Comprehensive Forecasting of Interconnected Socio-Economic Indicators as a Methodological Basis for Adopting Optimal Management

Oleksandr Yankovyi
Department of Enterprise Economics and Business Organization
Odesa National Economic University
Odesa, Ukraine
https://orcid.org/0000-0003-2413-855X

Nataliia Shmatko
Department of Management, Innovative Entrepreneurship and International Economic Relations
National Technical University "Kharkiv Polytechnic Institute"
Kharkiv, Ukraine
http://orcid.org/0000-0002-4909-252X

Oleksii Hutsaliuk
Department of Economics, Management and Business Administration
International European University
Kyiv, Ukraine
https://orcid.org/0000-0002-6541-4912

Olena Kabanova
Department of Management
Classic Private University
Zaporizhzhia, Ukraine
https://orcid.org/0000-0002-2678-8401

Viktoriia Tomareva-Patlakhova
Department of Management
Classic private university
Zaporizhzhia, Ukraine
https://orcid.org/0000-0001-8910-2641

Yuliia Rud
Department of Management
Classic Private University
Zaporizhzhia, Ukraine
https://orcid.org/0000-0002-0328-5895

Abstract**—The article presents the theoretical achievements of technical cybernetics, in particular the principle of balancing variables. The essence of which in the final conclusion on the suitability of certain forecasts is determined by the degree of performance for the predicted values of interrelated variables of the balance ratio, fair to them in the prehistoric period. The principle of balance of variables and the criterion based on it plays the role of an external complement that carries new information about the studied process and can be used as a priority in comparison with the methods of choosing mathematical forms of trends of varying complexity. First of all, this applies to the criterion of maximizing the coefficient of determination, which is usually used in the selection of the best reference functions that approximate the isolated time series. It is proposed to use this principle and the criterion of balance of variables based on it when forecasting socio-economic indicators, between which there are additive and multiplicative links. The application of the criterion of balance of variables in determining the forecast of the trend of forecasting the number of births, deaths and natural population growth in Ukraine for 2020-2021 is illustrated. Despite the small sample, balanced results were obtained for short-term forecasting of Ukraine's demographic indicators. They can serve as a reliable methodological basis for planning measures to improve the demographic situation in the country, as well as for making optimal management decisions on the use of human resources in the economy.**

Keywords—variable balance criterion, predictor, trend model, prehistory period, bias period

I. Introduction

Forecasting is a methodological basis for planning - one of the most important management functions at the micro and macro levels of the economic system. The quality of management decisions depends on the accuracy and reliability of forecast estimates of socio-economic indicators. Therefore, the urgency of the problems of developing an effective forecasting methodology is not in doubt among both scientists and practitioners.

In modern science, a number of methods have been developed that have certain advantages and disadvantages and often provide quite satisfactory results in predicting socio-economic variables, considered in isolation [1; 2]. However, the situation fundamentally changes when the simultaneous forecasting of several interdependent indicators such as the number of the total population, men and women; birth rate, mortality and natural population growth; export, import and export-import balances, etc. Such variables are described by additive models $Y_1 = Y_2 + Y_3 +\ldots + Y_k$, in the left part of which there are some socio-economic features, and in the right - their components. The multiplicative dependences $Y_1 = Y_2 \times Y_3 \times\ldots \times Y_k$ could be easily reduced to the additive type by taking the logarithm of the left and right parts of the model.

II. Methods

The principle of balance of variables and the criterion based on it plays the role of an external complement that carries new information about the studied process and can be used as a priority in comparison with the methods of choosing mathematical forms of trends of varying complexity. The principle of balance of variables was first put forward in the theory of self-organization as one of the areas of technical cybernetics in the modeling of complex probability systems using the method of group consideration of arguments [3].

However, its application was limited to forecasting the development of technical and natural systems, not finding practical use in socio-economic research. Exceptions are separate publications of the authors of this article, which highlight the applied aspects of the principle of balance of variables and the criterion of balance of variables based on it [4; 5; 6]. Certain special aspects of the process of choosing management decisions and evaluation of their effectiveness are reasonably reflected in the scientific works [7; 8; 9], including forecasting the producer's cash flows [10; 11].

III. Theoretical Basis of Forecasting Using the Balance of Variables Principle

In its content, the principle of balance of variables is the main in the whole set of principles of mathematical and statistical modeling and forecasting, as it is based on absolutely reliable information about the future state of the socio-economic system: no matter what the individual values of variables, that the same relationships between the traits will follow follows from the essence of the studied traits [12; 13].

The rationality of such an approach is not in doubt, because we know for sure about the future only what is in it and what took place in the past and in the present. The principle of balance of variables and the criterion based on it plays the role of an external complement that carries new information about the studied socio-economic process [14; 15; 16]. This means that it should be used as a priority in comparison with already developed mathematical and statistical theory and tested in practice methods for selecting mathematical forms of models (trend equations of varying complexity) in the construction of predictive predictors. First of all, this applies to the criterion of maximizing the coefficient of determination (maxR^2), which is usually used in the selection of the best reference functions that approximate isolated time series.

We will discuss the theoretical aspects of determining and using the criterion of balance of variables on the example of a simpler relationship between the studied socio-economic indicators [17]:

$$Y_1 = Y_2 + Y_3. \quad (1)$$

The role of Y_1 can be, for example, the birth rate, Y_2 - mortality, Y_3 - the natural population growth of the country (region, city).

Let the problem be to obtain the best comprehensive forecast of variables Y_1, Y_2, Y_3 for some period of bias L based on known values of these indicators for a certain period of prehistory N. The process of forecasting using the criterion of balance of variables in this case consists of the following main stages.

1. Definition of a homogeneous set of observations, which presents information about the dynamics of the studied socio-economic characteristics. A more important condition for ensuring a homogeneous set of observations is the comparability of statistics on interrelated indicators over time in the territory of their origin. That is, the information for each period of time (year, month, week, etc.) must relate to the same territory – country, region and city. Fulfillment of this requirement is usually ensured on the basis of a qualitative typological grouping [18].

2. Selection of possible reference functions that most accurately describe the studied time series. The time series of each indicator, taken from the prehistory period, can be represented by trend models [19]:

$$Y_i = f_i(X) + \varepsilon_i, \quad (2)$$

where X is the time factor (X = 1, 2,…, N); ε_i is a component that reflects the action of random factors.

To determine the type of basic mathematical functions $f_i(X)$ at this stage, traditional methods of statistical modeling are widely used: visual study of graphs, calculation of successive differences, growth characteristics, etc. parameters. As a result, in the second stage, from the whole list of reference functions, the number of which can be quite large, several main ones are selected (n_i, i = 1, 2, 3), which most adequately describe the interconnected processes being studied. Next, the least squares method determines the parameters of each of the selected support functions $f_i(X)$, including the coefficient of determination R^2, as well as the criteria of mathematical statistics (Fisher's F-test, Student's t-test, etc.). Extrapolation is performed and point forecast values of indicators for a given bias period $L = l_1 - l_2$ are calculated by substituting the corresponding time value $X = N + L$ in $f_i(X)$.

3. Sequential search of a limited number of reference functions selected in the previous stage. The degree of observance of the initial balance ratio for all points of the bias period $L = l_1 - l_2$ acts as an estimation of each combination. The number of all possible combinations m is determined by the product

$$m = n_1 \times n_2 \times n_3, \quad (3)$$

where n_1, n_2, n_3 - the number of reference functions selected for approximation the relevant interrelated indicator.

For example, if for three variables in the second stage 5 reference functions are selected, then $m = 5 \times 5 \times 5 = 125$.

The best combination of reference functions corresponds to the minimum value of the criterion of balance of variables B_j, which is calculated by the formula:

$$B_j = \min\left[\frac{\sum_{l_1}^{l_2}(Y_1 - Y_2 - Y_3)^2}{\sum_{l_1}^{l_2} {Y_1}^2}\right]. \quad (4)$$

That is, the degree of compliance with the initial balance ratio (1) for all points of the bias period acts as an assessment of each combination of forecast trends. The value of B_j characterizes the total relative imbalance of the complex forecast of the studied indicators. The closer it is to zero, the more correctly the chosen predictor and vice versa. Therefore, criterion (4) allows to optimize the definition of the predictor on the set of selected reference functions.

4. Selection of the optimal length of the prehistory period N. We recommend to carry out this stage only on large samples ($N > 20$) in the presence of outdated observations at the beginning of the time series. Indeed, it is very important to correctly determine the length of the prehistory period N, to capture current trends in the development of the studied processes, so that the predicted values do not contradict the current trends of interrelated socio-economic indicators. An iterative approach is used for this purpose. Its essence is to form new truncated from the beginning of the time series Y_1', Y_2', Y_3' by excluding from the original series G (G = 1, 2, ..., N / 2) the first levels. Trend models are constructed for each truncated series, forecast values are determined, and B_j values are found. In other words, the behavior of the criterion of balance of variables with a gradual reduction in the length of the prehistory period N is investigated.

5. Determining the interval values of the forecast. This stage completes the process of comprehensive forecasting of interrelated socio-economic indicators and is a normal calculation of confidence intervals according to a known scheme.

$$Y_{i\,N+L} \pm \Delta_i \quad (5)$$

where $Y_{i\,N+L}$ is the best point forecast for the i-th indicator; Δ_i is the marginal error of the i-th point forecast.

IV. PRACTICAL APPLICATION OF THE CRITERION OF THE BALANCE OF VARIABLES IN THE COMPREHENSIVE FORECASTING OF DEMOGRAPHIC INDICATORS OF UKRAINE

Consider the practical use of this theory for a comprehensive forecast of indicators of the number of live births Y_1, the number of deaths Y_2 and natural population growth in Ukraine Y_3 for 2020-2021 (Table I), between which there is a relationship (1).

It should be noted that since 1991, Ukraine has seen a negative natural increase in population. At the same time, the birth rate tends to decrease with a fairly stable mortality rate.

Consider briefly the main stages of comprehensive forecasting of live births, deaths and natural population growth in Ukraine for 2020-2021.

TABLE I. BIRTH, DEATH AND NATURAL INCREASE RATES POPULATION OF UKRAINE FOR 2010-2019

Years	**Number of live births, thousand people (Y_1)**	**The number of dead, thousand people (Y_2)**	**Natural increase, thousand people (Y_3)**
2010	497.7	698.2	-200.5
2011	502.6	664.6	-162.0
2012	520.7	663.1	-142.4
2013	503.7	662.4	-158.7
2014	465.9	632.3	-166.4
2015	411.8	594.8	-183.0
2016	397.0	583.6	-186.6
2017	364.0	574.1	-210.1
2018	335.9	587.7	-251.8
2019	308.8	581.1	-272.3

Source: compiled by the authors based on [20].

1. As the last six years in Ukraine there is an armed conflict provoked by the Russian Federation, since 2014 the official Ukrainian statistics do not take into account information from the temporarily occupied territory of the Autonomous Republic of Crimea, Sevastopol and part of the temporarily occupied territories in Donetsk and Luhansk regions. Therefore, it was decided to limit the study to data for 2014-2019 (N = 6), which are quite comparable and homogeneous from a territorial point of view (see the last six lines of Table 1).

2. To determine the type of reference mathematical functions $f_i(X)$ at this stage we used qualitative theoretical analysis and visual study of graphs of variables Y_1, Y_2, Y_3. Thus, for live-born Y_1 it was proposed to use only one reference function (n_1 = 1): linear, which is presented in Fig. 1. Its accuracy (R^2 = 0.9805) indicates a steady decline in birth rates in recent years with an average annual absolute decline of 29.89 thousand people.

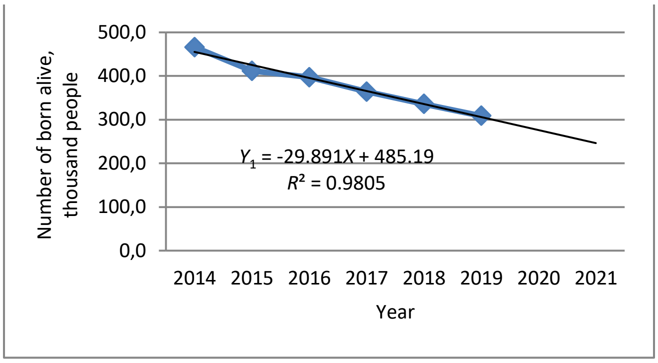

Fig. 1. Actual, aligned with a linear trend and predicted for 2020-2021 the number of born alive in Ukraine, thousand people

Source: built by the authors.

Point forecast of the number of live births according to a linear trend:

for 2020: $Y_{1\ 2020}$ = -29.891 × 7 + 485.19 = 275.95 thousand people;

for 2021: $Y_{1\ 2021}$ = -29.891 × 8 + 485.19 = 246.06 thousand people.

For the number of dead Y_2 on the basis of graphical analysis, it was proposed to use three support functions (n_2 = 3): parabolic, logarithmic and power, which are presented in Fig. 2-4. Moreover, the highest accuracy of approximation (R^2 = 0.8988) provides a parabola of the 2nd degree.

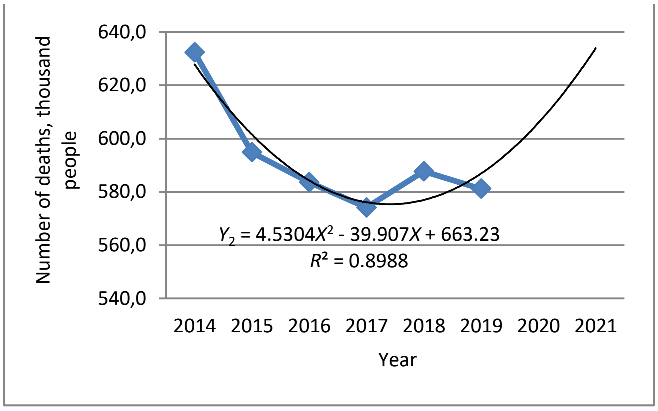

Fig. 2. Actual, aligned by parabolic trend and predicted for 2020-2021 the number of deaths in Ukraine, thousand people

Source: built by the authors.

Point forecast of the number of deaths according to the parabolic trend:

for 2020: $Y_{2\ 2020} = 4.5304 \times 7^2 - 39.907 \times 7 + 663.23$ = 605.87 thousand people;

for 2021: $Y_{2\ 2021} = 4.5304 \times 8^2 - 39.907 \times 8 + 663.23$ = 633.92 thousand people.

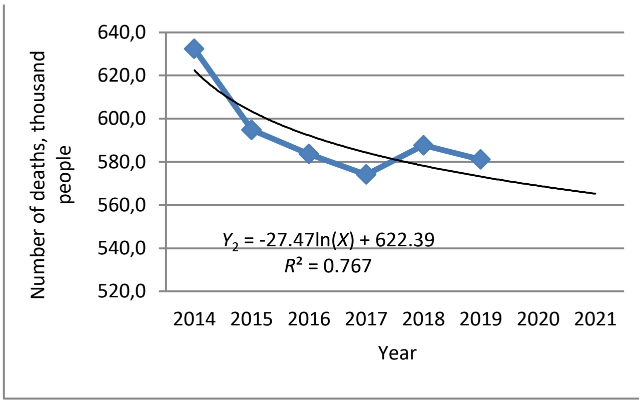

Fig. 3. Actual, aligned on a logarithmic trend and predicted for 2020-2021 the number of deaths in Ukraine, thousand people

Source: built by the authors.

Point forecast of the number of deaths according to the logarithmic trend:

for 2020: $Y_{2\ 2020}$ = -27.47 × ln(7) + 622.39 = 568.94 thousand people;

for 2021: $Y_{2\ 2021}$ = -27.47 × ln(8) + 622.39 = 565.27 thousand people.

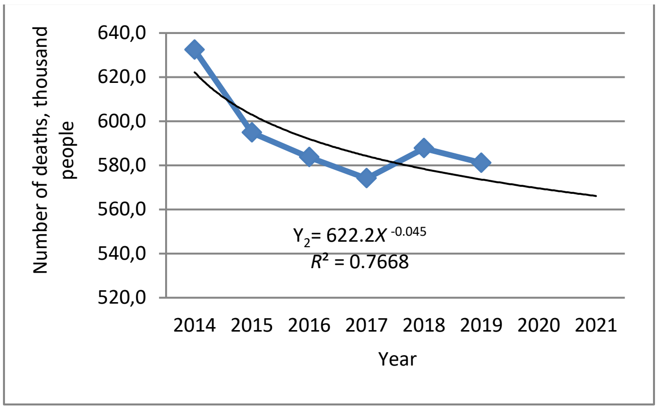

Fig. 4. Actual, aligned with the power trend and predicted for 2020-2021 the number of deaths in Ukraine, thousand people

Source: built by the authors.

Point forecast of the number of deaths according to the power trend:

for 2020: $Y_{2\ 2020}$ = 622.2 × $7^{-0.045}$ = 570.03 thousand people;

for 2021: $Y_{2\ 2021}$ = 622.2 × $8^{-0.045}$ = 566.62 thousand people.

For natural population growth Y_3 on the basis of graphical analysis it was proposed to use two reference functions (n_3 = 2): linear and parabolic, which are presented in Fig. 5, 6. Moreover, the highest accuracy of approximation (R^2 = 0.9774) is provided by a parabola of the 2nd degree.

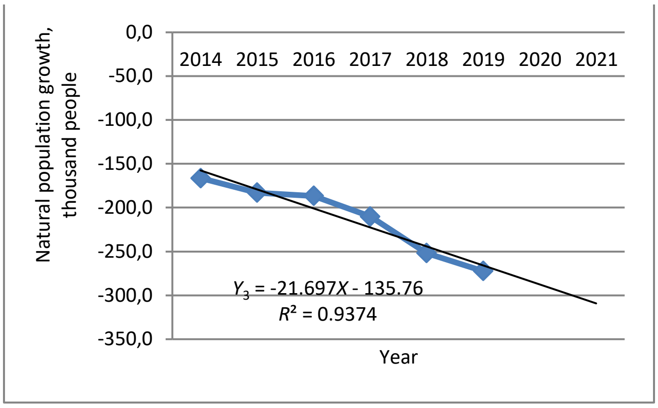

Fig. 5. Actual, aligned with the linear trend and predicted for 2020-2021 natural population growth in Ukraine, thousand people

Source: built by the authors.

Point forecast of natural population growth according to a linear trend:

for 2020: $Y_{3\ 2020}$ = -21.697 × 7 – 135.76 = -287.64 thousand people;

for 2021: $Y_{3\ 2021}$ = -21.697 × 8 – 135.76 = -309.34 thousand people.

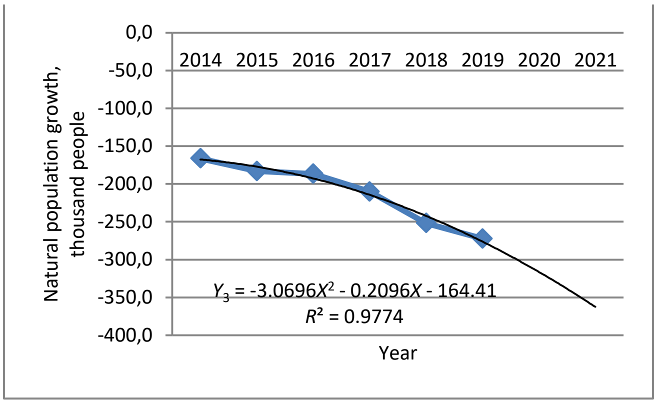

Fig. 6. Actual, aligned by the parabolic trend and predicted for 2020-2021 natural population growth in Ukraine, thousand people

Source: built by the authors.

Point forecast of natural population growth according to the parabolic trend:

for 2020: $Y_{3\ 2020}$ = -3.0696 × 7^2 – 0.2096 × 7 – 164.41 = -316.29 thousand people;

for 2021: $Y_{3\ 2021}$ = -3.0696 × 8^2 – 0.2096 × 8 – 164.41 = -362.54 thousand people.

3. By formula (3) we find the total number of combinations of reference functions selected at the previous stage (m = 1 × 3 × 2 = 6) and calculate the imbalance by formula (4) for each potential predictor.

Thus, six predictors claim the role of the best combination of support functions for complex forecasting of the number of live births Y_1, the number of deaths Y_2 and the natural increase of the population of Ukraine Y_3 for 2020-2021. For each of them in Table II shows the results of the calculation of the criterion of balance of variables by formula (4).

TABLE II. DETERMINING THE OPTIMAL PREDICTOR FOR COMPREHENSIVE FORECASTING FOR 2020-2021 OF THE NUMBER OF LIVE BIRTHS, THE NUMBER OF DEATHS AND NATURAL POPULATION GROWTH IN UKRAINE

Combination	Predictor	Unbalance B_j
1-1-1	Y_1 = -29.891X + 485.19; Y_2 = 4.5304X^2 – 39.907X + 663.23; Y_3 = -21.697X – 135.76.	0.058181
1-1-2	Y_1 = -29.891X + 485.19; Y_2 = 4.5304X^2 – 39.907X + 663.23; Y_3 = -3.0696X^2 – 0.2096X – 164.41.	0.006048
1-2-1	**Y_1 = -29.891X + 485.19; Y_2 = -27.47ln(X) + 622.39; Y_3 = -21.697X – 135.76.**	**0.000922**
1-2-2	Y_1 = -29.891X + 485.19; Y_2 = -27.47ln(X) + 622.39; Y_3 = -3.0696X^2 – 0.2096X – 164.41.	0.017711
1-3-1	Y_1 = -29.891X + 485.19; Y_2 = 622.2$X^{-0.045}$; Y_3 = -21.697X – 135.76.	0.001225
1-3-2	Y_1 = -29.891X + 485.19; Y_2 = 622.2$X^{-0.045}$; Y_3 = -3.0696X^2 – 0.2096X – 164.41.	0.016502

Source: calculated by the authors.

For example, the imbalance of the optimal combination 1-2-1 (in Table II highlighted in bold), which meets the criterion min B_j, according to formula (4) is calculated as follows:

$$B_{1-2-1} = \frac{(275.95 - 568.94 + 297.64)^2 + (246.06 - 565.27 + 309.94)^2}{275.95^2 + 246.06^2} = 0.000922.$$

The closest in accuracy (Bj = 0.001225) to the optimal is the combination 1-3-1, in which the number of deaths is approximated not by a logarithmic curve, but by a power function close to it in accuracy.

4. In this study, we did not optimize the length of the prehistory period N, because this stage is recommended only for large samples (N > 20), when there are outdated observations at the beginning of the time series. In the case under discussion, the sample was small (N = 6) and it revealed a fairly deep min Bj = 0.000922, which indicates the successful solution of the problem.

5. Since the principle of balance of variables and based on it criterion (4) is used only to solve the problem of choosing a forecast predictor, the confidence intervals of the forecast are calculated according to the known scheme (5), for example, in the system STATISTICA [3]. In this study, they are presented together with the corresponding point forecasts in Table III.

V. CONCLUSIONS

Thus, if the identified patterns of demographic processes in Ukraine remain in the near future, the point and 95 percent interval forecasts of the number of live births, deaths and natural population growth in Ukraine for 2020-2021 will be as presented in Table. III.

Thus, the best predictor, which provides the most balanced forecast estimates of the number of live births, the number of deaths and the magnitude of natural population growth in Ukraine for 2020-2021, is the following:

$$\begin{cases} Y_1 = -29.891X + 485.19 \\ Y_2 = -27.47\ln(X) + 622.39 \\ Y_3 = -21.697X - 135.76 \end{cases} \quad (6)$$

Note that the obtained predictor (6) did not include the most accurate trend models of mortality and natural increase (parabolas of the 2nd degree with coefficients of determination of 0.899 and 0.977, respectively). This means that in the case of their isolated application, the researcher is in danger of receiving a "fan" of contradictory forecasts that violate the balance sheet (1).

The simplicity and accessibility of the proposed approach to comprehensive forecasting of interrelated indicators based on the principle (criterion) of balance of variables in combination with its successful results, even in small samples will help increase the use of effective ideas and methods of technical cybernetics in socio-economic research.

TABLE III. THE RESULTS OF COMPREHENSIVE FORECASTING FOR 2020-2021 DEMOGRAPHIC INDICATORS OF UKRAINE, THOUSAND PEOPLE

YEARS	Number of born alive			Number of deaths			Natural population growth		
	point forecast	*the lower limit of 95% of the confidence interval*	*the upper limit of 95% of the confidence interval*	*point forecast*	*the lower limit of 95% of the confidence interval*	*the upper limit of 95% of the confidence interval*	*point forecast*	*the lower limit of 95% of the confidence interval*	*the upper limit of 95% of the confidence interval*
2020	275.95	253.15	298.74	568.94	547.02	590.85	-287.64	-317.96	-257.32
2021	246.06	217.88	274.23	565.27	541.01	589.52	-309.34	-346.81	-271.87

Source: calculated by the authors.

REFERENCES

[1] A.Karatzias, K. G. Power, J. Flemming, F. Lennan, V. Swanson, “The role of demographics, personality variables and school stress on predicting school satisfaction/dissatisfaction: Review of the literature and research findings”, Educational Psychology, vol. 22(1), pp. 33-50, 2002.

[2] M. G. Pittau, R. Zelli, A. Gelman, “Economic disparities and life satisfaction in European regions”, Social indicators research, vol. 96(2), pp. 339-361, (2010).

[3] V. Borovikov, G. Ivchenko, Forecasting in the STATISTICA System in the Windows Environment, M.: Finansy i statistika, 1999.

[4] K. Kostetska, N. Khumarova, Y. Umanska, N.Shmygol, V. Koval, “Institutional qualities of inclusive environmental management in sustainable economic development”, Management Systems in Production Engineering, vol. 28 (2), pp. 15-22, 2020.

[5] M. Yeshchenko, V.Koval, O. Tsvirko, “Economic policy priorities of the income regulation”, Espacios, vol. 40 (38), pp. 11, 2019.

[6] O. Baklanova, M. Petrova, V. Koval, “Institutional Transmission in Economic Development”, Ikonomicheski Izsledvania, 29(1), 68-91, 2020.

[7] M. Mezzetti, F. C. Billari, “Bayesian correlated factor analysis of socio-demographic indicators”, Statistical Methods and Applications, vol. 14(2), pp. 223-241, 2005.

[8] M. Petrova, V. Koval, M. Tepavicharova, A.Zerkal, A. Radchenko, N.Bondarchuk, “The interaction between the human resources motivation and the commitment to the organization”, Journal of Security and Sustainability Issues, vol. 9(3), pp. 897-907, 2020.

[9] O. Hutsaliuk, O. Yaroshevska, N. Shmatko, I. Kulko-Labyntseva, A. Navolokina “Stakeholder approach to selecting enterprise-bank interaction strategies”, Problems and Perspectives in Management, vol. 18, pp. 42–55, 2020.

[10] G. M. Righetto, R. Morabito, D. Alem, “A robust optimization approach for cash flow management in stationery companies”, Computers & Industrial Engineering, vol. 99, pp. 137-152, 2016.

[11] S. Kvasha, L. Pankratova, V.Koval, R.Tamošiūnienė, “Illicit financial flows in export operations with agricultural products”, Intelellectual Economics, vol. 13(2), pp. 195-209, 2019.

[12] I. Gontareva, V. Babenko, N. Shmatko, O. Litvinov, H. Obruch “The Model of Network Consulting Communication at the Early Stages of Entrepreneurship”. WSEAS Transactions on Environment and Development, vol. 16, pp. 390–396, 2020.

[13] O. Yankovyi, Yu. Goncharov, V. Koval, T. Lositska, “Optimization of the capital-labor ratio on the basis of production functions in the economic model of production”, Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, vol. 4, pp. 134-140, 2019.

[14] E. A.Padawer, J. M. Jacobs-Lawson, D. A. Hershey, D. G. Thomas, “Demographic indicators as predictors of future time perspective”, Current Psychology, vol. 26(2), pp. 102-108, 2007.

[15] N. Ahmad, A proposed framework for business demography statistics, Measuring entrepreneurship, Springer, Boston, MA, 2008, pp. 113-174.

[16] H. d’Albis, “Demographic structure and capital accumulation” Journal of Economic Theory, vol. 132(1), pp. 411-434, 2007.

[17] G.Khatskevich, A.Pranevich, Y. Karaleu, Analytical forms of productions functions with given total elasticity of production. International Conference on Information Systems Architecture and Technology, Springer, Cham, 2019, pp. 276-285.

[18] O. Hutsaliuk, O. Yaroshevska, O. Kotsiurba, A. Navolokina, “Exploring financial parameters and innovative orientation of banks as criteria for selecting financial partners for enterprises”, Banks and Bank Systems, vol. 15, pp. 118–131, 2020.

[19] A. G. Ivakhnenko, H. R. Madala, Inductive learning algorithms for complex systems modeling, London, Tokyo: CRC Press, 1994.

[20] State Statistics Service of Ukraine, “Population of Ukraine”, 2019. [Online]. Available: http://www.ukrstat.gov.ua.