E3S Web of Conferences **254**, 10005 (2021) https://doi.org/10.1051/e3sconf/202125410005
FARBA 2021

Results analysis of agricultural organizations subsidy

R.B. Shestakov*, *N.A.* Yakovlev, *E.I.* Lovchikova, *G.P.* Zvereva, and *A.S.* Volchenkova,

FSBEI HE Orel State Agrarian University named after N.V. Parakhin, Orel, Russia

Abstract. The article presents some issues of the methodology of the study of state support for agricultural production in the aspect of efficiency. The dynamic series of profitability for certain types of products sold by agricultural organizations with and without budget subsidies are used as source data. First, a preliminary analysis and pairwise comparison of the series using nonparametric methods is carried out. Further, in order to use complete information on all the data, the methods of reduction to three variables are applied. In order to compare different methods, the principal component method and autoencoding using a neural network were used. The transformed data was also compared with and without subsidies. Conclusions and practical recommendations on the results obtained, as well as directions for the development of scientific research, are formulated.

1 Introduction

The specifics of the agro-industrial complex in any country is its rather strong dependence on state support. At the same time, monitoring of the results of various forms of support and determining its significant effect or effectiveness is a very difficult aspect.

One of the most important forms of state support for agricultural production is subsidy from the budget. The methodology of the Ministry of Agriculture of the Russian Federation for assessing the effectiveness of subsidy use involves determining it as the level of achievement of the result of subsidy use, the ratio of the actual value of the result of subsidy use at the end of the reporting year to the planned value of the result of subsidy use [1]. Nevertheless, this method is quite subjective and ambiguous. Some authors, for example, suggest searching the dependence of the gross value added created in agriculture on the federal budget expenditures and the total expenditures of regional budgets [2] and observe a significant effect. Others consider it mandatory to study not only economic, but also social, environmental and other impacts [3]. All approaches have their advantages and disadvantages, and, as a result, the study of support effects is a voluminous and non-trivial task that requires a complex of various approaches.

In this work, we propose a general method for determining the presence of a significant statistical effect from the use of subsidies by agricultural organizations using indicators of product profitability.

* Corresponding author: nir_paper@inbox.ru

2 Materials and methods

The indicators of product profitability sold by agricultural organizations were used as the initial data (based on the consolidated reports on financial and economic activities of large and medium-sized agricultural organizations receiving state support from the federal budget) in Russia for the period 2014-2018 [4] (Table 1). The data are given for the main types of products with and without subsidies from the budget: grain, sunflower seeds, sugar beet, potato, vegetables, milk, livestock and poultry, wool, eggs.

The product profitability is a coefficient that shows the ratio of profit to the cost of product production and sales (cost price). This kind of indicator signals the amount of profit that will give the company one ruble invested in the production process. Profitability can be calculated both for the organization as a whole, and for individual areas of activity, types of products.

The given dynamic series will be used for pairwise comparisons using the corresponding methods given below.

Table 1. Profitability of products sold by agricultural organizations, %

Years	Grain	Seeds of sunflower	Sugar beet	Potato	Vegetables	Milk	Cattle	Pigs	Sheep, goats	Wool	Eggs
Without subsidies from the budget											
2014	24.3	48.6	38.6	34.9	17.8	23.7	-35.9	36.6	-7.9	-56.3	12.8
2015	39.5	90.9	78.9	23.9	26.6	19.5	-27.6	28.5	-7.9	-49.5	17
2016	32.8	70.5	56.2	4.7	7.4	18.5	-29.9	19.7	-9.2	-39.5	13.5
2017	18.6	42	13.2	19.8	4.1	25	-30.8	23.8	-12.5	-40.3	5.8
2018	25.6	33.2	27.6	22.9	12.6	14.5	-30.8	35.2	-8.4	-37.1	9.2
with subsidies from the budget											
2014	30.5	52.8	40.5	38.2	19.5	33	-33	37.3	7.4	-56.2	13.6
2015	44.9	94.1	80.9	26.9	29.1	26.6	-25.1	29	9.1	-46	17.5
2016	37.0	73.1	58.1	5.8	9	28.2	-27.3	19	5.3	-26.7	13.9
2017	21.4	42.2	13.4	22.3	7.5	32.3	-28.7	24.1	2.0	-27.8	6.6
2018	29.0	33.3	27.8	26.9	16.6	23.9	-28.5	35.8	11.0	-23.7	10.2

Source: according to the FSSS of the Russian Federation [4]

Studying the data in Table 1, the problematic points concerning the analysis of the source data can be immediately noticed. First of all, this is a very small sample with a significant number of variables (features), which will require a special approach for analysis. A special comparative nonparametric test is needed to determine the significant differences in each attribute depending on the accounting for the budget subsidy. In this situation, the Kruskal-Wallis criterion is suitable (H-test). It is used to test the null hypothesis that the median of all groups is equal. To increase the test efficiency, the significance threshold was raised up to 10%.

In order to use the most complete information contained in all the series, we apply the methods of dimension reduction. The most well-known of these is the principal component analysis (PCA) method. It is one of a class of "unsupervised" statistical learning algorithms used to explain high-dimensional data by reducing the data into a smaller number of features called principal components. The PCA preserves a significant amount of information, which can be controlled by setting the number of components.

Further, to compare the results, we use another method of dimension reduction, known as "autoencoding". This is, in fact, a special architecture of artificial neural networks [5,6]. The right part after the "bottleneck" (Fig.1), which reduces the dimension, is used only at the training stage of the model.

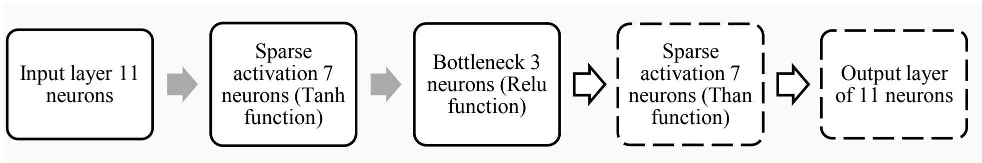

Fig. 1. Configuration of the used auto-encoder in the Python software environment [7]

Source: compiled by the authors

An automatic encoder without hidden layers is almost equivalent to the linear main component method, and when using additional layers of neurons, it makes it possible to consider non-linear connections and interactions. The main difficulty of autoencoding, like any other neural network, is the selection of architectonics and hyperparameters.

3 Results and discussion

A preliminary or "exploratory" analysis (exploratory data analysis, EDA) of the data is shown in table 2. It includes the definition of the mean, median, and standard deviation. A preliminary pairwise test was also performed between the same products for the presence of significant differences.

Table 2. Exploratory Data Analysis (EDA)

Indicators	Grain	Seeds of sunflower	Sugar beet	Potato	Vegetables	Milk	Cattle	Pigs	Sheep, goats	Wool	Eggs
Without subsidies from the budget											
mean	28.2	57.0	42.9	21.2	13.7	20.2	-31.0	28.8	-9.2	-44.5	11.7
median	25.6	48.6	38.6	22.9	12.6	19.5	30.8	28.5	-8.4	-40.3	12.8
standard deviation	8.1	23.4	25.5	10.9	8.9	4.2	3.0	7.2	1.9	8.1	4.3
With subsidies from the budget											
mean	32.6	59.1	44.1	24.0	16.3	28.8	-28.5	29.0	7.0	-36.1	12.4
median	30.5	52.8	40.5	26.9	16.6	28.2	28.5	29	7.4	-27.8	13.6
standard deviation	8.9	24.6	26.3	11.8	8.7	3.8	.9	7.7	3.5	14.3	4.1
H-test											
Sample difference hypothesis (Kruskal-Wallis criterion, $p < 10\%$	-	-	-	-	-	+	-	-	+	-	-

Source: calculated by the authors

Table 2 shows that the medians and mean values of the product series when taking into account the subsidy differ more, as expected, although with different "intensity". The largest standard deviation is shown by "sugar beet" in both parts of the table, and the smallest - "sheep and goats" without subsidies, cattle - with subsidies.

Statistically significant differences within the specified limits (+) are observed only in milk producers and in a separate segment of animal husbandry. This can also be seen by the apparent difference in the medians, for example, for milk 19.5 and 28.2%, respectively. Moreover, for goat and sheep breeding, the transition from the negative zone to the positive one is noticeable.

Before starting the analysis for the main components, the data was previously normalized using the Yeo-Johnson transformation, which, in fact, is analogous to the Box-Cox transformation, but works with zero and negative values of variables.

Table 3. Analysis results and transformed data using PCA

Periods	**Components and transformed data**		
	1	**2**	**3**
Without subsidies from the budget			
Dispersion load distribution (96%)	47%	30%	19%
2014	0.10	-2.66	-1.59
2015	3.71	0.31	-0.15
2016	0.31	2.73	-0.21
2017	-3.42	0.73	-0.78
2018	-0.70	-1.11	2.73
With subsidies from the budget			
Dispersion load distribution (95%)	49%	31%	15%
2014	0.58	3.06	-0.80
2015	3.65	-0.50	-0.30
2016	0.31	-2.67	-0.21
2017	-3.51	-0.40	-1.19
2018	-1.03	0.51	2.51
Sample difference hypothesis (Kruskal-Wallis criterion, $p < 10\%$	-	-	-

Source: calculated by the authors

In table 3, it can be seen that the load distribution for the main components is identical for data with and without subsidies. All components collected 96 and 95% of the volatility of the original array, respectively. The first components containing 47 and 48% are the most informative. Nevertheless, the comparative analysis failed to capture significant differences in the new variables.

As mentioned above, the auto-coding algorithm can capture non-linear relationships in the data, and thus increase the information content of the results. To proceed to the analysis, we will first scale the data in the range from 0 to 1. We will also focus on the three output components to preserve the relevance of the previous analysis.

The results of processing with the auto-encoder are shown in Table 4.

Table 4. Analysis results and transformed data using auto-coding

	New features and transformed data		
	1	**2**	**3**
Without subsidies from the budget			
2014	-0.29	0.86	0.15
2015	0.65	-0.01	0.53
2016	0.15	-0.51	0.42
2017	-0.14	-0.26	-0.62
2018	0.88	0.02	-0.54
2014	-0.65	0.28	0.35
2015	-0.91	-0.94	-0.97
2016	-0.16	-0.89	-0.96
2017	0.32	0.13	-0.75

2018	0.65	-0.83	0.10
Sample difference hypothesis (Kruskal-Wallis criterion, $p < 10\%$	-	-	-

Source: calculated by the authors

A comparative analysis of the results of table 4 also showed that there were no significant differences between the components of the data with and without subsidies.

Thus, let's summarize the results of the analysis based on the available data:

1) There are no significant differences (with rare exceptions) in the dynamics of profitability in certain sectors of agricultural production, with and without subsidies from the budget.

2) There are also no significant differences in the aggregate information in the dynamics of profitability with and without subsidies from the budget

First, when discussing the results, the question arises about the number of subsidies and how organization profitability is sensitive to their changes. Can a noticeable difference in profitability arise by increasing the subsidy? Apparently, a more significant leap for agribusiness profitability lies not only in the absolute figures of the subsidy, but also in the ability of organizations to use it more efficiently.

Previous studies of the authors have shown a noticeable disparity in the development and support of certain sectors of agricultural production and the influence of price factors [8]. Moreover, in addition to subsidies, it is necessary to make greater use of other support methods, for example, those related to taxation [9]. Structural and innovative factors are needed in a slightly different plane, but also affecting profitability. The use of digital ecosystems can transform traditional agricultural organizations into new types of organizations, which will increase the profitability of agriculture, the investment attractiveness of the agro-industrial complex, the employment of the rural population, and the level of rural development [10]. The direct impact on productivity and cost reduction is particularly important during periods of instability and general economic slowdown.

The study of differences in the "weighted" profitability can be specified as directions for further research. This methodology assumes that the share of the industry in the total volume of agricultural production is taken into account, and therefore the "impact" of changes in profitability for agribusiness as a whole. In addition, an interesting study will be the "differences of the differences themselves" for other types of farms (small forms).

4 Conclusions

The authors tested the methodology for determining the impact of subsidies on the results of agricultural organizations, where the dynamic series of profitability with and without state subsidies are used as the initial data. The preliminary analysis and pairwise comparison of the variables using nonparametric methods is carried out. The methods of dimension reduction to three features are applied to use complete information on all the features. For the purpose of comparison, two different methods were used, including the principal component method and autoencoding using a neural network. Further, the obtained data are also compared using nonparametric methods. Almost no statistically significant differences were found. The problems of profitability increase, in addition to subsidies, lie in other areas of the agricultural economy.

References

1. Methodology for assessing the effectiveness of subsidy use, Russian Federal State statistics service (2020)
2. A.M. Kalinin, V.A. Samokhvalov, Problems of Forecasting, 5, 142 (2020)
3. A.V. Golubev, The economy of agricultural and processing enterprises, 9, 13 (2019)
4. Agriculture in Russia 2019 statistical collection (2019)
5. S. Gu, B. Kelly, D. Xiu, Journal of Econometrics (2020)
6. J. Zhang, H. Wang, H. Yang, Journal of PLA University of Science and Technology, 17(1), 31 (2016)
7. Tensorflow API (2020)
8. R. B. Shestakov, N.A. Yakovlev, IOP Conf. Ser.: Mater. Sci. Eng, 753 (2020)
9. O.A. Pankratova, Taxes and Taxation, 3, 1 (2020)
10. A.I., Solodovnik, V.K. Dokalskaya, Bulletin of Agrarian Science, 6(87), 150 (2020)