METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH
X, 2009, str. 107 - 120

THE USEFULNESS OF MULTINOMIAL LOGIT MODELS IN EXAMINING THE RELATIONSHIP BETWEEN THE OWNERSHIP FORM OF AGRICULTURAL ENTERPRISES AND THEIR FINANCIAL-ECONOMIC RESULTS

Stanisław Jabłonowski, Andrzej Kluza
Katedra Informatyki, SGGW
e-mail: staja@mail.sggw.pl

Summary: From the year 1994 onwards, the magazine "Nowe Życie Gospodarcze" has been publishing financial-economical data of the best three hundred agricultural enterprises selected from groups of such enterprises that had chosen to participate. Makers of these lists have taken into account enterprises which have come into being from the property of state treasury in the nineteen nineties. The present paper takes up the problem of utilizing (to a rather limited extent) of multinomial logit models for studying the relation between the ownership form of agricultural enterprises and some financial-economic characteristics, describing economic effectiveness.

Keywords: agricultural enterprises, ownership forms, economical effectiveness indicators, multinomial logit models

STUDY OBJECTIVE

Our investigation is restricted to the so-called “top 300 lists”, presented by “Nowe Życie Gospodarcze”. In this data, we find a description of economic entities that have arisen after the year 1989 from the estate that had remained in the hands of A.W.R.S.P.’s (then A.N.R.[1]). In the paper, the question is raised whether it is possible to build econometric models where the qualitative explained variable corresponds to the form of property, while explanatory variables are constituted by financial-economic characteristics of agricultural industry enterprises. We ask also weather building such models makes sense, since it is not possible to use these models to observe causal links between the explanatory variables and the explained variable (which here is not an effect but cause of differences in the financial-

[1] A.W.R.S.P. = Agency for Agricultural Property of State Treasury; A.N.R. = Agricultural Property Agency

economic results). So, we want to focus on examination of existence of statistical links between features, ignoring indices of cause and effect. The causal link between explanatory variables and the explained one is important in qualitative variables' models [Cramer, 2001]. An approach presented below will be useful, in spite of that we obviously won't profit fully of such models. In the case of our data there exists a number of obstacles in building an econometric model in which a form of ownership would be an explanatory or one of explanatory variables (it will be explained later). So, analyzing the populations' data, it would be necessary to use common descriptive statistics only, and to use also some measurements of dependences between features.

The bodies from "top 300 lists" have various ownership and legal forms. Models are built separately for each year of the eleven year-time period under analysis. The models consider the set of three hundred entities featured by "Nowe Życie Gospodarcze" on the "top 300 list" for a given year.

The authors attempt to study the following idea: does a statistical link between economic performance and the ownership form of companies in these sets exist or not? We want to know which financial-economic features are significant in the constructed models. Does the utilization of models allow us to obtain greater benefits regarding the analysis of the sets under investigation or do we arrive at better results from the usual descriptive statistics?

An econometric model with a qualitative explained variable with many variants may take the form of a multinomial logit model. This form has been used in the paper.

DATA DESCRIPTION

Lists of three hundred best companies and other agricultural entities that had come into being from the property of state treasury have been scientifically processed by the I.E.R.i G.Ż (Institute of Agricultural Economics and Food Economy), while the initiators of their preparation have been the A.W.R.S.P. (Agency for Agricultural Property of State Treasury) [now A.N.R. - Agricultural Property Agency] and the magazine "Nowe Życie Gospodarcze".

The first list concerned the year 1994 and included only two hundred units. However, in the following years, the lists described three hundred best companies of the agricultural sector (from among those that would respond to the survey). The entities taken into account have been enterprises that had emerged from the property of state treasury and that remained in the hands of the A.W.R.S.P (restructured into the A.N.R. in 2003). They received varying ownership and legal status.

Rankings of companies were drawn up on the basis of responses to questionnaires sent out to a large number of units (e.g., to 2110 units in 1998). Usually, only less than 20% addressees would dispatch their answers (for example, in 1998, 386 entities completed the survey). The basic rating was originally created on the basis of the added value indicator, then on the basis of the profitability index and, since 2000, according to a special synthetic measurement tool (then slightly modified).

Data concerning the best companies have been published in „Nowe Życie Gospodarcze" and supplemented by comments in the form of articles written by eminent specialists in the field of agricultural economics [e.g. Leopold, Ziętara, 1999]. The sole standard descriptive statistics have been used in the comments. The data was used in many analyses presented in other titles [e.g. Franc-Dąbrowska, 2008; Grontkowska, 2007]. This paper utilizes eleven lists from the years 1995 - 2005.

In individual lists, the following information is given for particular enterprises: current position in the rating and the position from the previous year (if the object was featured on the list the previous year), name of the company and its voivodship of origin (unless there was no consent), the form of property management (ownership), the EKD (later PKD) code which denotes the dominant socio-economic activity of an economic entity. Further, the features' definitions and their designations have been given. The designations come from the authors' research of this paper and are used in the further description. The absence of certain digits in the numbering is due to the fact that some features published in the ratings have been neglected, since they do not contribute significant to any of the econometric models constructed. Detailed information on the measurement tools and indicators included on the lists can be found e.g. in [Guzewicz et al., 2006]. The successive features are as following:
indicator of profitability of economic activity (operating activity in the 1995) (W6), return on total assets (W7), return on equity (W8), the rate of value added (W9), current ratio at year-end (W10) and quick ratio at year-end (W11), equity to assets ratio (W12), financial results to total debt ratio (W13), labour efficiency (W14), total emoluments (W15), sales income (W16), net financial balance (W19), average employment in year (W20), arable land area (W22) and soil quality class (W23). In addition, two variables have been attached: employment per 100 hectare of arable land (W24) and land productivity (emoluments per 1 hectare of arable land) (W25), calculated on the basis of data from the lists.

MODELLING METHOD

The model type that frequently becomes used in case the feature to be explained has a small number of variants, but still more than two, is the multinomial model of qualitative variables, otherwise known as the model of multiple choice. However, in case of disordered categories, where determination of the order of the nominal feature variations proves impossible, researchers usually choose the logit multinomial model. Such models are used in different fields. For example Mikhail, Walther and Willis [1998] use such a model to research the phenomenon if a change of brokers' office by an financial analyzer has had place on effect of the quality of prognosis made by him. Cramer [2003] analyzes different stages of private cars' ownership according to different features of households. Seo and Mendelsohn [2008] describe how South American farmers adapt to climate changes by changing crops by developing a multinomial logit model of farmer's choice of crops. In the present analysis, we are interested in the form of property management that a companies utilize. This feature accepts 9 variants for the data in question. These are forms of ownership (abbreviations used in the "top 300 lists" are given in brackets): the enterprises purchased from the Agency (PZA), mixed units, i.e. these that have purchased some part of the land, while the rest is leased (PM), entities in that a company established by employees holds the lease of the utilized plot (DSP), leases concluded by natural persons (DOF), other leases (DP), companies wholly owned by the Agency (JSA) and administered enterprises (A). Number of administered enterprises (A) covered by the ratings gradually decreased (from 102 in 1995 to 0 in 2004), but since 2001, agricultural production cooperatives (RSP) and the other units (PJ) are also included in the ratings. There are only between 2 and 10 cases of the latter entity (PJ) and this code name usually denotes units that belong to scientific institutions. In this paper, entities with the legal form A (since 2001) and PJ have been disregarded. Overall, it can be said that the entities with the legal form A and JSA have conducted business on behalf of their mother Agency, while the rest - on their own account.

For particular "top 300 lists" from consecutive years, multinomial logit models have been built. The explained variable is a form of ownership, but its new categories are included, having applied four methods of classifying ownership forms – as shown in the tab. 1 (the symbols SP1, SP2, SP3, SP4, that denote the explained variable, correspond to the way of aggregating):

Table 1. Methods of assigning the form of property to the created variants

Property's form → ----------------------- Way of clustering	PZa	PM	DSP, DOF, DP	JSA	A (until 2000)	RSP (since 2001)
Sp1 (4 variants)	1	1	2	3	4	4
Sp2 (5 variants)	1	2	3	4	5	5
Sp3 (3 variants)	1	1	1	2	3	3
Sp4 (3 variants)	1	1	2	3	3	3

Source: authors' decisions

Potential explanatory variables have been mentioned on the "top 300 lists", but in the course of building the models in question some of them have been eliminated. Finally, for each of the eleven years, one model has been left, featuring the explained variable (out of SP1, SP2, SP3 and SP4) that occurred in the model with the best characteristics.

In the section "Study Objective" the following is stated: the resulting models are to describe the sets of 300 companies from the "top 300 lists". Unfortunately, the models cannot be generalized to cover the entire set of enterprises that had arisen from state treasury property managed by the A.W.R.S.P. (and later the A.N.R.) for particular years. The "top 300 lists" cannot be treated as representative samples selected from the entire set because it is not the case, as described above. Therefore, conclusions derived from constructed models refer only to entities listed on the "top 300 list" for the year in question.

MATHEMATICAL APPROACH TO THE MODEL

The applied econometric model type derives from Theil [Theil, 1969]. In this model, the aim is to estimate the values connected with probabilities of obtaining different categories by the explained variable. If Y is a qualitative variable that adopts J+1 variants and if P_{ij} denotes the probability of obtaining the j-th variant by this variable for the i-th element of the sample (population, once generalized), then the multinomial logit model explains the quotients of probabilities $P_{ij}/(P_{i0} + P_{ij})$, j = 1, ... J, i = 1, ... n. Index j = 0 corresponds to the chosen basic variant. This explanation is done by means of a function that represents cumulative distribution of logistic distribution. The argument of this function is a linear combination of explanatory variables and the vector of structural parameters of the model. However, there is a number of structural parameter vectors. The vectors are different for each j (j = 1, ..., J).

Let us denote the vector of structural parameters for the category j of k+1 size by $\boldsymbol{\beta}_j$, j = 1, ..., J, if we assume that there are k explanatory variables (the free coefficient on the first position).
In the case of a model built on the basis of individual data it can be summarized as follows: $\frac{P_{ij}}{P_{i0} + P_{ij}} = F(\mathbf{x}_i^T \boldsymbol{\beta}_j)$, where $\mathbf{x}_i$ is the vector (of dimension k +1) consisted of the number 1 and of explanatory variables' values for the i-th element of the sample (population), i = 1, ..., n , j = 1, ..., J[2].

Since F (.) is the logistic function, finally: $\frac{P_{ij}}{P_{i0} + P_{ij}} = \frac{1}{1 + \exp\left(-\mathbf{x}_i^T \boldsymbol{\beta}_j\right)}$, j = 1, ..., J for i-th object, i = 1, ... n.

[2] As the ratios of certain probabilities are explained and not single probabilities, so it can be shown, that there is no contradiction with the condition $\sum_j P_{ij} = 1$.

Hence, it can be deduced that: $\text{Ln}\frac{P_{ij}}{P_{i0}} = \mathbf{x}_i^T\boldsymbol{\beta}_j$ for each category j = 1, ..., J [Gruszczyński, 2002]. Predicted category for the i-th element of the sample (after generalization, for any element of the population) is selected after evaluation of J vectors of structural parameters' estimates. The estimates of expressions $\text{Ln}\frac{P_{ij}}{P_{i0}}$, j = 1, ..., J, are to be calculated. Then this category of J+1 of the Y variable's variants possible to be chosen is to be selected for which the estimate of the likelihood of its occurrence is the highest.

MODEL VERIFICATION

Evaluating parameters of a multinomial logit model is most often carried out by means of the principle of maximum likelihood [Cramer, 2001]. Most considerations may be regarded as a generalization of the binomial logit model, see e.g. [Dudek et al., 2006].

Long [Long, 1997] states that the maximum likelihood method should be applied for a large sample of at least 100 items, but preferably of about 500. The significance of parameters may be assessed in various ways, the most important being the likelihood quotient test. It is used for a larger set of parameters, while Wald statistics is used for individual structural parameters. All these statistical approaches have a chi-square distribution. Here a description of Wald statistics is skipped, indicating only that it has 1 degree of freedom.

The following provides a more detailed information on the likelihood quotient test. Logarithm of this ratio is calculated by the formula: $LR_j = -2(\ln\hat{L}_{Rj} - \ln\hat{L}_{UR})$, where $\ln\hat{L}_{Rj}$ means the logarithm of maximum of likelihood function for the model with the free coefficient, but which does not include the variable X_j from the k explanatory variables X_1, ..., X_k (or without free coefficient and with all variables, when j = 0). The expression $\ln\hat{L}_{UR}$ is the logarithm of maximum of likelihood function for the full model, i.e. with the free coefficient and k variables.

Statistics LR_j for large samples follows the chi-square distribution with J degrees of freedom. The verified hypothesis is: H_0: $\beta_{1j} = \ldots = \beta_{Jj} = 0$, against the alternative H_1: there exists r from the set {1,…,J}, that $\beta_{rj} \neq 0$. Statistics LR_j serves to its verification and the test is to be carried out separately for the successive variables and the free coefficient (j = 0,1, 2 ,..., k).

The second hypothesis refers to the absence of statistical significance of all structural parameters of the explanatory variables. The verified hypothesis is: H_0: $\boldsymbol{\beta}_r = [\beta_{r0}, 0, \ldots, 0]$ for all r = 1, ..., J, against the alternative H_1: there exists r from the set {1,…,J} and j from the set {1, 2,...,k}, that the parameter β_{rj} is different from zero. And here again the test of likelihood quotient is used to verify the hypothesis. Now it takes the form: $LR = -2(\ln \hat{L}_R - \ln \hat{L}_{UR})$, where $\ln\hat{L}_R$ denotes the logarithm of maximum of likelihood function for the model reduced to only a constant, while $\ln\hat{L}_{UR}$ is the logarithm of maximum of likelihood function for the full model, i.e. with the free coefficient and k variables.

For large samples the LR statistics follows the chi-square distribution with k*J degrees of freedom.

In assessing the compatibility of the model with empirical data a lot of different measurements are used. They are similar, in its conception, to the classical coefficient of determination R^2.

One of frequently used measurements is so-called McFadden pseudo-R^2 [Manski et al., 1977]: $R^2_{McFadden} = 1 - \frac{\ln \hat{L}_{UR}}{\ln \hat{L}_R}$, where $\ln\hat{L}_R$ stands for the logarithm of maximum of likelihood function for the model reduced to only a constant, while $\ln\hat{L}_{UR}$ is the logarithm of maximum of likelihood function for the full model, i.e. with the free coefficient and k variables. $R^2_{McFadden}$, as well as classical R^2, accepts values between 0 and 1, a better model fit is characterized by its higher value.

Similarly, count-R^2 [Gruszczyński, 2002] takes values between 0 and 1, a better model fit is characterized by its higher value. It is the quotient of the number of correct predictions to the number of all elements of the population, for which predictions have been estimated.

After having verified the model, an interesting analysis can be made. E.g. from the formula: $\frac{P_{ij}}{P_{i0}} = \exp(x_i^T \boldsymbol{\beta}_j)$, in which the left side can be called the "odds ratio of j category to the basic category", the following conclusion can be easily derived: $\exp(\beta_{jl})$ informs how many times this odds ratio increases, when X_l increases by a unit, ceteris paribus. This ratio increases with the increase of X_l, if only $\beta_{jl} > 0$.

Similarly, it can be shown that $\frac{P_{ij}}{P_{ik}} = \frac{\exp(x_i^T \boldsymbol{\beta}_j)}{\exp\left(x_i^T \boldsymbol{\beta}_k\right)}$, and therefore we conclude that $\exp(\beta_{jl} - \beta_{kl})$ informs how many times the quotient $\frac{P_{ij}}{P_{ik}}$ increases if the variable X_l is increased by a unit, ceteris paribus. Hence, if $\beta_{jl} > \beta_{kl}$, then with the increase of X_l, the quotient increases, or the probability of category j relative category k increases.

RESULTS OF MODEL EVALUATION

The authors have decided to build models for successive years, involving all potential explanatory variables. Then the aposteriori elimination method has been used to obtain end models that have positively verified hypotheses about the significance (on the level of $\alpha = 0.1$ at the most) of structural parameters (all three hypotheses discussed above). Count-R^2 had to be greater than 50%, while the McFadden pseudo-R^2 had to equal circa 0.20 at least. For each of the eleven years, one model remained – the one with the explained variable (from among SP1, SP2, SP3 and SP4) that has occurred in the model with the best characteristics.

As the sample size for each year was about three hundred (in case of 1995, the data of two companies was removed because of their deficiencies; for some years, the sample size decreased slightly due to the omission of entities of legal form PJ and legal form A since 2001), so it can be stated that the requirements pertaining to the minimum number of statistical units were met [Long, 1997].

Table 2. Test results to prove the lack of statistical significance of all structural parameters of the model

Model	$-2\ln\hat{L}_R$ or $-2\ln\hat{L}_{UR}$	LR	Degrees of freedom	P-value
Only constant	634.31			
The full model	437.17	197.15	12	0

Source: the SPSS software results

The SPSS software version 11.5.0 (16 Nov. 2002) has been used to complete necessary calculations. Results of the model estimated on the basis of data concerning 291 objects from the list of 2005 are presented below. SP4 has been chosen as the explained variable, while the selected explanatory variables are W7, W8, W12 , W14, W20 and W24. Following the description of these results, the summary of findings for all models is also presented.

Table 3. Test results to prove the lack of statistical significance of successive structural parameters of the model

Parameter	$-2\ln\hat{L}_{Rj}$	LR_j	Degrees of freedom	P-value
Constant	443.42	6.25	2	0.044
W7	461.3	24.13	2	0
W8	492.62	55.45	2	0
W12	466.67	29.51	2	0
W14	466.09	28.92	2	0
W20	455.01	17.85	2	0
W24	467.21	30.05	2	0

Source: the SPSS software results

Table 4. Evaluation of parameters (B) for successive categories (SP4 is the explained variable), together with test results to prove the lack of statistical significance for individual structural parameters (3 is the variant of reference)

variant	parameter	B	Wald	degr. of freed.	P-value	95% confidence interval for Exp(B)		
							lower limit	upper limit
1	Constant	1.69	3.11	1	0.08			
	W7	-0.057	4.28	1	0.04	0.944	0.894	0.997
	W8	0.124	18.69	1	0.00	1.133	1.070	1.198
	W12	-0.056	24.45	1	0.00	0.945	0.925	0.967
	W14	0.036	20.29	1	0.00	1.036	1.020	1.052
	W20	-0.013	8.11	1	0.00	0.987	0.979	0.996
	W24	0.05	9.14	1	0.00	1.051	1.018	1.086
2	Constant	2.453	5.87	1	0.02			
	W7	-0.143	18.58	1	0.00	0.867	0.812	0.925
	W8	0.178	36.63	1	0.00	1.195	1.128	1.265
	W12	-0.046	15.53	1	0.00	0.955	0.934	0.977
	W14	0.027	10.82	1	0.00	1.027	1.011	1.043
	W20	-0.013	7.39	1	0.01	0.988	0.979	0.996
	W24	0.055	10.92	1	0.00	1.056	1.023	1.091

Source: the SPSS software results

We shall describe the selected explanatory variables more precisely: W7 is return on total assets (ratio of value added to total assets used by the company in the end of 2005), W8 - return on equity (ratio of net financial result to level of equity in the end of 2005), W12 - equity to assets ratio (ratio of equity to assets in the end of 2005), W14 - labour efficiency (ratio of value added to average employment in 2005), W20 - average employment in the year (average number of people employed in 2005), W24 - employment per 100 hectare of arable land (the variable W20*100 devided by arable land area). McFadden pseudo-R^2 for the model of 2005 is equal to 0.311.

Table 5. Prediction results according to the model for 2005 (the count-R2 can be found in the lower right corner, expressed in per cents)

Categories				
Observed	Predicted			
	1	2	3	Percentage of correct predictions
1	41	22	22	48.24
2	22	56	13	61.54
3	11	5	99	86.09
Percentage totally	25.43	28.52	46.05	**67.35**

Source: the SPSS software results

Table 6. Summary results for all models

Year	Explained variable	McFadden - R^2	Count - R^2	Explanatory variables
1995	Sp4	0.736	93.6%	W12, W16, W20
1996	Sp4	0.655	90%	W6, W7, W9, W10, W12, W13, W20, W23, W24, W25
1997	Sp4	0.602	86.3%	W6, W11, W12,W13, W16, W20
1998	Sp3	0.477	81.7%	W12, W14, W20
1999	Sp3	0.329	79%	W13, W19, W20
2000	Sp4	0.317	68.7%	W6, W9, W12, W20
2001	Sp4	0.324	68.7%	W6, W12, W14, W20
2002	Sp3	0.319	74.7%	W10, W12, W14
2003	Sp3	0.3	74.3%	W7, W12, W13, W20
2004	Sp4	0.187	56.8%	W6, W12, W13, W14
2005	Sp4	0.311	67.4%	W7, W8, W12, W14, W20, W24

Source: The author's own calculations based on the SPSS software results

SUMMARY OF RESULTS AND CONCLUSIONS

As shown by Table 6, in all models the variables explained are SP3 or SP4, so the remaining variables are these concerning three variants of the ownership form (cf. Tab.1). The best model is the model for 1995 data. Models for subsequent years provide a worse match, but McFadden-R^2 is generally higher than 0.3, and count-R^2 usually reaches at least 70%.

Testing the significance of remaining variables has given very good results. It may be stated that verification of statistical models has been positive. Thus, a statistical link is visible between financial-economic performance and the form of ownership of the enterprises from the "top 300 lists". As Table 6 indicates, this relationship is more pronounced in the earlier years of the time period in question. Not in every year the same economic characteristics become significant in the constructed models. The explanatory variables that occur in the greatest number of models are as follows: W12 (10 models), W20 (9), W6 (5), W13 (5), W14 (5), W7 (3). Other variables - W9, W16, W24 – have occurred in two models, and W8, W10, W11, W19, W23 and W25 – in one model. The variables that appears most commonly are: W12, which denotes the ratio of equity to assets at the end of the year - expressed as a percentage - and W20, meaning the average employment in the year (measured in the number of persons).

Interpretation of a model will be shown on the example of the 2005 model. SP4 has become the explained variable. It attains three variants: the first means entities wholly or partially purchased (PZA, PM) , the second - leases (DSP, DOF, DP) and the third - companies wholly owned by the Agency and agricultural production cooperatives (JSA, RSP) (in 2005, administered entities, A, were not present).

The direction of the impact of explanatory variables on the odds ratio of both variant 1 and variant 2 to variant 3 (category 3 is the basic one) remains the same, because the parameters corresponding to the variables retain the same signs for variant 1 and variant 2. And so: an increase in the variable W7 decreased these quotients, while an increase in the variable W8 increased them - and so on, as seen in Table 4. Unfortunately, because of the fact that the ownership form is the cause and not the effect of differencies of values of variables W7 or W8 we only can deduce that W7 reaches greater average values for variant 3 and W8 – lower than for variants 1 and 2.

Due to similar reasons the following analysis for e.g. variable W12 cannot be utilised well. For this variable (the ratio of equity to assets expressed in per cents), the estimated parameters are: β_{1W12}=-0.056, β_{2W12}=-0.046 (both significantly different from zero).

Therefore, the conclusion can be drawn that as the value of variable W12 increases, the quotient of probability of variant 1 to the probability of variant 3 decreases (0.945 times per unit of W12) and the ratio of probability of variant 2 to the probability of variant 3 also decreases (0.955 times per unit W12), ceteris paribus.

One could also presume that with an increase in the value of the variable W12 the ratio of probability of variant 1 to the probability of variant 2 decreases (0.99 times per unit of W12), ceteris paribus. It does not have to be so, however, due to the lack of significant difference between the parameters β_{1W12} and β_{2W12}, as seen in Table 4. One should concentrate on the 291 companies involved in the calculations, out of the three hundred entities found on the list for 2005.

Let us clear up why a form of ownership is not the explanatory variable, as the probable cause of economic discrepancies i.e. why the models show only a quantitative relationship between variables, without specifying the cause and effect relation. In the case of "top 300 lists" the enterprises in the ratings are ordered by a special synthetic measurement tool. This is a kind of weighted average of features

W6, W9, W13 i W14, which, by ratings' authors, best express different aspects of economical effectiveness. Because of many reasons, the approach presented in the present paper is more convenient to use. For example, it would be difficult to determine the explained variable which would show economical effectiveness. If the synthetic measurement instrument tool a.n., used in the analyzed ratings, is supposed to be the explained variable then only one variable (W14) from 4 variables included in it has a significant link to the form of ownership (for the 2005 model).

It would be also difficult to establish optimal grouping of the ownership forms into lower number of variants. They seem to be too numerous in the source data. It is visible from the presented approach that the variable corresponding to the form of property takes into account the optimum allocation of primary forms of ownership into new variants. It becomes clear which variables are significantly linked to the ownership form.

Another important question is whether it would not be sufficient to use the standard approach i.e. to calculate usual descriptive statistics for the analysis of these sets, possibly with carrying out tests to check the discrepancies in averages.

Of course it would be of worth to put into use the known measures of relationship between features [Encyclopedia …, 2008], like rank correlation coefficients (of Spearman, Kendall), or correlation coefficients based on chi-square test (Chuprov T-coefficient, Cramer's contingency coeff., Yule's coeff. of association …). However, application of these measures (of Spearman, Kendall) requires the values of the variable Y to be ranged and each chosen order is very arguable. On the other hand, setting ranks to values of potential eplanatory variables, which are continous, would be meaningful when the space of their possible values would be constrained. It could be done by laying down variants' values corresponding to certain brackets of these variables' values. Arrangement of number and borders of these brackets is very arguable, too. Correlation coefficients based on chi-square test also require an arrangement of number and scopes of intervals of continous variables' values. So, different results are obtained for different arrangments. Finally, calculations of these different coefficients may only provide an intuition about relationship between the variable Y and the others. It is impossible to fix objectively, which variables are stronger and which weaker correlated.

Without minimizing the rank of the standard approach, it appears that the development of presented models is at least competitive. For example, for the year 2005 the average values of the variable W14 may be calculated and the variant 1 of the variable SP4 becomes equal to 90.59, for the variant 2: 74.69 and for the variant 3: 51.5. From the analysis of variable W14 in the model of the year 2005 it can also be concluded that the average values of this variable for variants 1 and 2 are significantly higher than for variant 3.

First of all the approach selected by authors of the present paper shows how forms of ownership should be grouped into a smaller number of variants. In

addition, a comprehensive approach is imbedded in the model: the set of variables found in it has a significant link to the variable that corresponds to the form of ownership. Hence the conclusion which financial-economic characteristics mostly coincide with the variant of ownership form.
Recapitulating, despite of considerable limiting of profits flowed from qualitative variable's model's building, the approach seems to be advantageous anyway. First of all it allows to look at the examined populations comprehensively and it is competitive in relation to the standard approach.

LITERATURE

Cramer J.S. (2001), An Introduction to the Logit Model for Economists, TCL, London
Cramer J.S. (2003), Logit Models from Economics and Other Fields, Cambridge Univ. Press, Cambridge
Dudek H., Dybciak M. (2006), Application of the Logit Model to Analyze the Test Results, Zesz. Nauk. Ek. i Org.Gosp. Żywn., No. 60, Warsaw, p. 81-92 (in Polish)
Encyclopedia of Statistical Sciences (2008), J.Wiley&Sons
Franc-Dąbrowska J. (2008), Separation of Property from Management and Efficiency of the Management of Selected Factors of Production, Roczniki Nauk Rolniczych, Serie G, T.95, z.3 / 4, p. 70 – 82 (in Polish)
Grontkowska A. (2007), Changes in Assets and Results of ANR's Partnership Enterprises in the Years 1996 - 2005, R.N.R., S. G, T.94, z.1, p. 66 – 73 (in Polish)
Gruszczyński M. (2002), Models and Forecasts of Qualitative Variables in Finance and Banking, OWSGH, Warsaw (in Polish)
Guzewicz W., Kulawik J., Osuch D. (2006), How the Ranking has Been Made. Methodological Notes, N. Ż. Gosp., 22/430 (App. "Lista 300"), p. 24-25 (in Polish)
Leopold A., Ziętara W. (1999), The Fifth List, Nowe Życie Gospodarcze, 39/227, p. 19 – 25 (in Polish)
Long S. (1997), Regression Models for Categorical and Limited Dependent Variables, Sage Publ., Thousand Oaks, London
Manski C.F., McFadden D. (eds.) (1977), Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge, Mass.
Mikhail M.B., Walther B.R., Willis R.H. (1998), Does forecast accuracy matter to analysts, Social Science Research Network Electronic Library, www.ssrn.com
Seo S.N., Mendelsohn R. (2008), An analysis of crop choice: Adapting to climate change in South American farms, Ecological Economics, 67, pp.109-116.
Theil H. (1969), A Multinomial Extension of the Linear Logit Model, International Economic Review, 10, pp. 251-259.

Przydatność wielomianowych modeli logitowych do badania związku formy własności przedsiębiorstw rolniczych z wynikami finansowo-ekonomicznymi

Streszczenie: W artykule podjęty został problem pewnego ograniczonego wykorzystania wielomianowych modeli logitowych do zbadania związku formy własności przedsiębiorstw rolniczych z pewnymi ich charakterystykami finansowo - ekonomicznymi, opisującymi efektywność gospodarowania. Do badań wykorzystano dane z tzw. „list 300" publikowanych przez "Nowe Życie Gospodarcze". W kolejnych, traktowanych odrębnie latach wybierano po trzysta najlepszych przedsiębiorstw rolnych spośród tych, które chciały wziąć udział w rankingu. Twórcy list brali pod uwagę przedsiębiorstwa, które powstały z majątku Skarbu Państwa w latach dziewięćdziesiątych XX w

Słowa kluczowe: przedsiębiorstwa rolnicze, formy własności, wskaźniki efektywności gospodarowania, wielomianowe modele logitowe