Doi: https://doi.org/10.15407/dse2017.03.061
UDC 005.95/.96, 519.2
JEL CLASSIFICATION: C35, C55, J08, J64

Y.I. YURYK
PhD (Economics), Senior Researcher
Institute for Economic and Forecasting
of the National Academy of Sciences of Ukraine
26, Panasa Myrnoho, Kyiv, 01011, Ukraine
E-mail: yarina79@ukr.net

G.G. KUZMENKO
PhD (Economics), Expert of Retail Risk Modelling
Alior Bank S.A.
38 D, Lopushanska, Warszawa, 02-232, Poland
E-mail: agkuzmenko@ukr.net

THE POSSIBILITIES OF ESTIMATING RISK EVENTS DURING STRATEGIC MANAGEMENT OF HUMAN RESOURCES

The approach to solution of predicting, classification and risk events diagnosis problems on the labour market during strategic management of human resources are proposed, which has been tested on assessing the risk of unemployment among working population of Ukraine. Specifically, the authors has built a scoring model, which takes into account the joint influence of socio-demographic and professional-and-qualification characteristics of employees, and calculates points based on which it ranks the employees by the risk of the loss of work. It has been discovered that a portrait of employee with the highest probability of «bad events» is the following: single male, aged 15–22, living in rural areas, with profession according to diploma (certificate) – qualified agriculture and forestry employee, skilled tool worker, person working in maintenance, exploitation and monitoring of technological equipment, while being employed in another job, mainly performing the simplest tasks in such economic areas as agriculture and construction.

The scoring model was built using the method of binary logistic regression and the R, SPSS and MS Excel software.

On the basis of the model, one can not only structure the process of preparing possible solutions for risk management, but also carry out a preliminary assessment of the significance of the employee's processed characteristics associated with the likelihood of risk events.

A monitoring of the built scoring model is carried out in order to assess the risk of unemployment among working population of Ukraine. Based on the testing using such parameters as stability, discriminatory power (ranking efficiency) and calibration quality, the author confirmed the model's good predictive ability and adequate functioning.

The model for estimation of probability of unemployment among the employed population of Ukraine is presented primarily as an example of scoring application in HR field. The future prospects of creating such probabilistic models, such tool can be relevant for state institutions, for example employment bureau, as well as *employers, i.e. all parties involved in creation and implementation of HR management strategies. From large amount of data they accumulate on a daily basis the knowledge base for making conscious, not intuitive, strategic decisions and tactical steps can be obtained.*

Keywords: *strategic management, human resources, risk, scoring model, unemployment.*

Я.І. Юрик
канд. екон. наук, старш. наук. співроб., старш. наук. співроб.
Інститут економіки та прогнозування НАН України
01011, Україна, м. Київ, вул. П. Мирного, 26
E-mail: yarina79@ukr.net

Г.Г. Кузьменко
канд. екон. наук, експерт з моделювання ризику
АТ «Аліор банк»
02-232, Польща, м. Варшава, вул. Лопушанська, 38Д
E-mail: agkuzmenko@ukr.net

МОЖЛИВОСТІ ОЦІНЮВАННЯ РИЗИКОВИХ ПОДІЙ У ПРОЦЕСІ СТРАТЕГІЧНОГО УПРАВЛІННЯ ЛЮДСЬКИМИ РЕСУРСАМИ

Запропоновано підхід до вирішення задачі прогнозування, класифікації та діагностики ризикових подій на ринку праці в процесі стратегічного управління людськими ресурсами, котрий апробовано у ході оцінювання ризику безробіття серед зайнятого населення України. Зокрема, побудовано скорингову модель, яка, враховуючи спільний вплив соціально-демографічних і професійно-кваліфікаційних характеристик зайнятих, розраховує бали і на їх підставі ранжує працівників за рівнем ризику втрати ними роботи. Виявлено, що найвищу ймовірність настання «поганої» події мають працівники чоловічої статі віком 15–22 роки, не одружені, які проживають у сільській місцевості, з професією (спеціальністю) згідно з дипломом (посвідченням) – кваліфікований робітник сільського та лісового господарств; кваліфікований робітник з інструментом; робітник з обслуговування, експлуатації та контролю за роботою технологічного устаткування, якщо при цьому вони зайняті не за фахом, а виконують найпростіші роботи у таких видах економічної діяльності як сільське господарство або будівництво. Розроблення скорингової моделі здійснено за допомогою методу бінарної логістичної регресії з використанням програм R, SPSS та MS Excel. На базі розробленої моделі можна не тільки структурувати процес підготовки варіантів рішень з управління ризиком, а й виконати попередню оцінку значущості досліджуваних характеристик зайнятого, пов'язаних із імовірністю настання ризикової події.

Виконано моніторинг побудованої скорингової моделі для оцінки ризику безробіття серед зайнятого населення України. На підставі тестування за такими параметрами як стабільність, дискримінаційна здатність (ефективність ранжування) та якість калібрації підтверджено хороші прогностичні можливості й адекватність функціонування моделі. Побудована модель оцінки ймовірності безробіття серед зайнятого населення України є прикладом застосування скорингу в сфері HR. У перспективі подібні ймовірнісні моделі можуть стати актуальним інструментом діяльності державних інституцій, наприклад, служби зайнятості, і роботодавців, тобто всіх тих суб'єктів, які є учасниками формування та реалізації стратегій управління людськими ресурсами. Зі щоденно нагромаджуваних у них даних можна отримувати знання для свідомого, а не інтуїтивного формування стратегічних рішень і планування тактичних кроків.

Ключові слова: *стратегічне управління, людські ресурси, ризик, скорингова модель, безробіття.*

Я.И. ЮРИК
канд. экон. наук, старш. науч. сотруд., старш. науч. сотруд.
Институт экономики и прогнозирования НАН Украины
01011, Украина, г. Киев, ул. П. Мирного, 26
E-mail: yarina79@ukr.net

А.Г. КУЗЬМЕНКО
канд. экон. наук, эксперт по моделированию риска
АТ «Алиор банк»
02-232, Польша, г. Варшава, ул. Лопушанская, 38Д
E-mail: agkuzmenko@ukr.net

ВОЗМОЖНОСТИ ОЦЕНИВАНИЯ РИСКА В ПРОЦЕССЕ СТРАТЕГИЧЕСКОГО УПРАВЛЕНИЯ ЧЕЛОВЕЧЕСКИМИ РЕСУРСАМИ

Предложен подход к решению задачи прогнозирования, классификации и диагностики рисковых событий на рынке труда в процессе стратегического управления человеческими ресурсами, апробированный в ходе оценки риска безработицы среди занятого населения Украины. В частности, построена скоринговая модель, которая, учитывая общее влияние социально-демографических и профессионально-квалификационных характеристик занятых, рассчитывает баллы, ранжирующие работников по уровню риска потери ими работы. Показано, что самая высокая вероятность наступления «плохого» события – у работников мужского пола в возрасте 15–22 лет, не женатых, жителей сельской местности, профессия (специальность) которых согласно диплому – квалифицированный рабочий сельского и лесного хозяйств; квалифицированный рабочий с инструментом; рабочий по обслуживанию, эксплуатации и контролю за работой технологического оборудования, если при этом они заняты не по специальности, а выполняют простейшие работы в таких видах экономической деятельности, как сельское хозяйство или строительство. Проведен мониторинг построенной скоринговой модели для оценки риска безработицы среди занятого населения Украины. На основании тестирования по таким параметрам как стабильность, дискриминационная способность (эффективность ранжирование) и качество калибрации подтверждены хорошие прогностические возможности и адекватность функционирования модели.

Модель оценки вероятности безработицы среди занятого населения Украины представлена как пример применения скоринга в сфере HR. Что же касается дальнейших перспектив создания подобных вероятностных моделей, то актуальным такой инструмент может быть в деятельности государственных институтов, например, службы занятости, и работодателей, то есть субъектов, участвующих в формировании и реализации стратегий управления человеческими ресурсами. Ежедневно у них накапливается большой объем данных, из которых можно получать знания, необходимые для принятия осознанных, а не интуитивных стратегических решений и планирования тактических шагов.

Ключевые слова*: стратегическое управление, человеческие ресурсы, риск, скоринговая модель, безработица.*

Introduction. Developers of a human resource management strategy should allow for the likelihood of adverse events whose occurrence may affect the strategy's implementation.

Today one of the most successful examples of risk diagnosis and prediction is scoring. Now it is used mostly in banking. And along with its traditional use [1], it seems relevant to transfer the scoring experience of solving problems to other areas. In particular, the present work explores the possibility of building scoring models to assess risk events in the strategic management of human resources.

Dismissal and unemployment are the threats encountered, in the context of development and implementation of the strategies of human resource management, at all levels – country, administrative unit or company. In this context, and taking into account the available information, the essence of our investigation was the creation of scoring models to assess the risk of unemployment among the working population of Ukraine. The results of simulation

[1] Thus, using the tool of scoring enables a financial institution to assess the borrowers' credit rating (application scoring); the likelihood of loan repayment (behavioral scoring); the potential for full or partial repayment in case of the violation of repayment terms (collection scoring); the probability that a new client is not a fraud (fraud scoring); consumer response to a new offer (response scoring); the probability of future use of a banking product or switching to another provider (attrition scoring) [1, p. 255].

will give the basis for pattern recognition and automatic classification of workers by «job loss potential», which in turn will streamline the development of preventive measures to reduce the problems of unemployment in this country.

Overview of the literature suggests that the works dealing with credit scoring – of both theoretico-methodological and practical purpose – are numerous and their number continues to grow (see e.g. H. Abdou & J. Pointon [2]; R. Anderson [3]; D. Hand & W. Henley [4]; A. Kaminsky & K. Pysanets [5, 6]; E. Lewis [7]; Y. Liu [8–10]; E. Mays et al. [11]; N. Siddiqi [12]; S. Sohn et al. [13]; L. Tomas et al. [14, 15]). The published results of the applications scoring models in the non-banking sector are mostly focused in clinical practice and marketing (e.g. E. Kebebew et al. [16]; E. Malthouse [17, 18]; K. Milchakov & M. Shebalkov [19] and other), while the field of HR is presented by individual calculations. In particular, among the relevant are studies of V. Nadraga [20], in which the author considers approaches to estimation of social risks, including the risk of unemployment. This work is the first attempt to create a scoring system that takes socio-demographic and occupational characteristics of workers to estimate probability of their transition to unemployed status, which has determined the **aim and objectives of this research.**

Main results. Modern methods of building scoring models are supported by a variety of predictive analysis tools that belong to a broad class of technologies for advanced data analysis (data science).

The main predictive analysis tools include: statistical methods (linear and multiple logistic regression); classification tree or recursive partitioning algorithm; and neural network [1, p. 255–256].

At the core of our scoring model is the method of binary logistic regression, which can detect the dependence of a dichotomous variable on several independent factors. Among the latter, according to the research objective and theoretical hypotheses, we selected such employee's characteristics as age, place of residence, gender, marital status, level of education, obtained qualification, profession, and field of activity. The dependent variable takes two values – unemployed or busy and has a binomial distribution (Table 1).

To build any scoring models, one must have an information base that, firstly, is sufficient by volume and quality, and, secondly, has a historical antiquity. For our study, those conditions were met for raw data of the monthly sample surveys of population (households) about economic activity for 2010–2013. And the years 2014–2015 were used as a time interval for monitoring, i.e. the procedure to diagnose the adequacy of operation of the developed model.

Before model construction, it is necessary to give a precise definition of the «bad» case, and the whole of the information base:

firstly, should be divided into two groups: the primary sample comprising the observations on whose basis the model itself will be created, and the test sample including data that will not participate in the simulation, but will be used for initial validation, i.e. checking the quality of the model's predicting capacity prior to its use. Primary and test samples are based on random selection mechanism and typically at a ratio of 70–80 % and 30–20 % respectively to the original volume of the total data set;

and, *secondly*, should be converted into a suitable form for further analysis. There are two main approaches used when working with both quantitative and qualitative variables. The first one is the transformation of each output characteristic value into a separate binary variable. But this approach, although considered methodologically simple, is inconvenient as it requires a large number of variables. The second approach, which we actually used, implies conversion of each output characteristic value into an amount equal to the logarithm of ratio of the percentage of respective «good» cases to that of «bad» ones: $Ln(\text{Good}_i/\text{Bad}_i)$. As a result, each of the output characteristic values obtains a numerical parameter, which corresponds to its «riskiness». The parameters of the information base prepared for analysis are summarized in Table 2.

Table 1. **A set of possible variables for the scoring model to predict the risk of unemployment among the working population of Ukraine**

Employee's characteristics	**Output values (of employee's characteristic)**
Status of economic activity	1 – unemployed or «bad»; 0 – employed or «good» (for more detail see Table 2)
Age	**Quantitave change**: from 15 to 59 years old (i.e. able bodied persons)
Place of residence	**Qualitative change:** 1 – urban; 2 – rural
Gender	**Qualitative change:** 1 – female; 2 – male
Marital status	**Qualitative change:** 1 – married; 2 – unmarried; 3 – divorced; 4 – widowed; 5 – unmarried under 18
Educational level	**Qualitative change:** 1 – higher; 2 – basic higher; 3 – incomplete higher; 4 – complete secondary; 5 – basic secondary; 6 – primary general; 7 – no primary
Profession (specialty) according to the certificate (diploma)	**Qualitative change:** 1 – legislators, senior officials, managers, stewards; 2 – professionals; 3 –specialists; 4 – technical staff; 5 – trade and services staff; 6 – skilled workers of agriculture and forestry, fish farming and fishing; 7 – skilled workers with a tool; 8 – workers engaged in servicing, operation and control of the work of processing equipment, assembly equipment and machinery; 9 – elementary occupations; 10 – courses graduates; 11 – no profession
Main work profession	**Qualitative change:** 1 – legislators, senior officials, managers, stewards; 2 – professionals; 3 – specialists; 4 – technical staff; 5 – trade and services staff; 6 – skilled workers of agriculture and forestry, fish farming and fishing; 7 – skilled workers with a tool; 8 – workers engaged in servicing, operation and control of the work of processing equipment, assembly equipment and machinery; 9 – elementary occupations
Main work activity	**Qualitative change:** 1 – agriculture, forestry, hunting; 2 – fish farming and fishing; 3 – mining industry; 4 – production and distribution of electricity, gas and water; 6 – construction; 7 – wholesale and retail trade; 8 – hotels and restaurants; 9 – transport and communication; 10 – financial activities; 11 – real estate; 12 – public administration; 13 – education; 14 – health care; 15 – other types of economic activity (incl. public and personal services, household activities, activities of extraterritorial organizations, work abroad, activities of households as producers of goods and services for personal consumption).

Source: compiled by the authors based on sample surveys of the population (households) on economic activity.

Table 2. **The main parameters of the information base for estimating the risk of unemployment for Ukraine's employed population**

The period when data were collected for: • **model building;** • **monitoring the model's performance**	**2010–2013** **2014–2015**		
Definition of «bad» event	Transfer of the employed to unemployed status at least once in the next 14 months [2] (hence the «good» case is that which is not «bad»).		
Data description*	**Sample for model construction**		**Sample for monitoring**
	primary (80 %)	**testing (20 %)**	
Number of observations	19514	4866	5747
Number of «bads»	815	200	222
Bad rate	4.18 %	4.11 %	3.86 %

* Information base prepared using the software of R.
Source: author's calculations based on sample surveys of the population (households) on economic activity.

Before direct modeling, an analysis of the scoring variables and testing of their predictive power are carried out, for which purpose the index of their *information* value (IV) is calculated. The higher the information value of the variable, the more weight it has in terms of usefulness during the model construction.

Based on the calculations, it was found (Table 3), that, despite different predictive power, all of the factors may be selected for scoring (being IV in each case greater than 0.02), and the most significant in predicting likelihood of the occurrence of our «bad» case are such worker's characteristics as age, profession (actual) and main work activity.

Table 3. **Evaluation of the prognostic power of variables in the probabilistic simulation of unemployment risk among employed population of Ukraine**

No.	Variable	IV
1.	Age	0.156
2.	Main work profession	0.147
3.	Main work activity	0.129
4.	Place of residence	0.094
5.	Marital status	0.087
6.	Educational level	0.083
7.	Profession (specialty) according to the certificate (diploma)	0.083
8.	Gender	0.041

Source: author's calculations based on sample surveys of the population (households) on economic activity.

[2] 14 months: the period is conditioned by specific features of sample surveys of the population (households) on economic activity: the monthly volume of the households sample totality is formed considering the scheme of rotation at which each selected household is interviewed 6 times: 3 months in a row – a 9-month break – another 3 months in a row.

One of the conditions of logistic regression is the absence of multicollinearity. In case of its presence, in order to build the model, it is necessary to find the best options for the exclusion of closely correlated variables.

Analysis based on pairwise correlation matrix revealed existence of a close relationship between employee's educational level and his profession both according to diploma (correlation coefficient 0.932) and according to main work (0.692); besides, a relatively high correlation between the qualification obtained from education and actually performed work (correlation coefficient 0.688) was discovered. Thus, the model will include primarily such variables as employee's age, place of residence, marital status, gender, profession according to diploma and activity according to main work. And, in order to eliminate multicollinearity and keep the maximum information load conveyed by such factors as educational level and actual occupation, and in so doing, raise predictive power of the future model, a new variable «whether the employee performed the work according to obtained qualification» was introduced. The answer to this question (yes / no) in combination with data about profession according to diploma allows to form some idea of both the educational level of the worker's formal training, and in fact about the actually performed work.

Coefficients of the logistic regression, which connects socio-demographic and professional qualification characteristics of employed population in Ukraine with the probability of their unemployment, are presented in Table 4.

An important characteristic of any statistical model is its reliability, which in the case of logistic regression is characterized by the ability to distinguish «good» cases from «bad» ones. Gini index, Kolmogorov–Smirnov test and area under the ROC curve (Table 5) are the main indicators for assessing the quality of the model's classification capabilities.

Given the fact that our logistic regression is an application model with a relatively small number of input variables, the obtained values of parameters (Table 5) show a good quality of binary classifier.

Table 4. **The logistic regression coefficients for estimating the risk of unemployment among employed population of Ukraine**

Variable	Coefficient	Wald statistic	Sig. level of the Wald statistic*
Age	–0.785	58.772	0.0000
Main work activity	–0.543	21.618	0.0000
Place of residence	–0.577	20.247	0.0000
Profession (qualification) according to diploma	–0.614	18.254	0.0000
Correspondence of actual work to obtained qualification	–0.572	10.274	0.0010
Marital status	–0.348	6.189	0.0130
Gender	–0.459	5.839	0.0160
Constant	–3.129	7010.158	0.0000

* If the significance level of Wald statistic is less than 0.05, then the variable is useful for model.
Source: author's calculations based on sample surveys of the population (households) on economic activity.

Table 5. **Indicators for quality evaluation of the model for estimating the risk of unemployment among employed population of Ukraine**

Indicator	Primary sample (80 %)	Test sample (20 %)
Gini index [3]	34.6 %	33.6 %
Kolmogorov–Smirnov test [4]	26.1 %	23.0 %
Area under the ROC curve (AUC [5])	0.673	0.668

Source: author's calculations based on sample surveys of the population (households) on economic activity.

The final stage of scoring model development is scaling (calibration), i.e. a technique of converting primary scoring points into a scoring scale, which is more convenient to use, in particular, with a range from 0 to 1000. The results of this process are transformed into a scorecard (Table 6).

The result of the model's calibration is assigning each employee a rating class according to the risk of occurrence of «bad» event. For this purpose, we chose an appropriate scale that best reflects the constructed model (Table 7).

Quality of the model's calibration can be verified using tests, whose task is to determine to what extent the true values of the studied parameter correspond to the forecasted ones.

Table 6. **Scorecard for estimating the risk of unemployment among employed population of Ukraine**

Variable	Values	SCORE	Comment
Age	15–22	44	The younger the worker, the higher the probability of occurrence of «bad» event
	23–25	68	
	26–47	83	
	48–51	98	
	52–55	110	
	56–59	130	
Main work activity	Agriculture, forestry and fishery; Construction	74	For those employed in agriculture or construction, risk of «bad» event is higher
	Manufacturing; Financial activities; Real estate and business services; Other economic activities	81	
	Governance	85	
	Trade. Hotels and restaurants; Transport and communications	92	
	Mining industry; Electricity, gas and water; Education	102	
	Health care and social assistance	114	
Place of residence	Rural	75	In rural areas, risk of «bad» event is higher
	Urban	95	

[3] Gini index moves the values of area under the ROC curve to the range from 0 to 1 (or 0–100 %).
[4] Value range of Kolmogorov–Smirnov statistics is from 0 to 100 %.
[5] Calculated value range of the indicator of area under the ROC curve (AUC) may be within the interval of 0.5–1.

Ending of Table 6

Variable	Values	SCORE	Comment
Profession, qualification according to certificate (diploma)	Skilled workers of agriculture and forestry, fish farming and fishing; Skilled workers with tools; Workers in servicing, operation and control of the work of processing equipment, assembly equipment and machinery	73	Skilled workers engaged in agriculture; those with tools; those engaged in servicing, operation and control of processing equipment are professions with a higher risk of «bad» event
	Technical staff; Trade and services staff; Elementary occupations; Course graduates; No profession	81	
	Specialists	89	
	Legislators, senior officials, managers, stewards; Professionals	103	
Working according to obtained qualification	No	80	Those engaged not in accordance with obtained qualification have a higher risk of «bad» event
	Yes	96	
Marital status	Unmarried	74	Unmarried workers have higher risk of «bad» event
	Married	87	
	Divorced; Widowed	93	
Gender	Male	80	Male workers have higher risk of «bad» event
	Female	91	

Source: author's calculations based on sample surveys of the population (households) on economic activity.

One such test uses a standard normal distribution. Testing is carried out for each rating class and overall sample and may be unilateral or bilateral. In terms of management, the most important is the risk of underpredicting the probability of losing a job, so Table 8 shows the results of detection of the corresponding significant deviations in the calibration of the constructed model.

Table 7. **Scale to determine the employee's rating class according to the risk of occurrence of «bad» events**

Total points	Probability that the employee will be unemployed at least once during the next 14 months	Rating class
(0; 530]	11.67 % and more	A7
(530; 560]	7.28–11.67 %	A6
(560; 590]	4.46–7.28 %	A5
(590; 620]	2.70–4.46 %	A4
(620; 650]	1.62–2.70 %	A3
(650; 680]	0.97–1.62 %	A2
(680; 1000]	0–0.97 %	A1

Source: author's calculations based on sample surveys of the population (households) on economic activity.

Table 8. **Testing the quality of the model's calibration for estimating the risk of unemployment among employed population of Ukraine, %***

Rating class	Bad $rate_{observed}$	Bad $rate_{forecast}$	$BR_o - BR_f$	Tolerance ranges for significance level			
				90 %	95 %	99 %	99.9 %
A7	13.64	14.60	–0.97	2.06	2.64	3.73	4.96
A6	6.89	8.37	–1.48	0.85	1.09	1.54	2.05
A5	6.53	5.54	0.99	0.43	0.55	0.77	1.03
A4	3.41	3.45	–0.05	0.30	0.38	0.54	0.71
A3	2.00	2.19	–0.19	0.28	0.36	0.51	0.68
A2	1.17	1.36	–0.19	0.37	0.47	0.67	0.89
A1	0.00	0.81	–0.81	0.72	0.93	1.31	1.74
All classes	4.18	4.18	0.00	0.18	0.24	0.33	0.44

* Calibration Test using Standard Normal Distribution (one-sided test).
Source: author's calculations based on sample surveys of the population (households) on economic activity.

Based on the obtained data, we can conclude on the good quality of the model's calibration, while a deviation (underestimation of risk) is observed only in one of the seven classes A5 (Table 8)[6].

Ranking of employed population by risk level of «bad» event is shown on Fig. 1, and considering scorecard data (Table 6) it was discovered, that the highest probability (18.2 %) of becoming unemployed at least once in every 14 months have male workers aged 15–22 years, unmarried, living in rural areas and having the following professions (specialties) (according to diploma (certificate)): skilled workers of agriculture and forestry; skilled workers with a tool; workers engaged in servicing, operation and control of work process equipment, who at the same time are not employed by profession. Actually, they are workers who perform the simplest works in agriculture or construction.

People, circumstances, and conditions may vary over time due to dynamic social and economic development. In the language of probabilistic modeling and in the context of the present study, this means that the influence of certain employee's characteristics on the probability of unemployment does not remain constant. And, for scoring model to continue functioning, periodical check of quality of its performance is required.

Monitoring of our rating system for adequacy of calculated risk parameters implied its diagnosing by such criteria as stability, discriminatory power (ranking efficiency) and the quality of calibration.

Stability tests are carried out to confirm compliance of the current data composition to the sample that is used for model development. Significant differences in the data set could reduce the quality of the model performance and hence require its re-building.

[6] Binomial test can be used as alternative for testing compliance between estimated probability and actual unemployment, which returned a similar result.

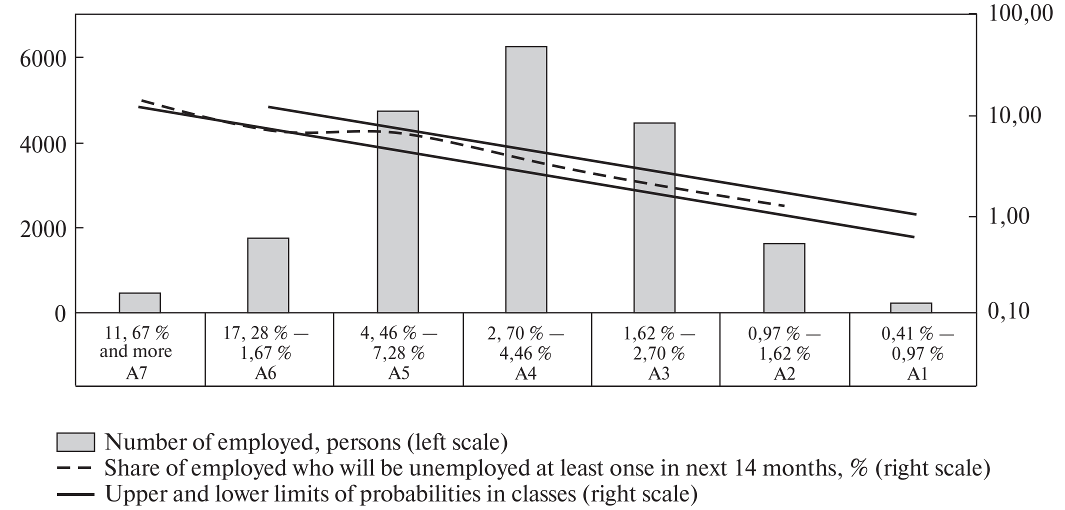

Figure 1. **Rating based classification of the employed population depending on the level of risk of their transition to the status of unemployed**

Source: constructed by the authors based on sample surveys of the population (households) on economic activity.

Based on the calculated index PSI (*Population Stability Index*), we conclude on the absence of difference between the samples used to build and monitor the model for evaluating the probability of unemployment among employed population in Ukraine (Table 9).

As a rule, discriminatory power of the model decreases over time since model's development. The scale of such decrease should be assessed in terms of acceptability, tracking the change of the indicators used to assess the model's classification possibilities during its construction, which are, as we mentioned before, the Gini index, the Kolmogorov–Smirnov test and the area under the ROC curve (Table 10).

Table 9. **Monitoring the model for estimating the likelihood of unemployment among employed population in Ukraine: a test for stability**

No.	Variable	PSI*
1.	Age	0.0052
2.	Obtained qualification	0.0042
3.	Main work activity	0.0083
4.	Place of residence	0.0307
5.	Marital status	0.0001
6.	Profession (specialty) according to certificate (diploma)	0.0093
7.	Gender	0.0000

* PSI < 0.1 – no difference; PSI – [0.1; 0.2] – insignificant differences; PSI > 0.2 – significant differences.
Source: author's calculations based on sample surveys of the population (households) on economic activity.

Table 10. **Monitoring the model for estimating the likelihood of unemployment among employed population in Ukraine: discriminatory power**

Indicator	Sample for model development	Sample for model monitoring
Gini index	34.6 %	34.2 %
Kolmogorov–Smirnov test	26.1 %	24.9 %
area under the ROC curve (AUC)	0.673	0.671

Source: author's calculations based on sample surveys of the population (households) on economic activity.

As can be seen on Table 10, the studied indicators have only changed slightly (6 %), hence the effectiveness of the model's ranking of the depending on the level of risk of their transition to unemployed status has not deteriorated.

For the purposes of monitoring, the quality of calibration is defined using standard tests with normal and binomial distributions in the same way as during the model development.

It should be reminded that, in terms of management, the most relevant is the risk of underprediction the probability of the employee's job loss and, as confirmed by the results of model testing, in any of the classes and in overall sample, the probability of «bad» event is not underestimated (Fig. 2).

However, in the monitoring sample in the rating class A7, i.e. among the employed with higher «potential» of losing their jobs, the likelihood of unemployment is slightly overestimated (Fig. 2, Tab. 11).

This situation is well expected, because the monitored period (2014–2015) coincides with a crisis in economy of Ukraine. This means that, during that time, people were less likely to look for better jobs and tried to keep their current positions.

To sum up, let's note that testing by criteria of sample's stability, discriminatory power and calibration quality confirms the adequacy of the model of estimating the probability of unemployment among employed population of Ukraine and currently no evidence of the deterioration in its performance is discovered.

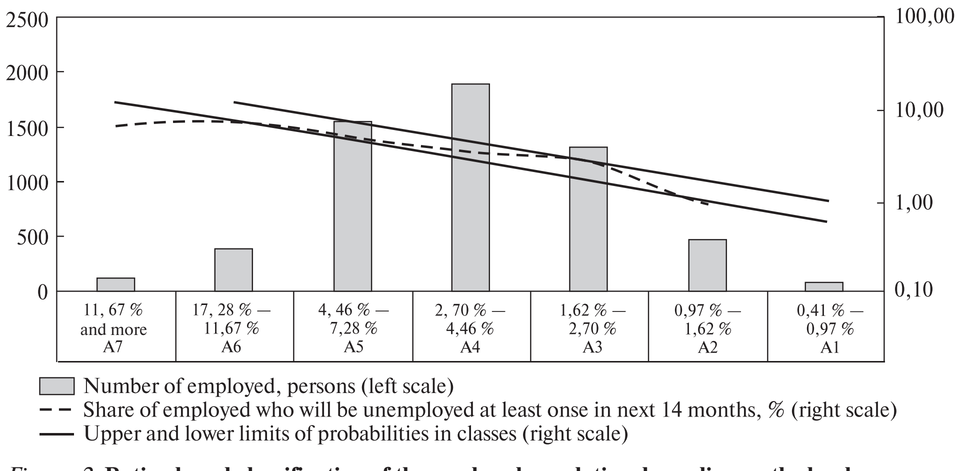

Figure. 2. **Rating based classification of the employed population depending on the level of risk of their transition to unemployed status (a monitoring sample)**

Source: constructed by the authors based on sample surveys of the population (households) on economic activity.

Table 11. **Model monitoring for estimating the likelihood of unemployment among employed population of Ukraine: quality of calibration, %***

Rating class	Bad $rate_{observed}$	Bad $rate_{forecast}$	$BR_f - BR_o$	Tolerance ranges for significance level			
				90 %	95 %	99 %	99.9 %
A7	6.36	14.66	8.29	4.32	5.55	7.84	10.42
A6	8.07	8.78	0.71	1.85	2.38	3.36	4.46
A5	5.24	5.70	0.46	0.76	0.98	1.38	1.83
A4	3.40	3.47	0.07	0.54	0.69	0.98	1.30
A3	2.75	2.19	–0.56	0.57	0.67	0.94	1.25
A2	0.86	1.37	0.51	0.69	0.89	1.25	1.66
A1	0.00	0.84	0.84	1.39	1.78	2.52	3.35
All classes	3.86	4.14	0.28	0.34	0.43	0.61	0.81

* Calibration Testusing Standard Normal Distribution (one-sided test).
Source: author's calculations based on sample surveys of the population (households) on economic activity.

Conclusion. Thus, for the first time, the development of a scoring model for solving the tasks of forecasting, classification and diagnosis of risk events on the labor market was introduced. In particular, the model, which is based on the joint influence of socio-demographic and professional qualification characteristics of employees, calculates points and on their basis ranks the employed population in terms of risk (probability) of the loss of job (unemployment).

Results of the testing of the model's quality and empirical evaluation confirming the theoretical assumptions actually prove the possibility and prospects of the application of scoring for the diagnosis and prediction of risk events on the labor market.

On the basis of the model, one can not only structure the process of preparing possible solutions for risk management, but also carry out a preliminary assessment of the significance of the studied employee's characteristics associated with the probability of «bad» event.

Let's note, that our model for estimating the probability of unemployment among employed population of Ukraine is represented primarily as an example of the application of scoring – a tool that empowers HR-analysts for the purpose of strategic management of human resources and allows a transition from retrospective descriptive data to the level of forecast.

As for the range of tasks and applications of such probabilistic models, their use can be relevant in the activities of both public institutions, such as employment services and employers, i.e. all those entities who participate in the formation and implementation of strategies of the management of human resources and who daily gather and accumulate large amounts of data from which they can «extract» the knowledge to make strategic decisions and plan tactical steps not intuitively, but on the basis on the latter.

The combination of HR-data (socio-demographic and professional qualification characteristics of workers, data on the key performance indicators, information on career and administrative change, information about the progress and success of training, etc.) makes scoring a suitable tool for the needs of human resource management in the context of strategic perspective. Moreover, it makes available not only application scoring, but also behavioral scoring, which takes into account motivational factors that determine individual's labor behavior and the associated probability of the occurrence of certain investigated events.

LITERATURE

1. *Юринець Р.В.* Економетрична модель оцінювання кредитного позичальника відповідно до експертної оцінки / Р.В. Юринець // Науковий вісник НЛТУ України: зб. наук.-техн. праць. – Вип. 19.5. – Львів : РВВ НЛТУ України, 2009. – С. 254–258.
2. *Abdou H.A.* Credit scoring, statistical techniques and evaluation criteria: A review of the literature / Abdou H.A., Pointon J. // Intelligent systems in accounting, finance and management. – 2011. – № 18 (2–3). – P. 59–88.
3. *Anderson R.* The credit scoring toolkit: theory and practice for retail credit risk management and decision automation / Anderson R. – New York : Oxford University press, 2007. – 790 p.
4. *Hand D.J.* Statistical classification methods in consumer credit scoring: a review / Hand D.J., Henley W.E. // Journal of the Royal Statistical Society, Series A. – 1997. – 160. – P. 523–541.
5. *Камінський А.Б.* Моделювання фінансових ризиків : монографія / Камінський А.Б. – К. : ВПЦ «Київський університет», 2006. – 304 с.
6. *Камінський А.Б.* Скорингові технології в кредитному ризик-менеджменті / А.Б. Камінський, К.К. Писанець // Бізнес Інформ. – 2012. – № 4. – С. 197–201.
7. *Lewis E.M.* An introduction to credit scoring / Lewis E.M. – San Rafael : The Athena Press, 1992. – 172 p.
8. *Liu Y.* New issues in credit scoring application / Liu Y. // Research paper, Institute of Information Systems, University of Goettingen. – 2001. – № 16. – 35 p.
9. *Liu Y.* A framework of data mining application process for credit scoring / Liu Y. // Research paper, Institute of Information Systems, University of Goettingen. – 2002. – № 1. – 45 p.
10. *Liu Y.* The evaluation of classification models for credit scoring / Liu Y. // Research paper, Institute of Information Systems, University of Goettingen. – 2002. – № 2. – 59 p.
11. Handbook of credit scoring / Mays E. (Ed.). – Chicago : Glenlake Publishing Company Ltd/Fitzroy Dearborn Publishers, 2001. – 382 p.
12. *Siddiqi N.* Credit risk scorecards: developing and implementing intelligent credit scoring / Siddiqi N. – New Jersey : John Wiley and Sons, 2006. – 208 p.
13. *Sohn S.Y.* Technology credit scoring model with fuzzy logistic regression / Sohn S.Y., Kim D.H., Yoon J.H. // Applied Soft Computing. – 2016. – 43. – P. 150–158.
14. *Tomas L.C.* A survey of credit and behavioral scoring: forecasting financial risk of lending to consumers / Tomas L.C. // International Journal of Forecasting. – 2000. – 16. – P. 149–172.
15. *Thomas L.C.* Credit scoring and its applications: SIAM monographs on mathematical modeling and computation / Thomas L.C., Edelman D.B., Crook J.N ; SIAM. — Philadelphia, USA, 2002. – 248 p.
16. *Kebebew E.* Predictors of single-gland vs multigland parathyroid disease in primary hyperparathyroidism: a simple and accurate scoring model / Kebebew E., Hwang J., Reiff E., Duh Q.Y., Clark O.H. // Archives of Surgery. – 2006. – 141 (8). – P. 777–782.
17. *Malthouse E.C.* Ridge regression and direct marketing scoring models / Malthouse E.C. // Journal of Interactive Marketing. – 1999. – 13 (4). – P. 10–23.
18. *Malthouse E.C.* Assessing the performance of direct marketing scoring models / Malthouse E.C. // Journal of Interactive Marketing. – 2001. –15 (1). – P. 49–62.
19. *Мильчаков К.С.* Скоринговые карты в медицине: обзор и анализ публикаций / К.С. Мильчаков, М.П. Шебалков // Врач и информационные технологии. – 2015. – № 1. – С. 71–79.
20. *Надрага В.І.* Соціальні ризики: сутність, аналіз, можливості впливу : монографія / В.І. Надрага, НАН України. Ін-т демографії та соціальних досліджень ім. М.В. Птухи. – Київ : ПП Сердюк В.Л., 2015. – 329 с.

REFERENCES

1. Yurynets, R.V. (2009). Ekonometrychna model otsiniuvannia kredytnoho pozychalnyka vidpovidno do ekspertnoi otsinky [Ekonometric Model of Evaluation of Credit Borrower Accordingly to Expert Estimation]. N*aukovyi visnyk NLTU Ukrainy - Scientific Bulletin of Ukrainian National Forestry University, 19.5*, 254-258 [in Ukrainian].
2. Abdou, H.A., & Pointon, J. (2011). Credit Scoring, Statistical Techniques and Evaluation Criteria: A Review of the Literature. *Intelligent Systems in Accounting, Finance and Management, 18, 2-3*, 59-88.

3. Anderson, R. (2007). *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*. Oxford: Oxford University Press.
4. Hand, D.J., & Henley, W.E. (1997). Statistical Classification Methods in Consumer Credit Scoring: A Review. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 160*, 3, 523-541.
5. Kaminsky, A.B. (2006). *Modeliuvannia finansovykh ryzykiv [Modeling of Financial Risk].* Kyiv: Kyiv National Taras Shevchenko University [in Ukrainian].
6. Kaminsky, A.B., & Pysanets, K.K. (2012). Ckorynhovi tekhnolohii v kredytnomu ryzyk-menedzhmenti [Scoring Technologies in Credit Risk-Management]. *Biznes Inform - Business Inform, 4*, 197-201 [in Ukrainian].
7. Lewis, E.M. (1992). *An Introduction to Credit Scoring.* San Rafael, CA: Athena Press.
8. Liu, Y. (2001). *New Issues in Credit Scoring Application*. Research paper 16/2001. Institute of Information Systems. University of Goettingen.
9. Liu, Y. (2002). *A Framework of Data Mining Application Process for Credit Scoring.* Research paper 01/2002. Institute of Information Systems. University of Goettingen.
10. Liu, Y. (2002). *The Evaluation of Classification Models for Credit Scoring.* Research paper 02/2002. Institute of Information Systems. University of Goettingen.
11. Mays, E. (Ed). (2001). *Handbook of Credit Scoring*. Chicago: Glenlake Pub.
12. Siddiqi, N. (2006). *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring*. Hoboken. NJ: Wiley.
13. Sohn, S.Y., Kim, D.H., & Yoon, J.H. (2016). Technology Credit Scoring Model with Fuzzy Logistic Regression. *Applied Soft Computing, 43*, 150-158.
14. Thomas, L.C. (2000). A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers. *International Journal of Forecasting, 16*, 2, 149-172.
15. Thomas, L.C., Crook, J., & Edelman, D. (2002). *Credit Scoring & its Applications: Siam Monographs on Mathematical Modeling and Computation*. Siam. USA.
16. Kebebew, E. (2006). Predictors of Single-gland vs Multigland Parathyroid Disease in Primary Hyperparathyroidism. *Arch Surg Archives of Surgery, 141,* 8, 777-782.
17. Malthouse, E.C. (1999). Ridge Regression and Direct Marketing Scoring Models. *Journal of Interactive Marketing, 13*, 4, 10-23.
18. Malthouse, E.C. (2001). Assessing the Performance of Direct Marketing Scoring Models, *Journal of Interactive Marketing, 15*, 1, 49-62.
19. Milchakov, K.S., & Shebalkov, M.P. (2015). Skoringovyie kartyi v meditsine: obzor i analiz publikatsiy [Scorecards in Medicine: Analytic Review]. *Vrach i informatsionnyie tehnologii - Doctor and Information Technology, 1*, 71-79 [in Russian].
20. Nadraga, V.I. (2015). *Sotsialni ryzyky: sutnist, analiz, mozhlyvosti vplyvu [Social risks: nature, analysis, ability to influence].* Kyiv: Serdyuk V.L. [in Ukrainian].

Article received on 06.07.2017.

Стаття надійшла до редакції журналу 06.07.2017.