World Applied Sciences Journal 5 (2): 177-181, 2008
ISSN 1818-4952

Classification of the Variables Used in Wheat Production According to Their Productivity by Discriminant Analysis: the Case of Thrace Region

S. Konyali

Department of Agricultural Economics,
Faculty of Agriculture, Namik Kemal University, Tekirdag, 59030, Turkey

Abstract: The purpose of this study is to determine if the variables used in wheat production are used productively or not. Within this scope, it was aimed to find out which districts in the research area used their inputs and outputs productively or unproductively by discriminant analysis. In the research, with the sampling method carried out in the Thrace region, it was found out that this research had to be conducted in 131 villages. In each village interview with 2 producers was seen enough and 262 questionnaires were conducted in 131 villages. The producers within the sample were randomly selected. In this analysis, according to the scale regarding the input from the questionnaire results, data was computed according to both constant (CRS) and variable (VRS) returns. Firstly, efficiency was computed according to CRS in the discriminant analysis. According to the function results of CRS model, it was found out that the districts used their inputs and outputs more unproductively. Because the amounts of seed and fertilizer were determined to be mostly used unproductively when input and output (production) coefficients were taken into consideration. As a result of the function, the districts used their inputs and outputs more unproductively with coefficient differences of 0.4 for seed, 0.5 for fertilizer, 0.1 for the amount of land and 0.2 for the amount of production. It was determined that the rate of the districts that were classified correctly was 84.2% in the productive and unproductive groups. Then efficiency according to VRS was computed. VRS model was tried for the districts to use their inputs and outputs more productively and the results were determined to be more positive. According to the function results of VRS (variable return to scale) model, it was found out that the districts used their inputs (excluding the amount of land) and outputs more productively. Because it was seen that the inputs and outputs were bigger than Y_0 function coefficients when the coefficients of Y_1 function were considered. The districts mostly used the amounts of seed and fertilizer productively. As a result of the function, the districts used their inputs and outputs more productively with coefficient differences of 1.7 for seed, 0.6 for fertilizer and 0.3 for the amount of production. Hence, according to the function results, when VRS model was used, districts used their inputs and outputs more productively than CRS model. As a result of the research, it was found out that the most important variables to classify the districts in the Thrace region as productive/unproductive were the amounts of seed and fertilizer.

Key words: Wheat production · Productivity · Discriminant analysis · Thrace region

INTRODUCTION

Wheat in Turkey is cultivated on an area of nearly 9.3 million hectares and the production is nearly 20 million tons despite the yearly change. In Turkey, wheat is produced on approximately 35% of the general agricultural lands. Productivity per decare was 220 kg as of the season of 2006/2007.

Wheat in the Thrace region is an important product for the region since it is one the most important source of income of the producers. In the region, usually seed and fertilizer were used in wheat production. Nitrogen fertilizer is important for region. According to the another research that was made in Thrace region, lands were insufficient of organic material and sufficient of phosphorus [1]. During the year 2007, in the region wheat was cultivated on an area of approximately 550 000 hectares and nearly 2,5 million tons were produced and the productivity per decare was 440-465 kg. Despite the yearly change, the Thrace region has approximately 6% of the total wheat

Corresponding Auhtor: S. Konyalı, Department of Agricultural Economics, Faculty of Agriculture, Namık Kemal University, Tekirdağ, 59030, Turkey

cultivable lands and 10.9% of wheat production in Turkey. Wheat production in the Thrace region (409 kg/da) doubles that of Turkey and it is one and a half times more than the world average (280 kg/da). Therefore, the Thrace region provides great contribution to the agriculture and economy in Turkey in terms of wheat cultivation, production and particularly, productivity.

In this study, it was aimed to determine if the variables used in wheat production were used productively or unproductively according to the questionnaire results conducted to the producers in the Thrace region, which is an important production region in the study. For this purpose, discriminant analysis was used. Discriminant analysis is a method that derives functions which divide the variables of X data set into two or more real groups and optimally appoint these units to the real groups, classes in the natural media according to the p value of these units [2]. In this analysis, according to the questionnaires responded by the producers in the Thrace region, input variables were determined as the amounts of seed, fertilizer and land and output variable was determined as wheat production.

MATERIALS AND METHODS

Research material was gathered from primary and secondary data resources. The information obtained from the questionnaires responded by the wheat producers in the Thrace region constituted the primary data of the research. Any kind of published researches, books, statistics and literature research in the field of wheat production, consumption, trade and economy constituted the secondary data resources of the research.

Research method was evaluated under two titles as sampling and data collection and economical analysis method.

Sampling and Data Collection Method: Since the Thrace region consists of 3 provinces, 25 districts and 721 villages, the volume of samples were narrowed because it was thought that sample group that would represent the region would involve a lot of producers. For this purpose, districts and villages were determined by contacting the Directorates of Ministry of Agriculture of Edirne, Tekirdağ and Kırklareli. The amounts of village lands were obtained from the Directorates of Ministry of Agriculture. Because the villages with a cultivable land of below 1000 decares and above 30 000 decares disrupt the normal distribution, extreme values were excluded out of the sampling. The parameters regarding the finite population formed in this way were computed and given below:

N = 702 (number of villages)
μ = 6532,8 da (average wheat cultivable land)
σ = 5039 da (population standard deviation)
D = 653,28 (sampling error)

According to the data above, sample volume was computed with the formula given below [3,4].

Sampling error (or estimation error) was accepted as (μ)±%10 of the arithmetical mean. In other words, the arithmetical mean of the sample could be different from the population mean [5]. On the other hand, the possibility of the sampling mean to be within this limit is accepted to be 90%.

$$n = \frac{[Z_{\alpha/2}\, \sigma / D]^2}{1 + 1 / N. [Z_{\alpha/2}\, \sigma / D]^2}$$

where:
N = Main mass (population)
n = Sample number
$Z_{\alpha/2}$ = z value (1.645)
σ = Population standard deviation
D = Sampling error (%10)

$$n = \frac{[1.645.\ 5039 / 653.28]^2}{1+1 / 702. [1.645.5039 / 653.28]^2} = 130.9 \cong 131$$

In the calculation carried out according to the formula above, it was found out that this research should be carried out in 131 villages. In each village, interview with 2 producers was seen enough and 262 questionnaires were conducted in 131 villages. The producers within the sample were randomly selected.

As a result of the sampling carried out, the 131 villages within the scope of the sampling were distributed by the proportional sampling method according to the proportions of wheat cultivable lands of 3 provinces composing the Thrace region (Tekirdağ, Edirne and Kırklareli) during the year 2006. According to this, the provinces composing the research area and the number of the samples selected from the villages are shown in Table 1.

As understood from the Table 1, more numbers of farmers (123 farmers) were interviewed since the amount of cultivable lands are more in Tekirdağ province, whereas interview with 68 farmers in Kırklareli, which has lower amount of lands was seen enough.

Table 1: Chosen of villages that applied questionnaire in the research

Provinces	Total area of cultivable land (da)	Ratio of lands(%)	No. of producer that applied guestionnaire
Tekirdağ	2309120	47	123
Edirne	1332233	27	71
Kırklareli	1313865	26	68
Total	4955218	100.0	262 √

Economical Analysis Method: All data were obtained from questionnaires that were suitable for their subject matters. All the information gathered from the producers were examined one by one and entered into databases by means of several programs according to the purpose of the research. For totally 262 questionnaires applied to the wheat producers, a general database was formed and a general coding plan was made according to the questions asked. Questionnaires were entered into the computer according to this coding plan.

As a result of the questionnaires, arithmetical means of the amounts of seed, fertilizer and land per decare were computed as the inputs of the districts and arithmetical mean of the wheat production was computed as the output of the districts. After the arithmetical means were computed according to the districts, the data was entered into the computer for the application of the discriminant analysis. In the study, districts were divided into two groups as productive and unproductive. After the data was computed separately according to both constant and variable returns regarding the input scale, efficiency was computed first for the CRS and then for the VRS in the discriminant analysis. CCR (Constant Return to Scale: CRS) models are used to determine the total efficiency of the decision units under constant return assumption according to the scale. BCC (Variable Return to Scale: VRS) models measure the efficiency score under variable return according to the scale. Efficiency scores found under this assumption are called as technical efficiency.

RESULTS AND DISCUSSION

After suitable selections were made for discriminant analysis, data and results obtained are shown in the tables below.

According to the test results in Table 2, group covariance matrices are homogenous and linear discriminant analysis can be applied.

As seen in Table 3, data can be separated by a discriminant function.

According to the results in Table 4, the discriminating power of the function is significantly high. (P= 0.049*).

Table 2: Result of test

Box's M		30.337
F	Approx.	2.205
	df1	10
	df2	1065.418
	Sig.	.016

Tests null hypothesis of equal population covariance matrices.

Table 3: Eigenvalues

Function	Eigenvalue	Variance %	Cumulative	Canonic Correlation
1	.889[a]	100.0	100.0	.686

a. First 1 canonical discriminant functions were used in the analysis.

Table 4: Wilks' Lambda

Function(s) Test	Wilks' Lambda	Chi-Square	df	Sig.
1	.529	9.540	4	.049

Table 5: Coefficient of Classification Functions

	productive	
	.00	1.00
seed	8.235	7.837
fertilizer	5.871	5.380
land	.848	.768
production	-1.635	-1.442
Constant	-166.805	-148.210

Fisher's linear discriminant functions

Classification functions in Table 5 are as follows; Y_0 = -166,805 + 8,235X_1 (seed) + 5,871X_2 (fertilizer) + 0,848X_3 (land) – 1,635X_4 (production) for unproductive 0 and Y_1 = -148,210 + 7,837X_1 (seed) + 5,380X_2 (fertilizer) + 0,768X_3 (land) – 1,442X_4 (production) for productive 1.

According to the function results of CRS model, it was found out that districts used their inputs and outputs more unproductively. Because the amounts of seed and fertilizer were determined to be used unproductively most when input and output coefficients were taken into consideration. Although the producers used their lands more unproductively according to Y_1 function, the coefficient of this was lower than those of seed and fertilizer. The effect of the amount of land on land productivity or unproductivity was very low in both functions. It was determined that the districts used their production amounts more unproductively as a result of these inputs. As a result of the function, the districts used their inputs and outputs more unproductively with coefficient differences of 0.4 for seed, 0.5 for fertilizer, 0.1 for land and 0.2 for the amount of production.

In Table 6, 11 units were correctly classified for unproductive 0 and 5 units for 1. There was no incorrect classification for 0, but 3 units were incorrectly classified for productive 1. In other words, originally classified

Table 6: Results of Classification

		Correct Classification Rate		
productive		.00	1.00	Total
Original Number	.00	11	0	11
	1.00	3	5	8
%	.00	100.0	.0	100.0
	1.00	37.5	62.5	100.0

a. 84,2 % of original grouped cases correctly classified

Table 7: Eigenvalues

Function	Eigenvalue	Variance %	Cumulative	Canonic Correlation
1	4.102[a]	100.0	100.0	.897

a. First 1 canonical discriminant functions were used in the analysis

Table 8: Wilks' Lambda

Function(s) Test	Wilks' Lambda	Chi-square	df	Sig.
1	.196	24.444	4	.000

Table 9: Coefficient of Classification Functions

	VAR00001	
	.00	1.00
seed	8.876	10.594
fertilizer	5.000	5.607
land	.318	.066
production	-.228	.499
Constant	-162.295	-236.164

Fisher's linear discriminant functions

Table 10: Results of Classification

		Correct Classification Rate		
VAR00001		.00	1.00	Total
Original Number	.00	17	0	17
	1.00	0	2	2
%	.00	100.0	.0	100.0
	1.00	.0	100.0	100.0

a. 100,0 % of original grouped cases correctly classified

samples, that is, the rate of the districts that were correctly classified in the productive and unproductive groups was 84.2%.

Then, efficiency according to VRS was computed in the discriminant analysis. After suitable selections were made for the discriminant analysis, data and results were obtained as shown below.

As seen in Table 7, data can be separated by a discriminant function.

According to the results in Table 8, the discriminating power of the function is significantly high. ($P= 0.000^{*}$).

Classification functions in Table 9 are as follows; $Y_0 = -162,295 + 8,876X_1$ (seed) $+ 5,000X_2$ (fertilizer) $+ 0,318X_3$ (field) $- 0,228X_4$ (production) for unproductive 0 and $Y_1 = -236,164 + 10,594X_1$ (seed) $+ 5,607X_2$ (fertilizer) $+ 0,066X_3$ (field) $+ 0,499X_4$ (production) for productive 1.

VRS model was tried for the districts to use their inputs and outputs more productively and the results were determined to be more positive.

According to the function results of VRS (variable return to scale) model, it was seen that the districts used their inputs and outputs more productively. Because it was found out that the inputs and outputs were bigger than Y_0 coefficients when the coefficients of Y_1 function were considered. The districts mostly used the amounts of seed and fertilizer productively. As a result of the function, the districts used their inputs and outputs more productively with coefficient differences of 1.7 for seed, 0.6 for fertilizer and 0.3 for the amount of production. Hence, according to the function results, when VRS model was used, districts used their inputs and outputs more productively than CRS model.

In Table 10, all the districts were correctly classified in the groups 0 and 1, that is, in the productive and unproductive groups. In other words, regarding their productive and unproductive levels, none of the 19 phenomenon (districts) was incorrectly classified.

Finally, while considering the function results of these models, the districts which used the inputs productively or unproductively can be discriminated by this analysis. In addition, as a result of this analysis, it was found out that the most important variances to classify the districts in the Trakya region as productive/unproductive were the amount of use of seed and fertilizer.

REFERENCES

1. Bellitürk, K. and T. Sağlam, 2005. A Research On The Urea Hydrolysis Rate In The Soils of Turkey. Pak. J. Biol. Sci., 8(3): 446-449.
2. Özdamar, K., 2004. Statistical Data Analysis with Package Programs (Multivariate Analysis). Renewed. 5. Edition, Kaan Bookstore, Eskişehir.
3. McClave, J.T. and P.G. Benson, 1988. Statistics for Business and Economics. 4th Edn. Dellen Pub Co., San Fransisco.
4. Arıkan, R., 1995. Research Techniques and Report Writing Tubitay Publications, Ankara
5. Gujaratı, D.N., 2001. Fundamental Econometrics, (Translated by: Ümit-Gülay Günlük ªenesen), Literatür Publications:33, Literatür Publishing, Istanbul.

6. Deligöz, A. and A. Gezer, 2006. Anatolian Black pine [Pinus nigra Arn. subsp. pallasiana (Lamb.) Holmboe] Statistical Analysis of Separation of Seed Stands according to the Characteristics of Some Cones and Seed. Süleyman Demirel University, Journal of Faculty of Forestry Series: A, No:2, ISSN: 1302-7085, Page: 13-22, Isparta.
7. Grene, H.W., 2000. Econometric Analysis. Fourth Edition, Prentice-Hall, Inc., New Jersey.
8. Johnson, E.D., 1988. Applied Multivariate Methods for The Data Analysts, Duxbury Pres.
9. Kayalıdere, K. and S. Kargın, 2004. Efficiency Study in Cement and Textile Sectors and Data Envelopment Analysis. Dokuz Eylül University, Journal of Institute of Social Sciences Volume:6, No:1, Izmir.
10. Yıldız, E., 2006. Economical Analysis of the Effects of Monetary and Fiscal Policies Applied in Turkey on Agricultural Production. T.Ü. Institute of Science, Dissertation, Edirne.