the actual and calculated value of the Basel AML index by model (16) is −0.590% with a maximum negative value at the level of −38.693% and a maximum positive deviation of 27.917%. In contrast, the average relative value of the difference between the actual and calculated values of the Basel AML index by model (21) is 0.999%. This deviation of model (21) in absolute value represents a value higher than model (16) by 0.409%, but the maximum negative value of the deviation calculated by model (21) is −26.057% and the maximum positive deviation is 25.196%. The minimum average deviation is observed in 2014 with model (21) where the value reached is −0.011%. On the contrary, the maximum average values of the difference between the actual and calculated values are reached for model

(16) in 2018 with a value of −11.448% and also in 2018 for model (21) with a value of −6.436%. The summarized statistical properties of the analyzed models according to individual years are presented in Table 7.

Table 7. Descriptive statistics of relative residuals of models (16) and (21) for monitored years.

Year	Model	Valid N	Mean	Median	Minimum	Maximum	Range	Quartile Range	Std. Dev.
2012	(16)	24	−0.590%	0.026%	−38.693%	27.917%	66.610%	18.831%	14.794%
	(21)	24	0.999%	−0.299%	−26.057%	25.196%	51.253%	13.543%	10.522%
2013	(16)	24	3.107%	2.787%	−29.499%	28.279%	57.779%	19.453%	13.877%
	(21)	24	2.501%	1.794%	−15.830%	22.613%	38.443%	8.280%	8.140%
2014	(16)	24	−5.405%	−1.015%	−67.841%	17.268%	85.109%	12.266%	18.041%
	(21)	24	−0.011%	2.361%	−42.987%	17.292%	60.279%	15.925%	12.738%
2015	(16)	24	−7.068%	−3.692%	−67.775%	16.458%	84.233%	13.622%	16.640%
	(21)	24	−3.390%	−0.525%	−45.807%	14.349%	60.155%	12.929%	12.559%
2016	(16)	24	−2.905%	−1.799%	−42.223%	12.364%	54.586%	12.269%	11.587%
	(21)	24	−1.005%	−0.155%	−35.519%	12.094%	47.613%	9.969%	9.608%
2017	(16)	24	−1.989%	0.302%	−46.641%	13.303%	59.944%	14.839%	12.640%
	(21)	24	1.946%	2.752%	−38.415%	16.320%	54.735%	8.669%	10.217%
2018	(16)	24	−11.448%	−8.404%	−69.684%	5.775%	75.460%	15.290%	15.176%
	(21)	24	−6.436%	−4.202%	−58.222%	8.520%	66.742%	11.299%	13.017%
2019	(16)	24	−9.139%	−9.041%	−34.784%	9.219%	44.003%	15.563%	10.610%
	(21)	24	−2.838%	−3.335%	−31.603%	11.164%	42.767%	9.894%	9.093%

Source: Compiled and calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

A certain problem of both models, PRM model (16) and model (21), are countries with a very low value of tax evasion, a high value of gross domestic product per capita (GDP per capita) and extreme values of the other monitored indices. A graphic representation of these four "problematic or critical" countries (Finland, Estonia, Slovenia and Sweden) from the point of view of the deviation of the actual and calculated value of the studied Basel AML index is presented in Figure 5 and a summary of statistical values in Table 8.

Table 8. Descriptive statistics of relative residuals of PRM (16) and nonlinear (21) models for "critical" EU countries.

Country	Model	Valid N	Mean	Median	Minimum	Maximum	Range	Quartile Range	Std. Dev.
Estonia	(16)	8	−8.841%	−11.764%	−25.568%	10.713%	36.281%	17.672%	13.048%
	(21)	8	−4.618%	−7.200%	−20.177%	11.939%	32.116%	18.538%	11.240%
Finland	(16)	8	−43.005%	−44.432%	−69.684%	−3.462%	66.222%	44.600%	25.555%
	(21)	8	−31.372%	−36.967%	−58.222%	1.684%	59.907%	28.543%	21.386%
Slovenia	(16)	8	−21.607%	−21.570%	−38.693%	−6.607%	32.086%	20.789%	12.174%
	(21)	8	−10.517%	−12.141%	−26.057%	3.127%	29.184%	12.359%	9.241%
Sweden	(16)	8	−11.817%	−11.503%	−24.435%	−2.720%	21.715%	9.638%	7.170%
	(21)	8	8.267%	11.979%	−10.508%	25.196%	35.705%	16.073%	11.634%

Source: Compiled and calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

In the third stage of the analytical part of the presented research study, we attempt to resolve the question of at which values of the independent variables (x_1 to x_{10}) the minimum value of the investigated variable (the Basel AML index representing the ML/TF risk in the context of EU member countries) will be reached. In other words, to find out the appropriate combination of input factors (x_1 to x_{10}) affecting the Basel AML index at which the level of risk of legalization of income from criminal activity and financing of terrorism will be minimal.

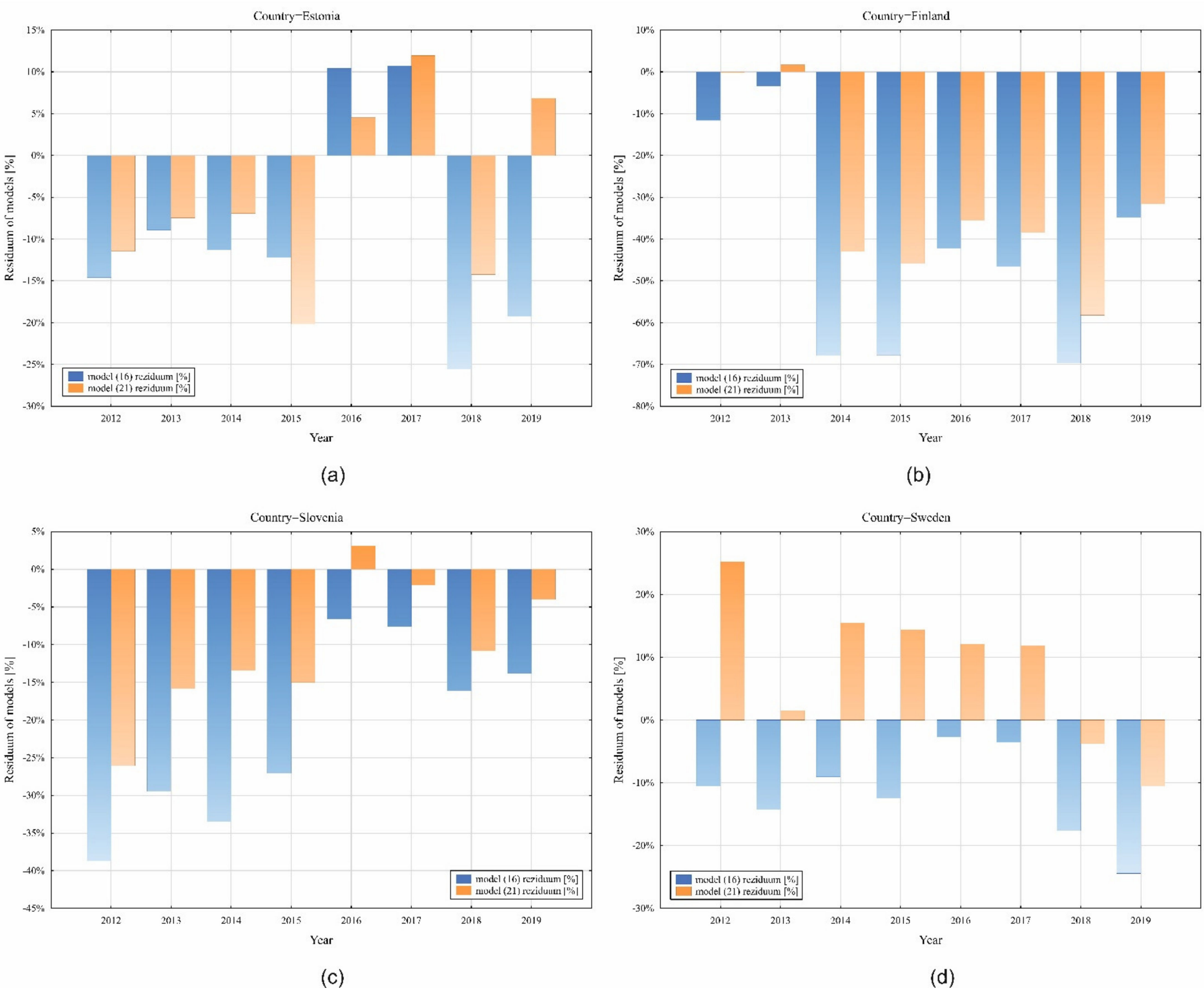

Figure 5. The relative value of the residuals of the analyzed PRM (16) and nonlinear (21) models, for the so-called critical EU countries (critical according to the analysis): (**a**) Estonia, (**b**) Finland, (**c**) Slovenia, (**d**) Sweden). Source: Compiled and calculated by the authors using Statistica 13.5, Matlab2019b and JPM 11.

According to Section 2.6, a mathematical programming problem (MP) in the broader sense expressed by Equation (15) is used for our purpose, hence presence of the constraints and nonlinearities have to be taken into account in our OP. Nonlinear regression model (21) is used as the objective function and optimization constraints are given by the minimum and maximum values of individual predictors (x_1 to x_{10}), as seen in Table 9 (standardized values) and Table 10 (natural scale). The objective function in the general form is as follows:

$$f_0(x_1, x_2, \cdots, x_{10}) \quad \rightarrow \quad \min \tag{22}$$

Table 9. Standardized values of optimization constraints for input variables x_1 to x_{10}.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
min	−2.52721	−2.30282	−1.67997	−2.67651	−1.80871	−2.14519	−2.52556	−2.17388	−2.39013	−1.19814
max	3.048832	1.813837	1.731918	1.897201	1.769436	1.40618	1.94668	2.757387	1.609693	3.76404

Table 10. Optimization constraints for input variables x_1 to x_{10} in natural scale.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
min	2.510	38.974	40.000	53.200	36.000	55.400	62.300	42.400	79.500	5 755.44 €
max	6.780	97.534	92.000	80.900	63.800	85.400	85.300	62.800	95.500	103 464.60 €

To run the optimization procedure for solving the defined OP—the minimization problem with objective function (21) subject to the optimization constraints (Tables 9 and 10), the criterion function f_0 and optimization constraints have been rewritten into a form suitable for optimization in the MATLAB2019b software environment. It should be mentioned that we used the "fmincon ()" solver for constrained nonlinear minimization and the interior point method (IPM) algorithm.

A graphical representation of the optimization process for the modeled value of ML/TF risk symbolized by the AML index is given in Figure 6.

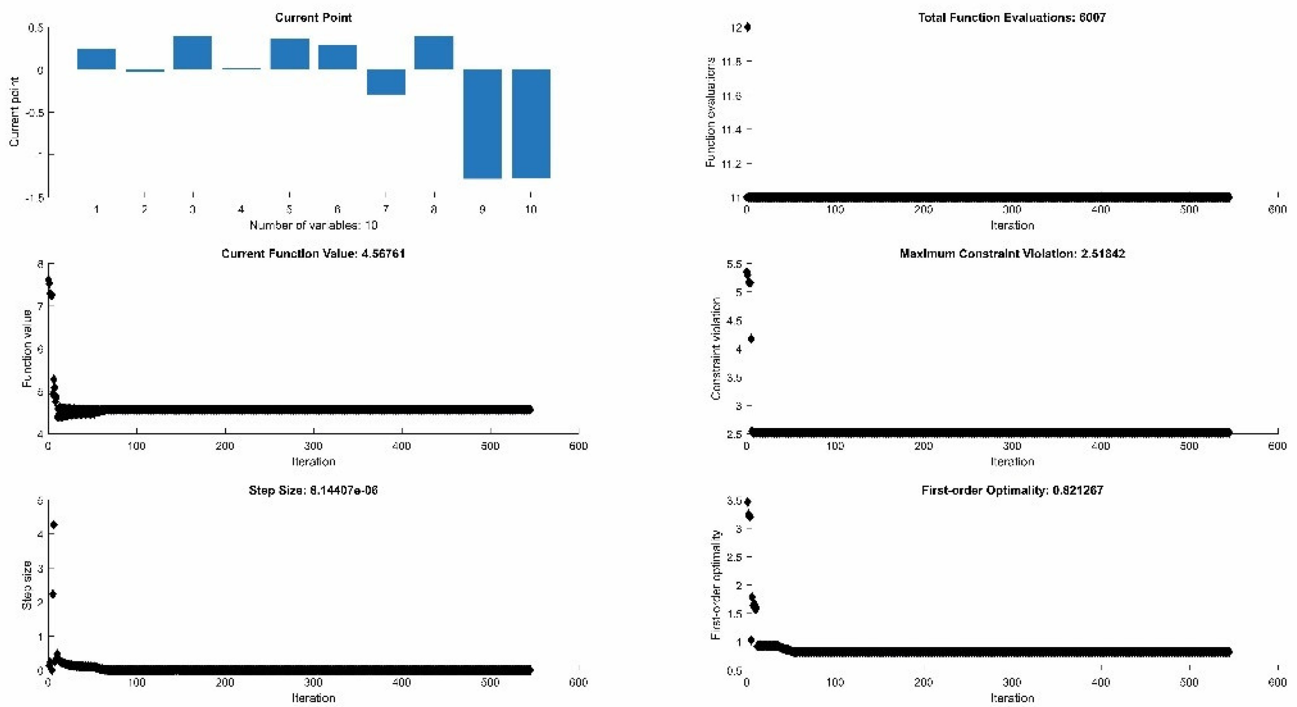

Figure 6. Graphical outputs of the optimization procedure. Source: Compiled by the authors using Matlab2019b.

From the results of data analysis and outputs of the optimization process, it is clear that the optimum (minimum) value of the ML/TF risk symbolized by the Basel AML index is Basel AML opt = 4.568. This value is achieved at the following values of individual predictors: WCI (x_1) = 75.203, CPI (x_2) = 65.208, EFI (x_3) = 71.772, GII (x_4) = 50.208, SEDA (x_5) = 76.647, DBI (x_6) = 76.775, GSCI (x_7) = 50.185, HDI (x_8) = 90.621, VAT_{GAP} (x_9) = 0.0168% and GDP per capita (x_{10}) = EUR 6 973.52.

Based on the abovementioned, it can be concluded that the optimal level of the value of the risk of legalization of income from criminal activity and financing of terrorism (expressed by Basel AML) can be achieved in the following way. Firstly, it is necessary for a certain country to achieve the minimum level of tax evasion expressed in the form of the size of the tax gap (VAT_{GAP}) and simultaneously to hold the value of GDP per capita at the lower limit. This first combination creates a logical conclusion that if there are legislative means that can minimize the amount of tax evasion and in parallel the performance of the country's economy is not great, there is a small space for committing criminal activity in the area of legalizing income from criminal activity.

At the same time, however, we must also consider the influence of the value of the Corruption Perceptions Index (CPI), which at the optimal value of the studied variable Basel AML is just below the average value (65.604) at the level of 65.207. Another conclusion can be drawn from the performed analysis. It can be said that it is not important that the index of perception of corruption be as small as possible, but that it moves at the level of the average of the studied EU member countries. On the other hand, to achieve the optimal value of the Basel AML index, it is necessary that the predictors—World Competitiveness Index (WCI), Economic Freedom Index (EFI), Sustainable Economic Development Index (SEDA), Doing Business Index (DBI) and Human Development Index (HDI)—would be higher than their average value within the studied EU member countries, as it is evident from the results of the monitored period of years.

5. Conclusions

Although multiple linear regression modelling provides useful models for many applications, sometimes it is necessary to applied nonlinear regression techniques to develop models that cannot be transformed into a linear format. The assumption of linearity between related variables when modelling economic phenomena often leads to a distorted explanation for real relationships within applied models. In this study we provide an example of such a situation. When using panel regression analysis, as a representative of linear methods, there is a distortion of the actual effect of the selected predictors on the change in the value of the Basel AML parameter under investigation. On the contrary, the use of nonlinear models allows a better understanding of this action while simultaneously capturing the impact of significant interactions, which allows a better understanding of relationships and the drawing of conclusions and consequences. Moreover, nonlinear models applied in economic practice reflect reality better—as it is declared in this study aiming to identify significant influences in the form of selected global indices, the amount of tax evasion and the value of gross domestic product per capita on the change in the value of the risk of legalization of income from criminal activity. This study declares the following strength and usefulness of these: that a nonlinear approach to modelling uncertainties and risks in economic issues enables an effective search for optimal values of the examined dependent variables, which is very difficult to capture with a linear approach.

It was the previous analyses that led us to the assumed regression model (polynomial multivariate model). Based on this, the predictors were determined: eight global indices, the size of the tax gap and the gross domestic product per capita value. Within the very selection of a specific form of the regression model, several variants of its form were analyzed based on general Equation (9). The main criteria for choosing a specific form of the model describing the investigated dependence of the Basel AML index on the selected predictors are the adequacy of the regression model, the maximum achieved value of the determination index and the adjusted value of the determination index, the minimum value of RMSE, AICc and BIC. Based on these criteria, the "most suitable" model was subsequently selected, whose estimates of regression coefficients, estimates of the standard error of the estimate of regression coefficients, their confidence intervals and their statistical significance are shown in Table 6 and the final form of the regression model is expressed by Equation (21). Although the determination index of model (21) reaches the value of 0.702151, according to the conclusions of Meloun et al. [77], it is acceptable for economic and humanitarian oriented research. At the same time, it must be said that within the analyzed variants of the models, developed model (21) reached the highest value precisely in this suitability indicator. It is a natural question to think about additional regressors that would increase the accuracy of the model, but this represents the next stage of planned research in the subject area.

Our study contributes to the relevant literature by providing empirical evidence on the positive relationship between the Basel AML index and defined global indicators, VATGAP and GDP per capita. However, determining the potential role of certain global socio-econometric indexes on the value of ML/TF risk (expressed by AML index) has never previously been investigated, as is known to us so far. Our study deepens the empirical results in prior literature, e.g., [4,23,25,29–33]. The novelty and main contribution of this study lies in two aspects—the two main goals (i) and (ii) of the presented study were achieved: (i) the new nonlinear mathematical-statistical computational model is developed, on the basis of it (ii) the statement of optimal values of factors affecting the Basel AML index. Thanks to the support of our research by the grant agency, we have achieved interesting results [69–71] (we deal with process-oriented management of financial management focusing on detection of tax evasion in terms of international business). We will continue the investigation in this field of research. Future research will be aimed at closely investigating the relationship of chosen factors (input and output variables) by applying a combination of nonlinear regression analysis and quantile regression analysis.

Author Contributions: Conceptualization, A.V., M.G. and A.K.; methodology, M.G., A.V. and A.K.; validation, M.G., A.V. and A.K.; formal analysis, M.G. and A.V.; investigation, M.G., A.V. and A.K.; resources, A.K., M.G. and A.V.; data curation, M.G., A.V. and A.K.; writing—original draft preparation, A.V. and M.G.; writing—review and editing, M.G., A.K. and A.V.; visualization, M.G. and A.K.; supervision, M.G., A.V. and A.K.; project administration, M.G.; funding acquisition, M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by a grant from the Research Grant Agency within the Ministry of Education, Science, Research and Sport of the Slovak Republic and Slovak Academy of Sciences, VEGA 1/0194/19.

Data Availability Statement: The data for this study were obtained from the official world indicators websites as listed in References [39–50] including links to publicly archived datasets. The data with an annual frequency from 2012 to 2019 are analyzed during the study.

Acknowledgments: The authors would like to thank the grant agency for supporting this research work through the project VEGA 1/0194/19 "Research on process-oriented management of financial management focusing on detection of tax evasion in terms of international business".

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hilal, W.; Gadsden, S.A.; Yawney, J. Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. *Expert Syst. Appl.* **2022**, *193*, 116429. [CrossRef]
2. Cohen, N.; Rubinchik, A.; Shami, L. Towards a cashless economy: Economic and socio-political implications. *Eur. J. Political Econ.* **2020**, *61*, 101820. [CrossRef]
3. Gandal, N.; Hamrick, J.; Moore, T.; Oberman, T. Price manipulation in the bitcoin ecosystem. *J. Monet. Econ.* **2018**, *95*, 86–96. [CrossRef]
4. Korauš, A.; Gombár, M.; Vagaská, A.; Bačík, R.; Korba, P.; Černák, F. Bitcoin Price as one of Basic Cryptocurrencies in Relation to the Basic Stock Market's Indicators. *Entrep. Sustain. Issues* **2021**, *9*, 552–569. [CrossRef] [PubMed]
5. Rogoff, K.S. *The Curse of Cash: How Large-Denomination Bills Aid Crime and Tax Evasion and Constrain Monetary Policy*; Princeton University Press: Princeton, NJ, USA, 2017. [CrossRef]
6. Araujo, L.; Camargo, B. Limited monitoring and the essentiality of money. *J. Math. Econ.* **2015**, *58*, 32–37. [CrossRef]
7. Galeazzi, M.A.; Mendelson, B.; Levitin, M. The anti-money laundering act of 2020. *J. Invest. Compliance* **2021**, *22*, 253–259. [CrossRef]
8. Bartolozzi, D.; Gara, M.; Marchetti, D.J.; Masciandaro, D. Designing the anti-money laundering supervisor: The governance of the financial intelligence units. *Int. Rev. Econ. Financ.* **2022**, *80*, 1093–1109. [CrossRef]
9. Tertychnyi, P.; Godgildieva, M.; Dumas, M.; Ollikainen, M. Time-aware and interpretable predictive monitoring system for Anti-Money Laundering. *Mach. Learn. Appl.* **2022**, *8*, 100306. [CrossRef]
10. Kuzmenko, O.; Šuleř, P.; Lyeonov, S.; Judrupa, I.; Boiko, A. Data mining and bifurcation analysis of the risk of money laundering with the involvement of financial institutions. *J. Int. Stud.* **2020**, *13*, 332–339. [CrossRef]
11. Ahmed, S.; Alshater, M.M.; Ammari, A.E.; Hammami, H. Artificial intelligence and machine learning in finance: A bibliometric review. *Res. Int. Bus. Financ.* **2022**, *61*, 101646. [CrossRef]
12. Demetis, D.S. Fighting money laundering with technology: A case study of Bank X in the UK. *Decis. Support Syst.* **2018**, *105*, 96–107. [CrossRef]
13. Farrugia, S.; Ellul, J.; Azzopardi, G. Detection of illicit accounts over the Ethereum blockchain. *Expert Syst. Appl.* **2020**, *150*, 113318. [CrossRef]
14. Gullo, V.; Montalbano, P. Financial transparency and anomalous portfolio investment flows: A gravity analysis. *J. Int. Money Financ.* **2022**, *128*, 102704. [CrossRef]
15. Isoluari, E.A.; Ameer, I. Money laundering as a transnational business phenomenon: A systematic review and future agenda. *Crit. Perspect. Int. Bus.* **2022**, *ahead of print*. [CrossRef]
16. Lánský, J.; Mihola, J.; Wawrosz, P. Mathematical Modelling of Qualitative System Development. *Mathematics* **2022**, *10*, 2752. [CrossRef]
17. Dobrowolski, Z.; Sułkowski, Ł. Implementing a Sustainable Model for Anti-Money Laundering in the United Nations Development Goals. *Sustainability* **2020**, *12*, 244. [CrossRef]
18. Hoskova-Mayerova, S.; Kalvoda, J.; Bauer, M.; Rackova, P. Development of a Methodology for Assessing Workload within the Air Traffic Control Environment in the Czech Republic. *Sustainability* **2022**, *14*, 7858. [CrossRef]
19. Panda, A.; Duplák, J. Comparison of theory and Practice in Analytical Expression of Cutting Tools Durability for Potential Use at Manufacturing of Bearings. *Appl. Mech. Mater.* **2014**, *616*, 300–307. [CrossRef]
20. Bekesiene, S.; Samoilenko, I.; Nikitin, A.; Meidute-Kavaliauskiene, I. The Complex Systems for Conflict Interaction Modelling to Describe a Non-Trivial Epidemiological Situation. *Mathematics* **2022**, *10*, 537. [CrossRef]

21. Bakare, E.A.; Hoskova-Mayerova, S. Optimal Control Analysis of Cholera Dynamics in the Presence of Asymptotic Transmission. *Axioms* **2021**, *10*, 60. [CrossRef]
22. Panda, A.; Zaloga, V.; Dyadyura, K.; Rybalka, I.; Pandova, I. Modelling Business Process of Manufacturing for Air Compressors. *TEM J.* **2019**, *8*, 430–436. [CrossRef]
23. Mikkelsen, D.; Pravdic, A.; Richardson, B. *Flushing Out the Money Launderers with Better Customer Risk-Rating Models. Risk Practice, 2019*; McKinsey & Company: New York, NY, USA, 2019; pp. 1–7. Available online: https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/flushing-out-the-money-launderers-with-better-customer-risk-rating-models (accessed on 20 September 2022).
24. Hrehová, S. Possibilities of Data Analysis Using Data Model. In Proceedings of the 4th EAI International Conference on Management of Manufacturing System (MMS 2019), Krynica Zdroj, Poland, 8–10 October 2019; EAI/Springer Innovations in Communication and Computing. Springer: Cham, Germany, 2020. [CrossRef]
25. Njegovanović, A. Digital Financial Decision with a View of Neuroplasticity/Neurofinancy/Neural Networks. *Financ. Mark. Inst. Risks* **2018**, *2*, 82–91. [CrossRef]
26. Goodell, J.W.; Kumar, S.; Lim, W.M.; Pattnaik, D. Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *J. Behav. Exp. Financ.* **2021**, *32*, 100577. [CrossRef]
27. Hrehova, S.; Matiskova, D. Possibilities of user interface design with the involvement of machine learning elements using Matlab. In Proceedings of the 23rd International Carpathian Control Conference (ICCC), Sinaia, Romania, 29 May–1 June 2022; pp. 153–157. [CrossRef]
28. Kuzior, A.; Vasylieva, T.; Kuzmenko, O.; Koibichuk, V.; Brożek, P. Global Digital Convergence: Impact of Cybersecurity, Business Transparency, Economic Transformation, and AML Efficiency. *J. Open Innov. Technol. Mark. Complex.* **2022**, *8*, 195. [CrossRef]
29. Mousavi, M.; Zimon, G.; Salehi, M.; Stępnicka, N. The Effect of Corporate Governance Structure on Fraud and Money Laundering. *Risks* **2022**, *10*, 176. [CrossRef]
30. Domashova, J.; Politova, A. The Corruption Perception Index: Analysis of dependence on socioeconomic indicators. *Procedia Comput. Sci.* **2021**, *190*, 193–203. [CrossRef]
31. Haque, M.A.; Raza Shah, S.M.; Arshad, M.U. Sustainable Economic Growth and FDI Inflow: A Comparative Panel Econometric Analysis of Low-Income and Middle-Income Nations. *Sustainability* **2022**, *14*, 14321. [CrossRef]
32. Methodology What's behind the Basel AML Index? Available online: https://index.baselgovernance.org/methodology (accessed on 20 September 2022).
33. Basel AML Index 2021: 10th Public Edition Ranking Money Laundering and Terrorist Financing Risks around the World. Available online: https://baselgovernance.org/sites/default/files/2021-09/Basel_AML_Index_2021_10th%20Edition.pdf (accessed on 19 August 2022).
34. Gnutzmann, H.; McCarthy, K.J.; Unger, B. Dancing with the devil: Country size and the incentive to tolerate money laundering. *Int. Rev. Law Econ.* **2010**, *30*, 244–252. [CrossRef]
35. Batrancea, L.; Rathnaswamy, M.K.; Batrancea, I. A Panel Data Analysis on Determinants of Economic Growth in Seven Non-BCBS Countries. *J. Knowl. Econ.* **2021**, *13*, 1651–1665. [CrossRef]
36. Akartuna, E.A.; Johnson, S.D.; Thornton, A. Preventing the money laundering and terrorist financing risks of emerging technologies: An international policy Delphi study. *Technol. Forecast. Soc. Change* **2022**, *179*, 121632. [CrossRef]
37. Ardizzi, G.; De Franceschis, P.; Giammatteo, M. Cash payment anomalies and money laundering: An econometric analysis of Italian municipalities. *Int. Rev. Law Econ.* **2018**, *56*, 105–121. [CrossRef]
38. Premti, A.; Jafarinejad, M.; Balani, H. The impact of the Fourth Anti-Money Laundering Directive on the valuation of EU banks. *Res. Int. Bus. Financ.* **2021**, *57*, 101397. [CrossRef]
39. Basel AML Index 2022: 11th Public Edition Ranking Money Laundering and Terrorist Financing Risks around the World. Available online: https://index.baselgovernance.org/api/uploads/221004_Basel_AML_Index_2022_72cc668efb.pdf (accessed on 22 August 2022).
40. World Competitiveness Ranking. Available online: https://www.imd.org/centers/world-competitiveness-center/rankings/world-competitiveness/ (accessed on 18 August 2022).
41. Corruption Perceptions Index. Available online: https://www.transparency.org/en/cpi/2021 (accessed on 17 August 2022).
42. 2022 Index of Economic Freedom. Available online: https://www.heritage.org/index/ (accessed on 11 August 2022).
43. Global Innovation Index 2019. Available online: https://www.wipo.int/publications/en/details.jsp?id=4434 (accessed on 19 August 2022).
44. Sustainable Economic Development Assessment. Available online: https://www.bcg.com/industries/public-sector/sustainable-economic-development-assessment (accessed on 17 August 2022).
45. The World Bank. Doing Business. Archive, Reports. Available online: https://archive.doingbusiness.org/en/reports/global-reports/doing-business-reports (accessed on 2 August 2022).
46. The Global Sustainability Competitiveness Index. Available online: https://solability.com/the-global-sustainable-competitiveness-index/downloads (accessed on 11 August 2022).
47. Human Development Index (HDI) Reports. Available online: https://hdr.undp.org/data-center/human-development-index#/indicies/HDI (accessed on 2 August 2022).

48. VAT Gap in the EU. Report 2021. Available online: https://op.europa.eu/en/publication-detail/-/publication/bd27de7e-5323-11ec-91ac-01aa75ed71a1 (accessed on 4 August 2022).
49. Eurostat. GDP and Main Components (Output, Expenditure and Income). Available online: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10_gdp&lang=en (accessed on 15 August 2022).
50. Eurostat. Population on 1 January. Available online: https://ec.europa.eu/eurostat/databrowser/view/tps00001/default/table?lang=en (accessed on 16 August 2022).
51. Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2009; 701p.
52. Afanasiev, V.N.; Kolmanovskii, V.B.; Nosov, V.R. *Mathematical Theory of Control Systems Design*, 1996th ed.; Kluwer Academic: Dordrecht, The Netherlands, 1996; p. 696.
53. Camacho, E.F.; Alba, C.B. *Model Predictive Control*, 2nd ed.; Springer Science & Business Media: New York, NY, USA, 2013; p. 427.
54. Świercz, R.; Oniszczuk-Świercz, D.; Chmielewski, T. Multi-Response Optimization of Electrical Discharge Machining Using the Desirability Function. *Micromachines* **2019**, *10*, 72. [CrossRef]
55. Vinter, R. *Optimal Control*; Modern Birkhäuser Classics; Birkhäuser Boston, Inc.: Boston, MA, USA, 2010; p. 507. [CrossRef]
56. Rao, S. *Engineering Optimization. Theory and Practice*, 4th ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009; p. 830.
57. Vagaská, A.; Gombár, M.; Straka, L'. Selected Mathematical Optimization Methods for Solving Problems of Engineering Practice. *Energies* **2022**, *15*, 2205. [CrossRef]
58. Ahmad, F.; Draz, M.U.; Yang, S.-C. Causality nexus of exports, FDI and economic growth of the ASEAN5 economies: Evidence from panel data analysis. *J. Int. Trade Econ. Dev.* **2018**, *27*, 685–700. [CrossRef]
59. Greene, W.H. *Econometric Analysis*, 7th ed.; Pearson Education Limited: Edinburgh, Scotland, 2012; pp. 383–494.
60. Baltagi, B.; Song, S.; Koh, V. Testing Panel Data Regression Models with Spatial Error Correlation. *J. Econom.* **2003**, *117*, 123–150. [CrossRef]
61. Donald, S.; Lang, K. Inference with Difference-in-Differences and Other Panel Data. *Rev. Econ. Stat.* **2007**, *89*, 221–233. [CrossRef]
62. Hausman, J.A. Specification Tests in Econometrics. *Econometrica* **1978**, *46*, 1251–1271. [CrossRef]
63. Im, K.S.; Pesaran, M.; Shin, Y. Testing for unit roots in heterogeneous panels. *J. Econ.* **2003**, *115*, 53–74. [CrossRef]
64. Bai, J.; Choi, S.H.; Liao, Y. Feasible generalized least squares for panel data with cross-sectional and serial correlations. *Empir. Econ.* **2021**, *60*, 309–326. [CrossRef]
65. Chatterjee, S.; Simonoff, J.S. *Handbook of Regression Analysis*; Wiley: New York, NY, USA, 2013.
66. Hsiao, C.-W.; Chan, Y.-C.; Lee, M.-Y.; Lu, H.-P. Heteroscedasticity and Precise Estimation Model Approach for Complex Financial Time-Series Data: An Example of Taiwan Stock Index Futures before and during COVID-19. *Mathematics* **2021**, *9*, 2719. [CrossRef]
67. Belkin, M.; Hsu, D.; Mitra, P.P. Overfitting or Perfect Fitting? *Risk Bounds for Classification and Regression Rules that Interpolate.* 2018. Available online: https://arxiv.org/abs/1806.05161 (accessed on 26 August 2022).
68. Vilček, I.; Řehoř, J.; Carou, D.; Zeman, P. Residual stresses evaluation in precision milling of hardened steel based on the deflection-electrochemical etching technique. *Robot. Comput. Integr. Manuf.* **2017**, *47*, 112–116. [CrossRef]
69. Browne, M.W.; Cudeck, R. Alternative Ways of Assessing Model Fit. In *Testing Structural Equation Models*; Bollen, K.A., Long, J.S., Eds.; Sage: Beverly Hills, CA, USA, 1993; pp. 136–162.
70. Cameron, A.C.; Windmeijer, F.A.G. An R-squared measure of goodness of fit for some common nonlinear regression models. *J. Econom.* **1997**, *77*, 329–342. [CrossRef]
71. Hebák, P.; Hustopecký, J.; Malá, I. *Vícerozměrné Statistické Metody*, 2nd ed.; Informatorium: Praha, Czech Republic, 2006; 240p.
72. Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E.; Tatham, R.L. *Multivariate Data Analysis*, 6th ed.; Prentice Hall: Hoboken, NJ, USA, 2005; 928p.
73. Akaike, H. Information theory and an extension of the maximum likelihood principle. In *Information Theory: Proceedings of the 2nd International Symposium*; Akademiai Kiado: Budapest, Hungary, 1974; pp. 267–281.
74. Gombár, M.; Korauš, A.; Vagaská, A.; Tóth, Š. Analytical View on the Sustainable Development of Tax and Customs Administration in the Context of Selected Groups of the Population of the Slovak Republic. *Sustainability* **2022**, *14*, 1891. [CrossRef]
75. Cicceri, G.; Inserra, G.; Limosani, M. A Machine Learning Approach to Forecast Economic Recessions—An Italian Case Study. *Mathematics* **2020**, *8*, 241. [CrossRef]
76. Korauš, A.; Gombár, M.; Vagaská, A.; Šišulák, S.; Černák, F. Secondary Energy Sources and Their Optimization in the Context of the Tax Gap on Petrol and Diesel. *Energies* **2021**, *14*, 4121. [CrossRef]
77. Meloun, M.; Militký, J. *Statistická Analýza Experimentálních Dat/Statistical Analysis of Experimental Data*; Academia: Prague, Czech Republic, 2004; p. 928.