ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 164, 2003

Piotr Błażczak, Feliks Wysocki***

APPLICATION OF SMALL AREA STATISTICS IN AGRICULTURE INVESTIGATIONS

Abstract. The paper presents investigative proposition for assessment of usefulness of SAS regressive methods to the analysis of economic phenomena in agriculture. Objects of investigation are administrative districts of Wielkopolska province. To statistical analysis one subjected separately two pair of features, in accordance with their reason-result connection. Results, obtained for the investigated, randomly selected sample of objects, are compared with the known values of indices, calculated for the whole population. Methodical aspects, leaning on regressive problems, are illustrated on statistical data originating from last farming register in Poland.

Key words: small area statistics, economics of agriculture, districts of Wielkopolska Province.

1. INTRODUCTION

Statistical investigations in Poland have been carried out for many decades, similarly as in many other countries. Their great significance, among other, is connected with obtaining appropriate assessments, describing the state of agriculture in respect of, for example, the level of crops, agrarian structure, farming effectiveness etc. Even in the nineties, the majority of statistical studies in Poland was based on reporting consisting in full investigation of statistic units. Presently, from the point of view of dynamic economic systems, and resulting from the need of possibly quick analyses, public reporting directs towards such technical investigations which will provide correct general conclusions about the whole economy, using for this purpose small by number sampling populations, briefly called samples (e.g. K o r d o s (1991), B r a c h a (1996. p. 249)).

* Ph. Doctor, Department of Mathematical and Statistical Methods, August Cieszkowski Agricultural University of Poznań.

** Professor, Department of Agricultural Economics and Management, August Cieszkowski Agricultural University of Poznań.

New territorial division of Poland, introduced on the 1st of January 1999, and resulting from this greater competences for self-goverment organizations of smaller administrative units (administrative districts and communes) increase the demand for current assessment of various indices corresponding to small areas. Public reporting in the scope of regional statistics perceived such need much earlier, looking for appropriate methods of statistical analysis (Szwałek and Zaremba (1992)).

In the eighties and nineties, the world public statistics directed its attention towards the utilization of investigation methods for regional needs, based on small area statistics (SAS). Kordos (1996) in his paper referred to the justification of this problem, examining the problem of regional investigations against a background of many aspects of the public statistics work.

SAS is understood as statistical method and statistical data collected, worked out and published for such administrative units as: villages, communes, settlements, towns and administrative districts. (Kordos (1992)). Statistical methods used in this scope are based on maximum utilization of information from small in number samples. For this reason methods of representative statistics are observed very rigorously, including the rules of units sampling (Bracha (1996)). There are utilized new techniques of estimation, typical only for SAS. Some of these methods utilize additional information in the form of auxiliary characteristic, values of which are known in the whole population. A very interesting review of these methods, against the background of regional needs, was given by Dehnel (1999). Similarly, Rao (1999) presented a review of estimation techniques and application of SAS during the last five years.

The problems connected with SAS may be a very useful tool for statistical description of agricultural problems. Actions of Chief Census Bureau (CCB) aim at adaptation of standard elaborated by EUROSTAT and mutual unification of technique and scopes of statistical investigations. In the Polish bibliography, there are not any works, presenting SAS investigation propositions for assessment of economic and agricultural indices. Some attempts in this scope were made by Błażczak *et al.* (1999).

This paper presents investigative proposition for assessment of usefulness of SAS regressive methods to the analysis of economic phenomena in agriculture. Results, obtained for the investigated, randomly selected sample, are compared with the known values of indices, calculated for the whole population. Consistence (inconsistence) of results was assessed by means of measures for problems of predictions accuracy investigation *ex post*.

The investigated population of 226 units are communes of Wielkopolska Province. Administrative districts, constituting 35 separable units of inference including communes, are the investigated domains, also called small areas (Table 1).

Table 1

Characteristics of administrative districts on account of division into communes

Number communes in the district	1	3	4	5	6	7	8	9	10	11	14	17
Number of districts	4	2	3	6	3	5	5	2	1	2	1	1

Four districts, constituting at the same time one commune, are town districts: Kalisz, Konin, Leszno and Poznań. Among the land districts, those with the greatest number of communes are respectively: Poznań District (17) and Konin District (14).

As the values of investigated main and auxiliary characteristics for communes were determined on the basis of 1996 farm census, published by Chief Census Bureau, thus the census data has been transformed into the actually valid territorial and administrative system of communes and districts.

2. METHODS OF INVESTIGATIONS

Small area statistics utilizes various techniques, enabling the estimation of unknown economic indices. Some of them (direct estimation of POS type) apply only to small areas, from which observations were included in the sample, whereas other techniques (synthetic, complex estimation) apply to all small areas. In the second case it applies to the method of regressive estimation, which uses beside the values of investigated characteristic also values of auxiliary characteristic known in relation to all units of population.

This paper considers both techniques of estimation of the expected investigated characteristic value in the small areas (districts). For this purpose use is made of forms of estimators given in the paper of Choudhry and Rao (1992) and denoted below by numbers (1) to (6):

– The simple expansion estimator

$$\hat{Y}_i(\exp) = \begin{cases} \frac{\tilde{N}}{\tilde{n}} \sum_{j \in s_i} y_{ij} & \text{when} \quad n_i \geqslant 1 \\ 0 & \text{when} \quad n_i = 0 \end{cases} \tag{1}$$

– The post-stratified estimator (POS type)

$$\hat{Y}_i(\text{pst}) = \begin{cases} \frac{N_i}{n_i} \sum_{j \in s_i} y_{ij} = N_i \bar{y}_i & \text{when} \quad n_i \geqslant 1 \\ 0 & \text{when} \quad n_i = 0 \end{cases} \tag{2}$$

– The ratio synthetic estimator

$$\hat{Y}_i(\text{syn}) = (\bar{y}/\bar{x})X_i \quad (3)$$

– The best linear unbiased predictor

$$\hat{Y}_i(\text{blup}^*) = n_i\bar{y}_i + (\bar{y}/\bar{x})X_i^* \quad \text{where} \quad X_i^* = X_i - n_i\bar{x}_i \quad (4)$$

– The sample size dependent estimator

$$\hat{Y}_i(\text{ssd}) = a_i\hat{Y}_i(\text{pst}) + (1 - a_i)\,\hat{Y}_i(\text{syn}) \quad (5)$$

$$\text{where} \quad a_i = \begin{cases} 1 & \text{when} \quad x_i \geqslant W_i, \\ (1/\delta)(w_i/W_i) & \text{when} \quad w_i < W_i, \end{cases}$$

$$\text{while } w_i = n_i/\tilde{n} \quad \text{and} \quad W_i = N_i/\tilde{N}.$$

– The different sample size dependent estimator

$$\hat{Y}_i(\text{ssd}^*) = a_i^*\hat{Y}_i(\text{reg}) + (1 - a_i^*)\,\hat{Y}_i(\text{syn}) \quad (6)$$

$$\text{where} \quad a_i^* = \begin{cases} 1 & \text{when} \quad x_i \geqslant W_i, \\ (w_i/W_i)^{h-1} & \text{when} \quad w_i < W_i, \end{cases}$$

$$\text{while } \bar{X}_i = X_i/N_i \quad \text{and} \quad \hat{Y}_i(\text{reg}) = N_i[\bar{y}_i + \frac{\bar{y}}{\bar{x}}(\bar{X}_i - \bar{x}_i)].$$

Individual denotations in the given formulae have the following meaning: i – index of i-th – small area, j – index of j-th – observation in i-th – small area, Y, (X) – symbol of investigated main (auxiliary) variable, $Y_i(\bar{Y}_i)$, $X_i(\bar{X}_i)$ – sums (mean) of characteristics value in i-th – small area, $\tilde{n}$ – number of sample units, $\tilde{N}$ – number of population units, n_i – number of sample units from i-th – small area, N_i – number of population units belonging to i-th – small area, $\bar{y}$, $\bar{x}$ – arithmetic means of variables from all sample units, $\bar{y}_i$, $\bar{x}_i$ – arithmetic means of variables from sample units of i-th – small area, δ, h – parameters arbitrarily selected, influencing the size of synthetic estimator participation in general value of estimator.

Formulae (1) to (6) relate to description of small areas (districts) in situation, when single-stage sampling scheme for sample is used. In compliance with C h o u d h r y and R a o (1992) suggestion we accept the generally used values of parameters $\delta = 1$, $h = 2$. S i n g h and other (1992) specify similar formulae for estimators, however they are extending considerations on the units of investigation with some importances assigned to them. In the scope of this paper this trend of considerations was not followed. Formulae (1) and (2) make use only of information concerning the main characteristic Y, the remaining formulae take into account also the auxiliary characteristic X. For these reasons formulae (1) and (2) allow for estimation of main characteristic only for these small areas, from which the commune units were randomly selected for sample.

Random selection of statistical units (communes) numbers for sample was carried out with the use of randomising function LOS ( ) in the programme of EXCEL 97 calculation sheet. It generates random values from uniform distribution. In this work, there was accepted the rule of sampling of 25 statistical units, what constitutes about 11% of units in general population. This is in agreement with the general tendency of samples random selection (minimum 10% units) for assessment of unknown parameters used in practice by the public statistics.

On the basis of randomly selected sample, values of estimators (1) to (6) for small areas have been determined. Next, for the values of estimator, defined by the formula (6), calculated for all small areas, values of measures for degree of prediction accuracy *ex post* were determined, expressed in formulae (7) to (12) [C i e ś l a k (1997)]:

– Mean square error of predictions *ex post*,

$$s = \sqrt{\sum_i (\hat{Y}_i(\text{ssd}^*) - Y_i')^2/(a-1)}, \tag{7}$$

– Coefficient of variation

$$w = 100^* s/\bar{y}, \tag{8}$$

– Theil's coefficient of divergence

$$I^2 = \sum_i (\hat{Y}_i(\text{ssd}^*) - Y_i')/\sum_i Y_i'^2, \tag{9}$$

– Mean error of prediction

$$I = \sqrt{I^2}, \tag{10}$$

– Coefficient of negative influence on prediction

$$I_1^2 = (\hat{\bar{Y}}(\text{ssd}^*) - \ddot{y})^2/((\sum_i Y_i'^2)/(a-1)), \tag{11}$$

– Relative index of negative influence on prediction

$$\tilde{I}_1^2 = 100^* I_1^2/I^2. \tag{12}$$

where: a – number of small areas, $Y'(Y_i')$ – general sum (of i-th – area) of main characteristic determined on the basis of all population units, $\ddot{y}$ – mean value of main characteristic calculated for all Y_i' small areas, $\hat{\bar{Y}}$ (ssd*)– mean value of main characteristic calculated for all $\hat{Y}_i$(ssd*) small areas.

Values (7) to (12) lead to conclusions, concerning the suitability of SAS estimators for the investigation of economic and agricultural phenomena.

3. RESULTS OF EMPIRICAL INVESTIGATIONS

Results of investigations will be illustrated by an example of assessment of agricultural production goods value (APG), calculated for one person fully employed in the farmsteads, for one randomly generated sample of 25 communal units. During investigation 500 such samples were randomly selected for presenting general conclusions.

Values of APG (in thousand PLN) and number of fully employed in the farmsteads (NFE) were assessed separately and treated in each case as the main characteristics. To each of them there was assigned one auxiliary variable, proper for the number of farmsteads considered by their users as evolutionary (NEF) and for the number of households engaged in on agricultural activity, the income of which from agricultural activity in relation to the general income of the household lies within the range $< 90 - 100\% >$ (NH).

Values of these characteristics for randomly selected units of the sample are given in Table 2.

Table 2

Values of main and auxiliary characteristics for randomly selected 25-element sample of communes in the order of population units numbers

Number of commune	The commune	APG	NFE	NEF	NH
3	Szamocin	4 125	514	53	54
6	Czarnków	887	67	8	4
9	Wieleń	6 857	941	104	119
26	Krobia	28 054	2 041	375	409
45	Koźminek	12 844	2 249	208	309
53	Bralin	6 847	733	95	136
62	Babiak	9 102	1 711	199	400
66	Kościelec	5 860	1 243	105	131
74	Grodziec	5 819	1 610	114	227
75	Kazimierz Biskupi	3 283	1 029	62	95
92	Zduny	9 264	703	107	151
100	Włoszakowice	8 634	1 266	105	141
103	Chrzypsko Wielkie	7 395	603	78	92
112	Rogoźno	13 244	802	126	163
122	Grabów nad Prosną	12 546	1 726	211	308
139	Chocz	4 525	875	68	91
148	Kórnik	16 375	1 334	172	251
166	Słupca	1 474	203	26	32
181	Obrzycko	5 135	511	58	96
183	Dominowo	10 031	616	98	142
194	Brudzew	5 215	1 665	103	215
202	Skoki	7 758	650	96	126
206	Wągrowiec	24 819	1 826	357	390
215	Złotów	1 187	66	7	4
225	Leszno	5 605	277	27	43

Source: Own elaboration.

Subjecting values of four investigated characteristics from the sample to initial statistic analysis values of basic characteristics and coefficients of linear correlation were obtained, which are given in Table 3.

Table 3

The basic descriptive characteristics of the randomly selected sample units

Descriptive characteristics		APG	NFE	NEF	NH
Kurtosis		2.93	–0.90	2.47	–0.24
Skewness		1.64	0.32	1.54	0.79
Bottom quartile		5 135	603	62	92
Maximum		28 054	2 249	375	409
Median		6 857	875	103	136
Mean		8 675	1 010	118	165
Minimum		887	66	7	4
Top quartile		10 031	1 610	126	227
Standard deviation		6 601	627	93	120
Coefficient of variation		76	62	79	73
Coefficient of linear correlation	APG		0.6646	0.9434	0.8149
	NFE			0.8226	0.9053
	NEF				0.9331

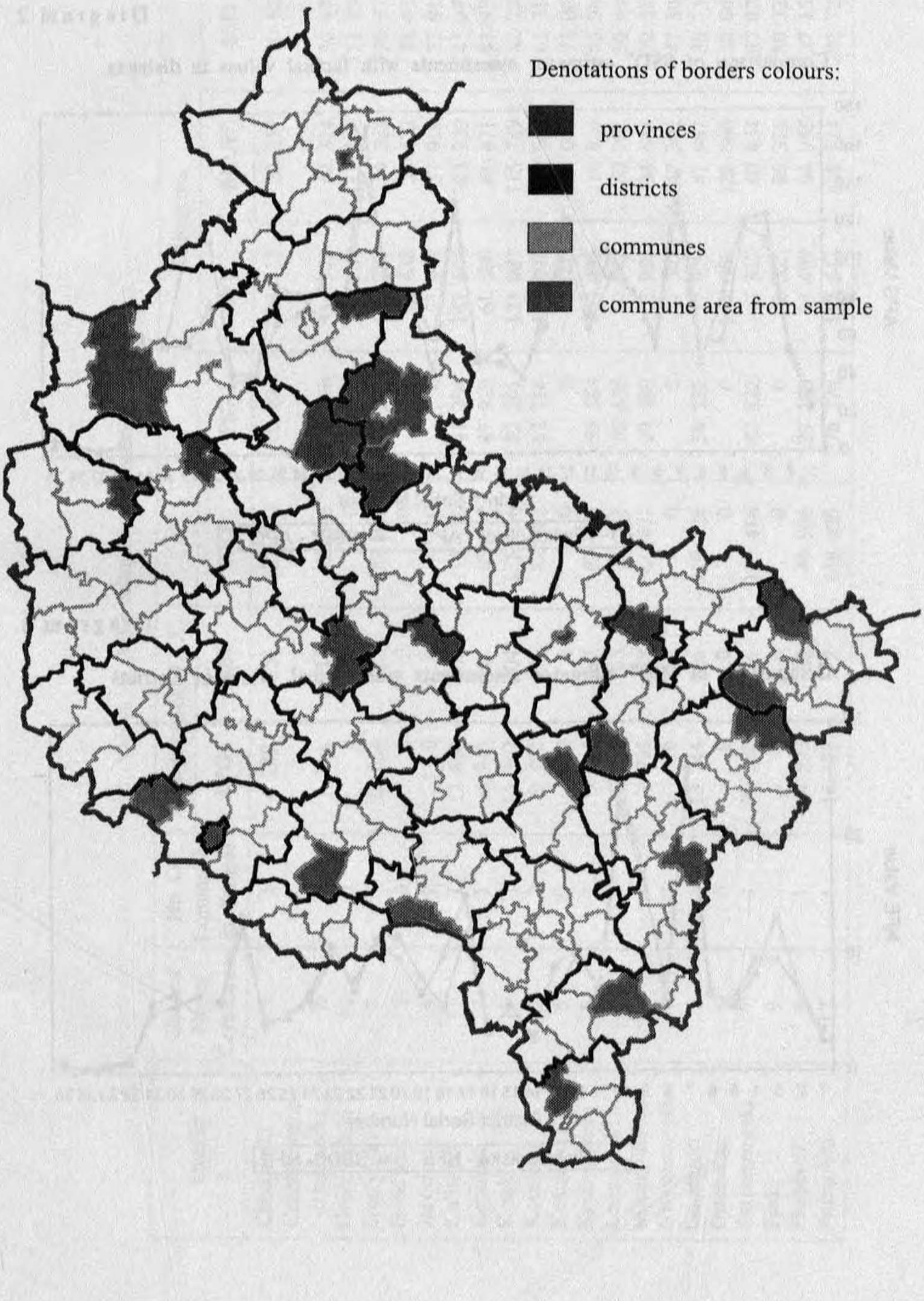
Source: Own elaboration.

Analysing the results obtained in Table 3 we state, that APG and NEF are characterized by substantial positive kurtosis and this testifies that these characteristics distribution is slenderer than normal distribution. Values of remaining characteristics show slight flattening of their distributions. It must be observed, that all considered characteristics show very great variation. It is indicated by the values of coefficient of variation, which exceeds 60% and 70%.

Spatial arrangement of sampled units with respect to population division into small areas (districts) is shown on Diagram 1. In addition, it is observed that the considered groups of main and auxiliary characteristics are closely correlated between themselves. The same conclusions to the above provides analysis of all population units. They confirm the justness of selection of characteristics pairs (Y – APG, X – NEF), (Y – NFE, X – NH) for regressive estimation of main characteristics (Y), utilizing values of auxiliary variables (X).

Diagram 1

Spatial distribution of sample units



Diagrams 2 and 3 show values of regressive estimators (SSD*) for two main characteristics from the sample and corresponding to them factual values from the investigated population.

Diagram 2

Comparison of SSD* estimator assessments with factual values in districts

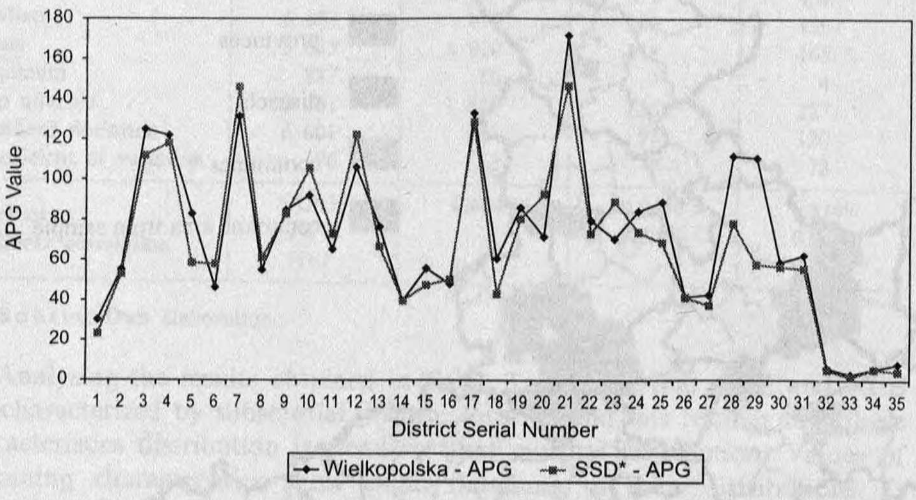


Diagram 3

Comparison of SSD* estimator assessments with factual values in districts

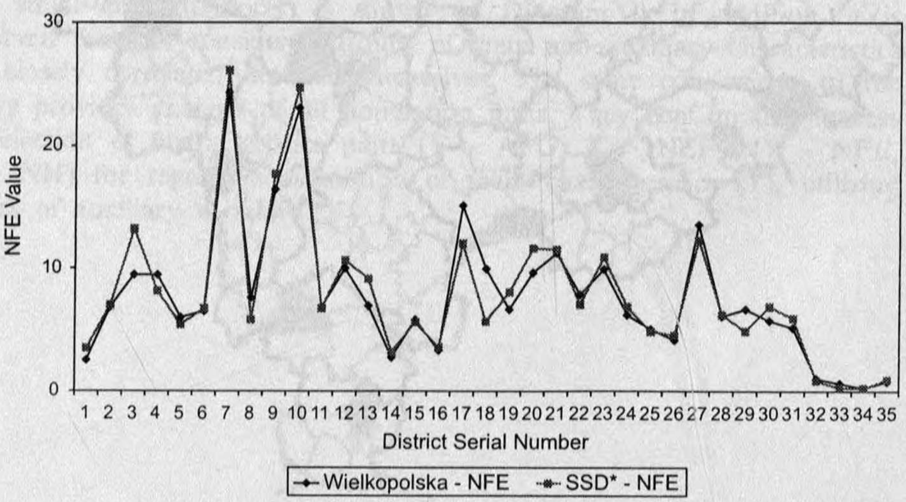


Table 4

Values of APG estimators assessed for districts

District	General No of Communes	No Of Communes in Sample	Characteristic Y, X		Estimators					
			APG	NEF	EXP	PST	SYN	BLUP*	SSD	SSD*
Chodzieski	5	1	4 125	53	37 290	20 625	22 113	22 357	20 625	23 334
Czarnkowsko--trzcianecki	8	2	7 744	112	70 006	30 976	54 770	54 314	30 976	52 943
Gnieźnieński	10	0	0	0	0	0	111 884	111 884	111 884	111 884
Gostyński	7	1	28 054	375	253 608	196 378	113 788	114 383	196 378	117 957
Grodziski	5	0	0	0	0	0	58 432	58 432	58 432	58 432
Jarociński	4	0	0	0	0	0	57 919	57 919	57 919	57 919
Kaliski	11	1	12 844	208	116 110	141 284	167 606	165 220	145 974	146 034
Kępiński	7	1	6 847	95	61 897	47 929	61 580	61 471	47 929	60 816
Kolski	11	2	14 962	304	135 256	82 291	123 087	115 789	82 291	82 950
Koniński	14	2	9 102	176	82 282	63 714	132 093	128 308	63 714	105 597
Kościański	5	0	0	0	0	0	73 003	73 003	73 003	73 003
Krotoszyński	6	1	9 264	107	83 747	55 584	113 495	114 924	55 584	122 070
Leszczyński	7	1	8 634	105	78 051	60 438	60 042	60 988	60 438	66 662
Międzychodzki	4	1	7 395	78	66 851	29 580	32 950	34 634	29 580	39 685
Nowotomyski	6	0	0	0	0	0	47 302	47 302	47 302	47 302
Obornicki	3	1	13 244	126	119 726	39 732	37 929	41 947	39 732	49 983
Ostrowski	8	0	0	0	0	0	128 066	128 066	128 066	128 066
Ostrzeszowski	7	1	12 546	211	113 416	87 822	63 337	60 434	87 822	43 010
Pilski	9	0	0	0	0	0	80 325	80 325	80 325	80 325
Pleszewski	6	1	4 525	68	40 906	27 150	95 409	94 955	27 150	92 684
Poznański	17	1	16 375	172	148 030	278 375	112 543	116 324	200 727	146 721

Table 4 (contd.)

District	General No of Communes	No Of Communes in Sample	Characteristic Y, X		Estimators					
			APG	NEF	EXP	PST	SYN	BLUP*	SSD	SSD*
Rawicki	5	0	0	0	0	0	73 149	73 149	73 149	73 149
Słupecki	8	1	1 474	26	13 325	11 792	92 407	91 977	11 792	88 968
Szamotulski	8	1	5 135	58	46 420	41 080	66 559	67 447	41 080	73 664
Średzki	5	1	10 031	98	90 680	50 155	54 331	57 186	50 155	68 607
Śremski	4	0	0	0	0	0	41 297	41 297	41 297	41 297
Turecki	9	1	5 215	103	47 144	46 935	58 358	56 031	46 935	37 416
Wągrowiecki	7	2	32 577	453	294 496	114 020	80 325	79 732	114 020	78 250
Wolsztyński	3	0	0	0	0	0	57 919	57 919	57 919	57 919
Wrzesiński	5	0	0	0	0	0	56 821	56 821	56 821	56 821
Złotowski	8	1	1 187	7	10 730	9 496	50 450	51 125	9 496	55 846
Kalisz	1	0	0	0	0	0	5 052	5 052	5 052	5 052
Konin	1	0	0	0	0	0	1 611	1 611	1 611	1 611
Leszno	1	1	5 605	27	50 669	5 605	1 977	5 605	5 605	5 605
Poznań	1	0	0	0	0	0	4 320	4 320	4 320	4 320

Table 5

Values of NFE estimators assessed for districts

District	General No of Communes	No Of Communes in Sample	Characteristic Y, X		Estimators					
			NFE	NH	EXP	PST	SYN	BLUP*	SSD	SSD*
Chodzieski	5	1	514	54	4 647	2 570	2 570	2 753	2 570	3 488
Czarnkowsko--trzcianecki	8	2	1 008	123	9 112	4 032	5 953	6 208	4 032	6 975
Gnieźnieński	10	0	0	0	0	0	13 221	13 221	13 221	13 221
Gostyński	7	1	2 041	409	18 451	14 287	11 385	10 924	14 287	8 157
Grodziski	5	0	0	0	0	0	5 408	5 408	5 408	5 408
Jarociński	4	0	0	0	0	0	6 711	6 711	6 711	6 711
Kaliski	11	1	2 249	309	20 331	24 739	22 869	23 227	24 406	26 110
Kępiński	7	1	733	136	6 626	5 131	6 491	6 392	5 131	5 798
Kolski	11	2	2 954	531	26 704	16 247	19 327	19 032	16 247	17 706
Koniński	14	2	2 639	322	23 857	18 473	20 000	20 669	18 473	24 683
Kościański	5	0	0	0	0	0	6 742	6 742	6 742	6 742
Krotoszyński	6	1	703	151	6 355	4 218	11 973	11 752	4 218	10 648
Leszczyński	7	1	1 266	141	11 445	8 862	6 326	6 729	8 862	9 150
Międzychodzki	4	1	603	92	5 451	2 412	2 992	3 032	2 412	3 152
Nowotomyski	6	0	0	0	0	0	5 598	5 598	5 598	5 598
Obornicki	3	1	802	163	7 250	2 406	4 130	3 934	2 406	3 544
Ostrowski	8	0	0	0	0	0	12 034	12 034	12 034	12 034
Ostrzeszowski	7	1	1 726	308	15 603	12 082	6 742	6 584	12 082	5 634
Pilski	9	0	0	0	0	0	8 051	8 051	8 051	8 051
Pleszewski	6	1	875	91	7 910	5 250	9 709	10 027	5 250	11 619
Poznański	17	1	1 334	251	12 059	22 678	13 337	13 136	18 304	11 515

Table 5 (contd.)

District	General No of Communes	No Of Communes in Sample	Characteristic Y, X		Estimators					
			NFE	NH	EXP	PST	SYN	BLUP*	SSD	SSD*
Rawicki	5	0	0	0	0	0	7 048	7 048	7 048	7 048
Słupecki	8	1	203	32	1 835	1 624	10 829	10 836	1 624	10 887
Szamotulski	8	1	511	96	4 619	4 088	7 427	7 351	4 088	6 817
Średzki	5	1	616	142	5 569	3 080	6 026	5 773	3 080	4 762
Śremski	4	0	0	0	0	0	4 448	4 448	4 448	4 448
Turecki	9	1	1 665	215	15 052	14 985	9 085	9 435	14 985	12 232
Wągrowiecki	7	2	2 476	516	22 383	8 666	8 492	7 811	8 666	6 109
Wolsztyński	3	0	0	0	0	0	4 809	4 809	4 809	4 809
Wrzesiński	5	0	0	0	0	0	6 797	6 797	6 797	6 797
Złotowski	8	1	66	4	597	528	5 567	5 609	528	5 900
Kalisz	1	0	0	0	0	0	875	875	875	875
Konin	1	0	0	0	0	0	294	294	294	294
Leszno	1	1	277	43	2 504	277	263	277	277	277
Poznań	1	0	0	0	0	0	948	948	948	948

Tables 4 and 5 show results of assessment of investigated main characteristics unknown values in the small areas respectively for APG and NFE characteristics. Values of expansive and POS type estimators have been calculated only in relation to the small areas, from which the communal units were randomly selected for sample. Therefore they have a limited range of application. However, they are simple direct estimations in investigation of main characteristics. Values of these estimators, in relation to the results known for the whole population of communes, indicate that for Gostyń, Poznań and Wągrowiec districts there took place a significant reassessment of both main characteristics values.

The results of assessment of regressive estimators values for small areas (districts) for both main characteristics, on the basis of sample and their actual values, show some agreement of results (Diagram 2). Similarly, as with reference to the EXP and POS estimators, significant divergences in APG may be observed only in case of Poznań, Wągrowiec and Wolsztyn districts. In spite of the stated substantial changeability of investigated main and auxiliary characteristics (Table 3), regressive estimations of SSD* seem to be satisfactory for all small areas. This fact is confirmed by measure values of degree of prediction accuracy. Table 6 shows values of measures for both main characteristics and their quotient APG/NFE, calculated in accordance with given formulae (7) to (12), describing soundness of prediction.

Table 6

Values of measures for degree of predictions accuracy for selected characteristics

Name of measure	APG	NFE	APG/NFE
Mean square error of predictions ex post	20 679.5975	1 921.9288	3 986.6258
Coefficient of variation	28.8246	24.6485	37.6681
Theil's coefficient of divergence	0.0624	0.0398	0.1186
Error of prediction	0.2499	0.1994	0.3444
Coefficient of negative influence on predict	0.0010	0.0041	0.0014
Relative negative influence on prediction	1.5391	10.3271	1.1861

Source: Own calculations.

Obtained values of measure $\hat{I}\frac{2}{1}$ being near zero testify, that the assessment of the analized indices on the basis of 25-element sample is not negatively influenced.

The presented results obtained for 25-element districts sample randomly selected from the 226-element population illustrate only some of various aspects of problems connected with the assessment of the unknown values of main characteristics. Random selection of 25-element samples was repeated 500 times by the authors. Additionally, they took into account modification of estimators by means of presented formulae (3) and (6), substituting in them assessment of regression $(\bar{y}/\bar{x})$ with coefficient of regression, calculated from the sample observation. This led to obtaining another form of SSD* estimator denoted by SSD**.

Values for each of these functions, calculated separately for APG and NFE, were used for formation of a value of index type characteristic APG/NFE. Afterwards, the coefficients of linear correlation were calculated for the values of index type characteristic APG/NFE for SSD* and SSD** and real values from the whole investigated population. Values of coefficients of linear correlation for SSD** in 325 random selections were higher than for SSD*. It means, that the use of coefficient of regression calculated from sample observation does not testify univocally correction of the estimated values of main characteristic. It is influenced by the observed great changeability of characteristics values (Table 3).

During every random selection, the number of districts was also noted, from which communes were selected for sample and various multiplicities of selected communes from districts. The results of multiplicities of randomly selected communes are given in Table 7.

Table 7

Results of 1000 simulations with 25 sampled elements

No of communes randomly selected from district for sample	0	1	2	3	4	5	6
Participation of districts with a given number of communes in the sample (%)	50.02	33.48	12.57	3.05	0.77	0.10	0.02

Source: Own calculations.

Simultaneously, a relative size of assessed main characteristics values deviation in relation to real values was registered for districts. Set of these results for four ranges of deviations is given in Table 8.

Table 8

Relative differences between assessed APG and NFE values and real values (%)

Range of relative differences (in %)	APG	NFE
<0 – 10>	39.39	32.66
(10 – 25>	39.37	32.26
(25 – 50>	17.80	26.29
Above 50	3.44	8.80

Source: Own calculations.

4. CONCLUSIONS

Preliminary results of investigation with application of SAS estimators in problems connected with agriculture show their usefulness. They lead to the following general conclusions:

– It is worth noticing, that investigation of small in number samples enables assessment of characteristics values also in these small areas, in which none of units was randomly selected.

– Degree of sample results adjustment to reality depends particularly on the levels of changeability of investigated main and auxiliary characteristics.

– Increase of sample size does not influence basically the accuracy of investigated indices assessment.

– Due to the nature of many agricultural problems, assessment of usefulness of SAS estimators for determination of economic indices should be directed on simultaneous consideration of many auxiliary variables.

REFERENCES

Błażczak P., Rozpiątkowski A., Siatkowski I. (1999), *Use of Small Area Estimators in Agricultural Problems*, [in:] *Small Area Estimation*, Conference Proceedings, IASS, 209–213.

Bracha Cz. (1996), *Teoretyczne podstawy statystyki reprezentacyjnej*, PWN, Warszawa, 248–267.

Choudhry G. H., Rao J. N. K. (1992), *Estimation of Small Area Estimators an Empirical Study*, [in:] *Small Area Statistics and Survey Designs*, v. I. Invited papers, GUS and PTS, Warszawa, 271–281.

Cieślak M. (1997), *Prognozowanie gospodarcze. Metody i zastosowania*, PWN, Warszawa.

Dehnel G. (1999), *Statystyka małych obszarów jako narzędzie oceny rozwoju ekonomicznego regionów*, AE Poznań, praca doktorska.

Kordos J. (1991), *Statystyka małych obszarów a badania reprezentacyjne*, „Wiadomości Statystyczne" **4**, 1–5.

Kordos J. (1992), *Podejście do statystyki małych obszarów w Polsce*, „Wiadomości Statystyczne" **10**, 1–5.

Kordos J. (1996), *Efektywne wykorzystanie statystyki małych obszarów*, „Wiadomości Statystyczne" **9**, 11–19.

Mateńko K., Kursa L., Wanke H. (1995), *Statystyka wsi, rolnictwa i gospodarki żywnościowej*, „Wiadomości Statystyczne" **4**, 16–20.

Rao J. N. K. (1999), *Some Recent Advances in Model-based Small Area Estimation*, „Survey Methodology" **25**(2), 175–186.

Singh M. P., Gambino J., Mantel H. (1992), *Issues and Options in the Provision of Small Area Data*, [in:] *Small Area Statistics and Survey Designs*, v. I. Invited papers, GUS and PTS, Warszawa, 32–72.

Szwałek S., Zaremba H. (1992), *Źródła danych statystycznych oraz szacunków dla małych obszarów*, „Wiadomości Statystyczne" **10**, 15–19.

Piotr Błażczak, Feliks Wysocki

ZASTOSOWANIE STATYSTYKI MAŁYCH OBSZARÓW W BADANIACH ROLNICZYCH

(Streszczenie)

Niniejsza praca przedstawia propozycję badawczą zmierzającą do oceny przydatności metod regresyjnych SMO w opisie cech ekonomicznych w rolnictwie. Obiektami badania są powiaty województwa wielkopolskiego. Analizie statystycznej poddano oddzielnie dwie pary cech, zgodnie z ich powiązaniem przyczynowo-skutkowym. Uzyskane wyniki dla rozpatrywanej wylosowanej próby obiektów są porównywane ze znanymi wartościami parametrów wyliczonymi dla całej populacji. Aspekty metodyczne, oparte na zagadnieniach regresyjnych, są ilustrowane na danych statystycznych pochodzących z ostatniego spisu rolnego w Polsce.