EKONOMETRIA ECONOMETRICS 32 • 2011

ISSN 1507-3866

Aneta Sobiechowska-Ziegert
Gdańsk Technical University

APPLICATION OF TIME-SERIES-CROSS-SECTION DATA IN THE CASE OF FORECASTING SALE IN AN ENTERPRISE

Summary: The article treats possibilities of using time-series-cross-section data for sales forecasting in an enterprise. Different kind of approaches to econometric forecasting using TSCS data are presented. Short-run forecasts of gasoline sales according to regions in Poland are determined and their quality with the use of the ex-post measures is estimated. The most accurate forecasting methods are then used to create a combined forecast. In the summary. advantages and disadvantages of the particular approaches are indicated as well as the conditions of their practical use.

Key words: forecasting, sale, time-series-cross-section data.

1. Introduction

Time-series-cross-section data (TSCS data) enables to conduct more accurate analysis than time series data or simply spatial data. Besides their ability to analyze the phenomena in time and space separately, they make economic variables research possible in these dimensions at the same time. Although they seem to be very useful, there are not many papers about forecasting based on this data in Polish literature, basically due to the low availability of statistical data. The other reason is that not in all cases can this data be used jointly. TSCS data is very often called panel data and are analyzed in a manner appropriate for panel data[1].

In the analyzed case, TSCS data for 16 objects (representing population) and 31 time periods will be considered. The aim of the paper is presenting the possibilities of sale forecasting based on TSCS data. The basic purpose of presenting short-run forecasts will be determining the future sale of gasoline 95Unleaded based on statistical data provided by the LOTOS Group, according to regions. In point of fact, in the scope of modelling and forecasting of the sale, every known method can be used. The choice, however, is determined most of all by the availability of statistical

[1] Because of the nature of the paper, the differences in definitions between TSCS data and panels will be omitted.

data. Heuristic methods are useful when forecasting the sale of new products is the case, whereas statistical and econometric methods can be used for forecasting products already existing on the market, when historical data are available. The approaches to forecasting which are presented in the paper are of an econometric type[2].

2. Determinants of fuel sales

In the presented research the data concerning unleaded gasoline sales in the period of January 2006 – July 2008 according to regions have been used[3]. The forecasting period is between August 2008 and October 2008. Total gasoline sales during the analysed period is presented in Figure 1. It is clear that the sales are characterised by the growing tendency for seasonal deviations. Total gasoline sales according to the regions during the analysis period is presented in Figure 2 and indicates a large variation in space.

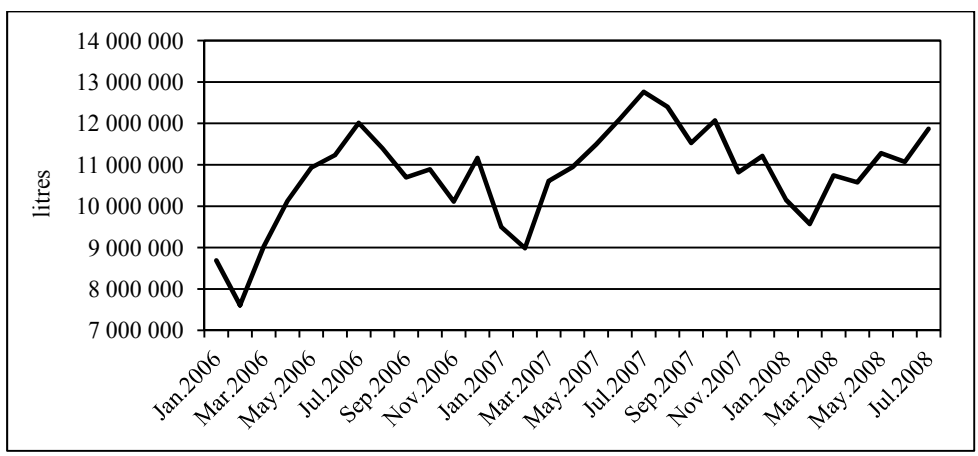

Figure 1. Unleaded gasoline sales time-series data

Source: Based on LOTOS SA data.

Based on the analysis concerning fuel markets in Poland and in the world [eg. Sobiechowska-Ziegert 2005; Dees et al. 2005], explanatory variables that show variability in time, in space or both in time and space have been chosen for econometric modelling. Table 1 presents a detailed list of variables being used in this type of research.

[2] Presented approaches to forecasting sales do not include macro-environment.

[3] Unleaded gasoline sales data concerning own stations and agency arrangements stations obtained from LOTOS Joint Stock Company (LOTOS SA).

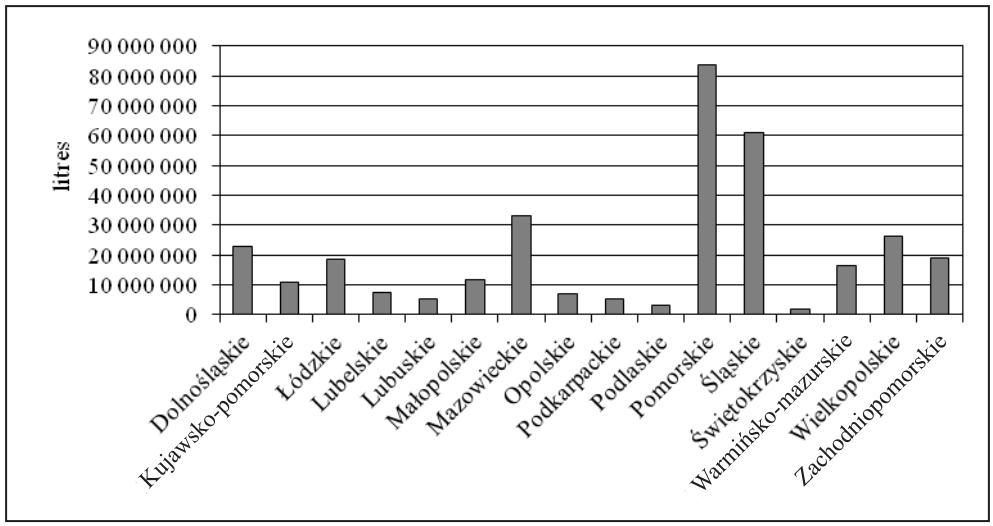

Figure 2. Unleaded gasoline sales during the analysis period; cross-section data

Source: Based on LOTOS SA data.

Table 1. Unleaded gasoline sales determinants

Sales determinants in time:	Sales determinants in space:
1. Gasoline retail price [PLN/l]	1. Gasoline retail price [PLN/l]
2. Average gross wages [PLN]	2. Average gross wages [PLN]
3. Substitution fuel retail price (LPG) [PLN/l]	3. Substitution fuel retail price (LPG) [PLN/l]
4. Seasonal dummies	4. Market power coefficient (intensity of competition)
5. Time trend	5. Dummies related to regions location

Source: own study.

The determinant that varies in time and space is gasoline retail price (see Figure 3).

Average gross wages are another factor that shows changes in time and space (see Figure 4). It can therefore be used for sale modelling according to time-series data, cross-section data or both time-series-cross-section data.

Despite the fact that LPG price (fuel substitutable for gasoline) shows variability in time and space, it will not be further considered because it appeared to be statistically insignificant.

The only factor that needs to be explained is market power coefficient (MPC). MPC factor usually takes the form of the quotient of the number of stations owned by one producer and the size of the region or the number of stations owned by one producer and the number of stations owned by the nearest competitor. Some authors also use the quotient of the total number of stations and the size of the region [Bello,

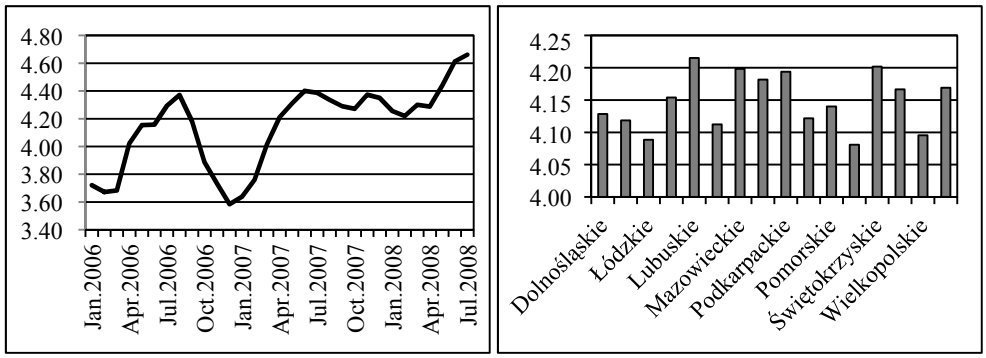

Figure 3. Average unleaded gasoline price [PLN/litre] during the analysed period

Source: Based on LOTOS SA data.

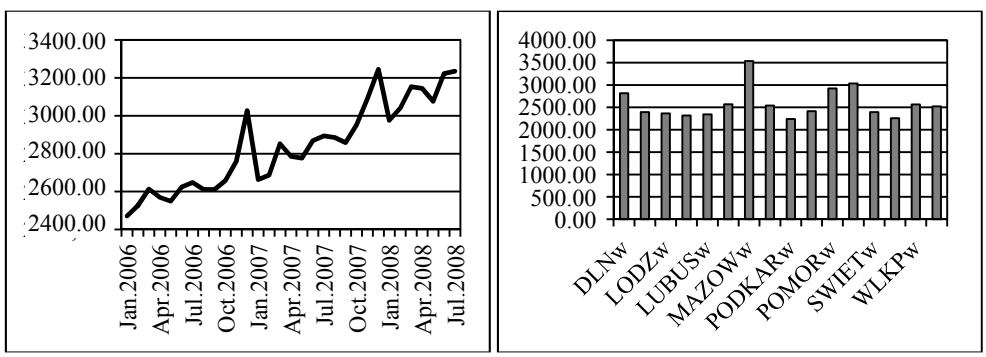

Figure 4. Average Gross wages in the corporate sector during the analysis period [PLN]

Source: Based on Central Statistical Office (GUS).

Cavero 2007]. In this case the first solution has been used and the MPC factor has been estimated according to the following formula:

$$MPC_i = \frac{GS_i}{RS_i} \tag{1.1}$$

where: MPC_i – market power coefficient of the producer in the region i,
GS_i – number of gasoline stations owned by LOTOS SA and the agent stations in the region i,
RS_i – size of the region i.

MPC in the analyzed period was characterized most of all by variability in space (see Figure 5). Variability in time appeared to be impossible to observe because of the data frequency.

In the scientific papers in the field of fuel market analysis, other explanatory variables which play the role of sale determinants are also used. These variables are related most of all to the location of the region, for example: the number of routes

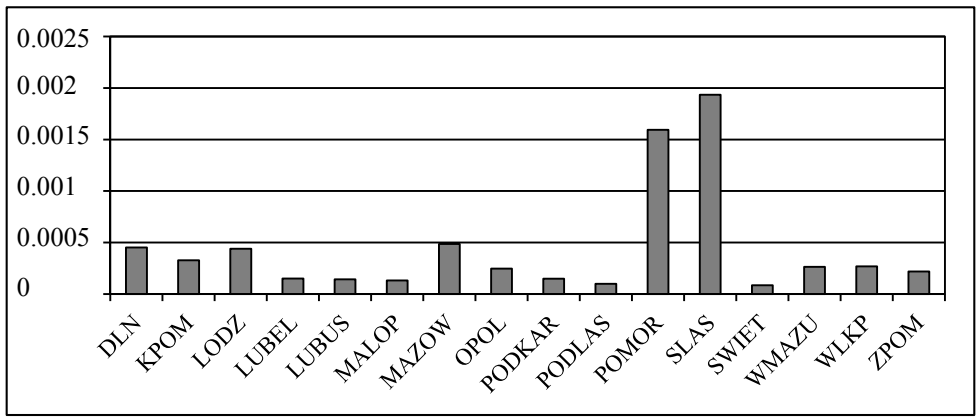

Figure 5. The average Market Power Coefficient (MPC) of the producer during the analysis period

Source: Based on LOTOS SA data.

and their purpose (expressways, highways), agricultural or industrial character of the region, proximity to other countries or regions [Banfi et al. 2005]. In the analyzed case all mentioned variables appeared to be statistically insignificant.

3. Forecasts of sales

To forecast unleaded gasoline sales every approach discussed before has been used. The forecasting period was between August and October 2008. Because of the availability of the actual values of dependent variable in the forecasting period, forecast accuracy measures have also been set. To increase accuracy, a combined forecast has been created using the methods regarded as best in this case.

In the simplest approach, using time-series data, the sale forecast has been set based on the following time-trend model[4]:

$$y_t = b_0 + b_1 t + b_2 t^2 + \sum_{j=1}^{11} c_j SR_j + u_t \tag{1.2}$$

where: b_0, b_1, b_2, c_j – structural parameters of the time-trend model,
SR_j – seasonal dummies related to the last season.

A detailed forecast for the individual regions has been obtained based on their average shares in total sales assuming that these shares will not change in the forecasting period[5].The results are presented in Table 2 and in Figure 6.

[4] The proposed parabolic time trend model described gasoline sales best. The estimation results allowed to use this model for forecasting.

[5] Assumption of maintaining the constancy of the shares in the future was made due to the short horizon of the forecast.

Table 2. Detailed forecast based on model 1.2.

1 Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error	Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error
DLN	Aug. 2008	790 377.0	–37 316.2	–5.0%	PODKAR	Aug. 2008	186 618.1	47 877.0	20.4%
	Sep. 2008	714 562.4	–13 201.0	–1.9%		Sep. 2008	168 717.4	56 152.4	25.0%
	Oct. 2008	735 980.2	10 693.3	1.4%		Oct. 2008	173 774.4	74 949.4	30.1%
KPOM	Aug. 2008	370 028.6	29 516.0	7.4%	PODLAS	Aug. 2008	112 310.5	39 438.2	26.0%
	Sep. 2008	334 534.7	64 561.8	16.2%		Sep 2008	101 537.4	31 518.9	23.7%
	Oct. 2008	344 561.8	84 712.5	19.7%		Oct. 2008	104 580.8	40 161.2	27.7%
LODZ	Aug. 2008	638 401.5	82 758.3	11.5%	POMOR	Aug 2008	2 876 256.2	336 471.9	10.5%
	Sep. 2008	577 164.8	–38 915.1	–7.2%		Sep. 2008	2 600 359.9	175 433.3	6.3%
	Oct. 2008	594 464.3	–39 330.1	–7.1%		Oct. 2008	2 678 301.2	283 359.0	9.6%
LUBEL	Aug. 2008	250 798.3	36 798.9	12.8%	SLAS	Aug.. 2008	2 091 593.7	36 839.3	1.7%
	Sep. 2008	226 741.2	45 969.9	16.9%		Sep. 2008	1 890 963.8	242 043.5	11.3%
	Oct. 2008	233 537.4	51 762.0	18.1%		Oct. 2008	1 947 642.2	385 268.3	16.5%
LUBUS	Aug. 2008	179 412.8	1 789.0	1.0%	SWIET	Aug. 2008	54 485.1	–6 808.6	–14.3%
	Sep. 2008	162 203.2	–6 096.0	–3.9%		Sep. 2008	49 258.8	–612.8	–1.3%
	Oct. 2008	167 064.9	–15 906.0	–10.5%		Oct. 2008	50 735.2	2 503.4	4.7%
MALOP	Aug. 2008	407 330.8	–56 859.0	–16.2%	WMAZU	Aug. 2008	560 029.2	–33 993.0	–6.5%
	Sep. 2008	368 258.8	–20 327.7	–5.8%		Sep. 2008	506 310.0	–9 100.7	–1.8%
	Oct. 2008	379 296.7	26 031.8	6.4%		Oct. 2008	521 485.8	–11 310.3	–2.2%
MAZOW	Aug. 2008	1 143 959.0	–26 937.8	–2.4%	WLKP	Aug. 2008	903 889.0	–30 379.7	–3.5%
	Sep. 2008	1 034 228.2	60 702.6	5.5%		Sep. 2008	817 186.1	33 341.4	3.9%
	Oct. 2008	1 065 227.3	128 894.2	10.8%		Oct. 2008	841 679.8	89 037.9	9.6%
OPOL	Aug. 2008	241 295.6	–28 323.1	–13.3%	ZPOM	Aug. 2008	656 058.7	–206 725.4	–46.0%
	Sep. 2008	218 150.1	–14 817.3	–7.3%		Sep. 2008	593 128.2	–148 940.6	–33.5%
	Oct. 2008	224 688.7	109.6	0.0%		Oct. 2008	610 906.2	–151 492.7	–33.0%

Source: own study.

Analyzing the results presented above, one can notice that the most accurate forecast has been obtained for: Dolnośląski, Warmińsko-mazurski and Wielkopolski, while the least accurate was for: Zachodniopomorski, Podkarpacki and Podlaski. The reason for it was the assumption of a constant share of sales that was varying in different directions and the strength of individual regions during the analysed period. The results achieved based on this approach could be improved by taking into account the company's strategic plans for entering new markets or omitting old ones. The forecasts obtained in this approach are on average underestimated and their average relative error is 15.8%.

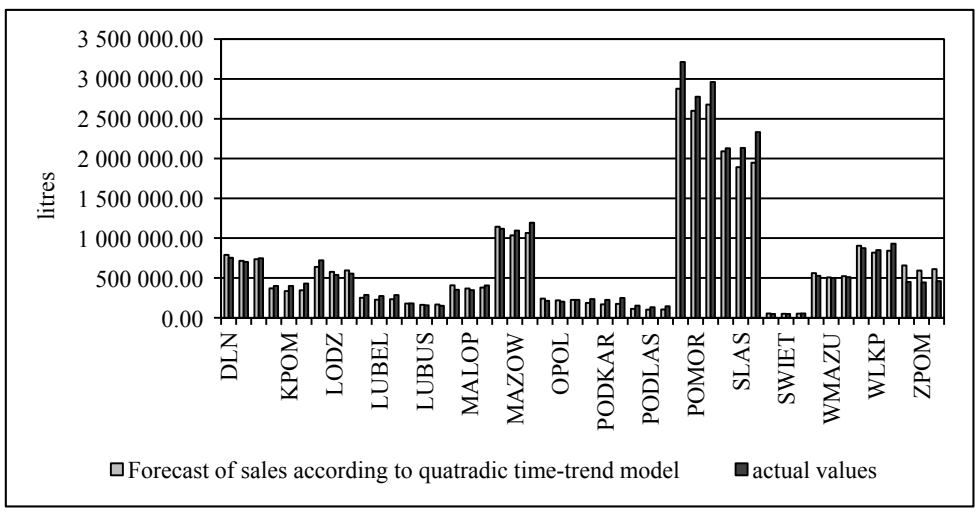

Figure 6. Forecast of unleaded gasoline sales according to model 1.2

Source: own study.

In a more laborious approach, different analytical forms and econometric models types have been used because of the heterogeneity of the regions[6]. For the purpose of forecasting, models which best reflected gasoline sales during the analysed period, have been chosen. A forecast for each region based on cause-effect models has been obtained using forecasting values of gross wages for individual provinces and producer's assumptions concerning future, desirable retail prices at gas stations according to the regions.The results have been presented in Table 3 and in Figure 7. While analyzing the results it can be observed that the above approach allowed to obtain accurate sales forecasts in most provinces. The exception is the region of Opole where the average relative ex-post error was 21.5% and the region of Lodz where the average, relative ex-post error was 17.6%. The forecast accuracy in this approach

[6] Variation coefficient for the total sale according to regions was 103.8%.

Table 3. Detailed forecasts based on individual econometric models for each region

2 Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error	Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error
DLN	Aug. 2008	737 258.9	15 801.9	2.1%	PODKAR	Aug. 2008	248 473.2	–13 978.0	–6.0%
	Sep. 2008	744 281.3	–42 919.9	–6.1%		Sep. 2008	261 271.2	–36 401.5	–16.2%
	Oct. 2008	726 444.4	20 229.1	2.7%		Oct. 2008	271 715.7	–22 992.0	–9.2%
KPOM	Aug. 2008	389 421.1	10 123.5	2.5%	PODLAS	Aug. 2008	144 682.7	7 066.0	4.7%
	Sep. 2008	383 063.3	16 033.2	4.0%		Sep. 2008	145 607.4	–12 551.0	–9.4%
	Oct. 2008	378 511.7	50 762.6	11.8%		Oct. 2008	147 180.3	–2 438.3	–1.7%
LODZ	Aug. 2008	670 916.5	50 243.3	7.0%	POMOR	Aug. 2008	3 249 137.1	–36 408.9	–1.1%
	Sep. 2008	655 399.3	–117 149.7	–21.8%		Sep. 2008	2 777 050.6	–1 257.4	0.0%
	Oct. 2008	688 589.1	–133 454.9	–24.0%		Oct. 2008	2 813 106.6	148 553.5	5.0%
LUBEL	Aug. 2008	267 399.7	20 197.4	7.0%	SLAS	Aug. 2008	2 018 974.0	109 459.1	5.1%
	Sep. 2008	251 500.4	21 210.7	7.8%		Sep. 2008	2 033 359.7	99 647.6	4.7%
	Oct. 2008	229 992.0	55 307.3	19.4%		Oct. 2008	2 055 233.4	277 677.0	11.9%
LUBUS	Aug. 2008	191 778.6	–10 576.8	–5.8%	SWIET	Aug. 2008	44 388.4	3 288.1	6.9%
	Sep. 2008	165 958.4	–9 851.2	–6.3%		Sep. 2008	50 093.2	–1 447.3	–3.0%
	Oct. 2008	165 791.9	–14 633.0	–9.7%		Oct. 2008	57 544.2	–4 305.6	–8.1%
MALOP	Aug. 2008	340 101.7	10 370.1	3.0%	WMAZU	Aug. 2008	596 822.7	–70 786.5	–13.5%
	Sep. 2008	330 380.4	17 550.7	5.0%		Sep. 2008	509 655.0	–12 445.7	–2.5%
	Oct. 2008	372 279.4	33 049.1	8.2%		Oct. 2008	518 332.9	–8 157.5	–1.6%
MAZOW	Aug. 2008	1 164 630.0	–47 608.8	–4.3%	WLKP	Aug. 2008	877 298.9	–3 789.6	–0.4%
	Sep. 2008	1 100 205.0	–5 274.2	–0.5%		Sep. 2008	873 512.9	–22 985.4	–2.7%
	Oct. 2008	1 156 500.0	37 621.6	3.2%		Oct. 2008	958 802.7	–28 085.0	–3.0%
OPOL	Aug. 2008	255 659.1	–42 686.6	–20.0%	ZPOM	Aug. 2008	456 023.7	–6 690.4	–1.5%
	Sep. 2008	242 074.3	–38 741.5	–19.1%		Sep. 2008	433 956.7	10 231.0	2.3%
	Oct. 2008	279 875.5	–55 077.1	–24.5%		Oct. 2008	411 761.7	47 651.9	10.4%

Source: own study.

can be improved by using econometric models better describing the sales in the object. Forecasts obtained in this approach are on average underestimated and their average relative error is 8.8%.

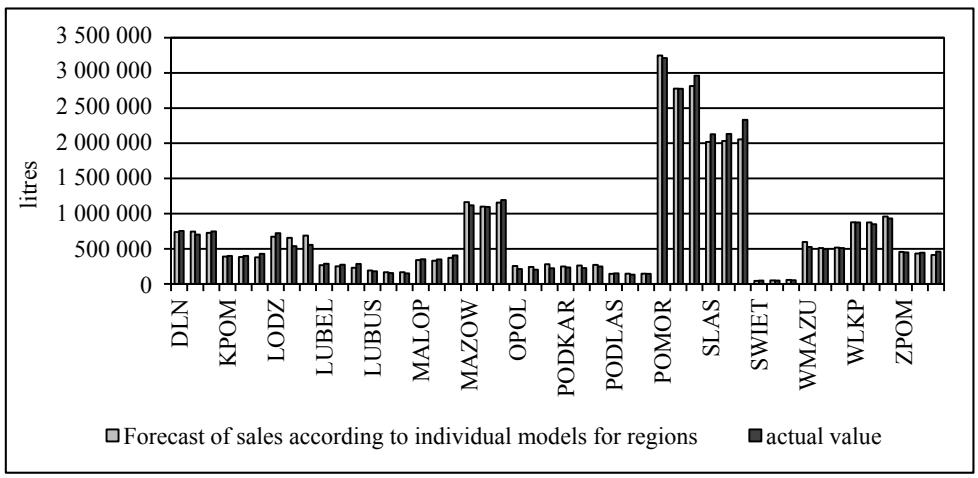

Figure 7. Forecast of unleaded gasoline sales according to models for the individual regions

Source: own study.

In the static approach, cross-section models for sales forecasting have been used. In the analyzed case, because of the seasonality of gasoline sales, cross-section models for homologous periods have been estimated. This means that the period of the last available observation for the forecast for August 2008 was August 2007, and so on. Using previously presented gasoline sales determinants, the following cross-section model has been estimated:

$$y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + b_3 x_{3i} + b_4 x_4 + u_i \tag{1.3}$$

where: x_{1i} – a retail price of gasoline in region i,
x_{2i} – the average gross wages in the industry for region i,
x_{3i} – MPC in region i,
x_4 – dummy variable with value 1 for Pomorskie region (location of the producer) and value 0 for other regions.

The sales forecast for the regions has been set using the producer's assumptions about retail prices, forecasting values of gross wages and the last available observation concerning MPC[7].The results are presented in Table 4 and in Figure 8.

[7] Due to the construction of the coefficient and the short-term forecast it has been assumed that in the forecasting period this ratio does not change.

Table 4. Detailed forecast based on model 1.3

3 Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error	Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error
DLN	Aug. 2008	814 599.0	–61 538.2	–8.2%	PODKAR	Aug. 2008	184 534.5	49 960.7	21.3%
	Sep. 2008	770 759.5	–69 398.0	–9.9%		Sep. 2008	285 189.3	–60 319.5	–26.8%
	Oct. 2008	915 632.8	–168 959.4	–22.6%		Oct. 2008	356 987.4	–108 263.7	–43.5%
KPOM	Aug. 2008	388 609.9	10 934.7	2.7%	PODLAS	Aug. 2008	398 051.7	–246 303.0	–162.3%
	Sep. 2008	502 167.5	–103 071.0	–25.8%		Sep. 2008	470 924.1	–337 867.7	–253.9%
	Oct. 2008	712 628.4	–283 354.2	–66.0%		Oct. 2008	707 488.5	–562 746.5	–388.8%
LODZ	Aug. 2008	388 101.6	333 058.2	46.2%	POMOR	Aug. 2008	3 302 774.4	–90 046.3	–2.8%
	Sep. 2008	451 174.9	87 074.7	16.2%		Sep. 2008	2 757 429.9	18 363.4	0.7%
	Oct. 2008	629 090.4	–73 956.2	–13.3%		Oct. 2008	3 019 902.5	–58 242.3	–2.0%
LUBEL	Aug. 2008	168 884.7	118 712.4	41.3%	SLAS	Aug. 2008	2 134 647.6	–6 214.5	–0.3%
	Sep. 2008	243 527.5	29 183.6	10.7%		Sep. 2008	2 185 933.4	–52 926.1	–2.5%
	Oct. 2008	372 357.7	–87 058.4	–30.5%		Oct. 2008	2 578 758.9	–245 848.5	–10.5%
LUBUS	Aug. 2008	105 778.2	75 423.6	41.6%	SWIET	Aug. 2008	169 707.0	–122 030.5	–256.0%
	Sep. 2008	291 129.4	–135 022.3	–86.5%		Sep. 2008	297 142.5	–248 496.6	–510.8%
	Oct. 2008	362 905.4	–211 746.5	–140.1%		Oct. 2008	310 472.7	–257 234.1	–483.2%
MALOP	Aug. 2008	395 941.5	–45 469.6	–13.0%	WMAZU	Aug. 2008	226 929.8	299 106.4	56.9%
	Sep. 2008	466 846.7	–118 915.6	–34.2%		Sep. 2008	316 446.5	180 762.8	36.4%
	Oct. 2008	671 923.1	–266 594.6	–65.8%		Oct. 2008	435 079.5	75 096.0	14.7%
MAZOW	Aug. 2008	1 204 499.9	–87 478.7	–7.8%	WLKP	Aug. 2008	567 773.4	305 735.8	35.0%
	Sep. 2008	1 242 787.6	–147 856.8	–13.5%		Sep. 2008	635 905.6	214 621.9	25.2%
	Oct. 2008	1 454 691.4	–260 569.8	–21.8%		Oct. 2008	867 669.0	63 048.7	6.8%
OPOL	Aug. 2008	387 949.6	–174 977.1	–82.2%	ZPOM	Aug. 2008	452 538.0	–3 204.7	–0.7%
	Sep. 2008	556 456.1	–353 123.3	–173.7%		Sep. 2008	586 807.6	–142 620.0	–32.1%
	Oct. 2008	651 860.4	–427 062.1	–190.0%		Oct. 2008	654 421.5	–195 008.0	–42.4%

Source: own source.

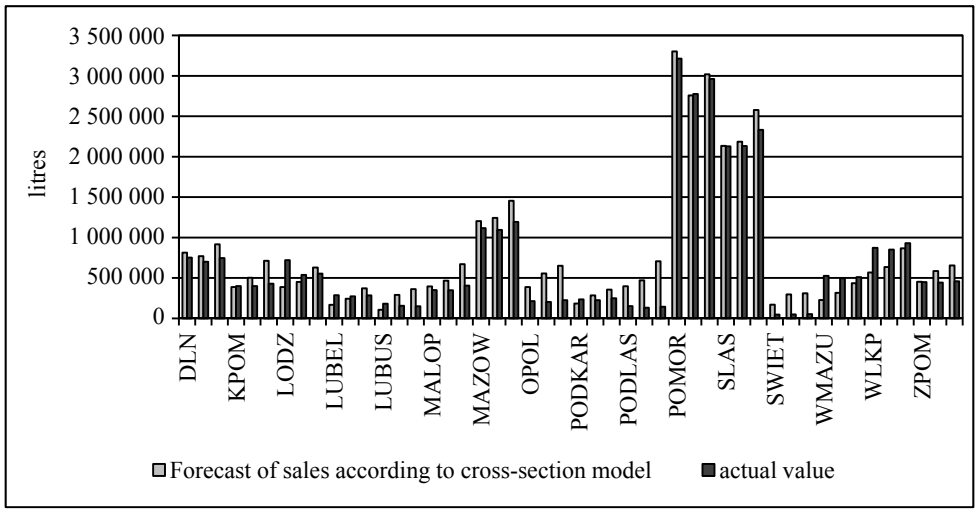

Figure 8. Forecast of unleaded gasoline sales according to cross-section models in homologous periods

Source: own study.

It can be noticed that the results obtained in this approach are much worse than in the previous one. The average ex-post error is about 28.2% and the forecasts are overestimated. The sales forecast for some provinces can be regarded as completely inaccurate (Świętokrzyski, Podlaski, Opolski). The reason for it is the heterogeneity of the objects that are not included in the presented cross-section models.

In the quasi-dynamic approach the general form of model 1.3 has been used. 31 cross-section models of sales for each month of the analyzed period have been estimated in the first place. They were used to check whether structural parameters' variations in time are statistically significant. The above mentioned model has been also estimated for the average values in the analyzed period (average values model). After receiving the series of structural parameters estimates, their linear trends have been estimated. It appeared that parameters b_0 and b_1 do not show significant changes in time. Parameters b_2 and b_4 are characterized by insignificant trends, they show, however, a significant seasonal variation. Parameter b_3 is characterized by a significant decreasing tendency as well as seasonal deviations. Therefore it was decided to use the following solution. For the purpose of the forecast, parameters b_0 and b_1 obtained from the average values model have been used as well as forecasting values of the parameters b_2, b_3 i b_4, according to their trends. The results are presented in Table 5 and in Figure 9.

While analyzing the presented results it should be noticed that in the regions where the sale is the lowest the forecast values are negative. This is because, in the case of cross-section modelling, when values of a dependent variable vary in space and additionally when the objects are heterogeneous, estimated parameters do not

Table 5. Detailed forecast based on model 1.3 – quasi-dynamic approach

4 Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error	Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error
DLN	Aug. 2008	604 641.8	148 419.0	19.7%	PODKAR	Aug. 2008	–4 445,7	238 940,9	101,9%
	Sep. 2008	620 740.8	80 620.6	11.5%		Sep. 2008	141 428,2	83 441,5	37,1%
	Oct. 2008	504 211.3	242 462.2	32.5%		Oct. 2008	26 246.0	222 477.7	89.4%
KPOM	Aug. 2008	189 590.0	209 954.6	52.5%	PODLAS	Aug. 2008	239 283.6	–87 534.9	–57.7%
	Sep. 2008	369 198.4	29 898.1	7.5%		Sep. 2008	350 115.0	–217 058.7	–163.1%
	Oct. 2008	356 152.3	73 121.9	17.0%		Oct. 2008	379 591.2	–234 849.2	–162.3%
LODZ	Aug. 2008	185 448.9	535 710.9	74.3%	POMOR	Aug. 2008	2 934 695.4	278 032.8	8.7%
	Sep. 2008	295 205.9	243 043.7	45.2%		Sep. 2008	2 621 150.9	154 642.4	5.6%
	Oct. 2008	263 912.5	291 221.6	52.5%		Oct. 2008	2 459 303.0	502 357.2	17.0%
LUBEL	Aug. 2008	–40 582.0	328 179.2	114.1%	SLAS	Aug. 2008	1 899 383.9	229 049.1	10.8%
	Sep. 2008	77 958.3	194 752.8	71.4%		Sep. 2008	2 068 271.6	64 735.6	3.0%
	Oct. 2008	32 399.6	252 899.7	88.6%		Oct. 2008	1 956 935.8	375 974.7	16.1%
LUBUS	Aug. 2008	–123 071.6	304 273.4	167.9%	SWIET	Aug. 2008	–40 237.3	87 913.8	184.4%
	Sep. 2008	136 663.2	19 443.9	12.5%		Sep. 2008	144 610.7	–95 964.8	–197.3%
	Oct. 2008	22 004.3	129 154.6	85.4%		Oct. 2008	–32 443.6	85 682.2	160.9%
MALOP	Aug. 2008	207 491.9	142 979.9	40.8%	WMAZU	Aug. 2008	29 036.7	496 999.4	94.5%
	Sep. 2008	323 860.9	24 070.2	6.9%		Sep. 2008	164 772.7	332 436.6	66.9%
	Oct. 2008	312 849.7	92 478.8	22.8%		Oct. 2008	93 865.1	416 310.3	81.6%
MAZOW	Aug. 2008	952 906.7	164 114.6	14.7%	WLKP	2008.08	382 544.2	490 965.1	56.2%
	Sep. 2008	1 078 201.7	16 729.0	1.5%		2008.09	509 058.2	341 469.3	40.1%
	Oct. 2008	935 592.4	258 529.2	21.7%		2008.10	496 379.6	434 338.1	46.7%
OPOL	Aug. 2008	146 750.7	66 221.8	31.1%	ZPOM	2008.08	251 675.7	197 657.5	44.0%
	Sep. 2008	398 171.5	–194 838.7	–95.8%		2008.09	451 871.2	–7 683.6	–1.7%
	Oct. 2008	263 218.0	–38 419.6	–17.1%		2008.10	281 841.9	177 571.6	38.7%

Source: own study.

fully reflect relations among variables. This method of forecasting enabled to achieve the average squared error of 35% which means that a quasi-dynamic approach should not be used in that type of analysis.

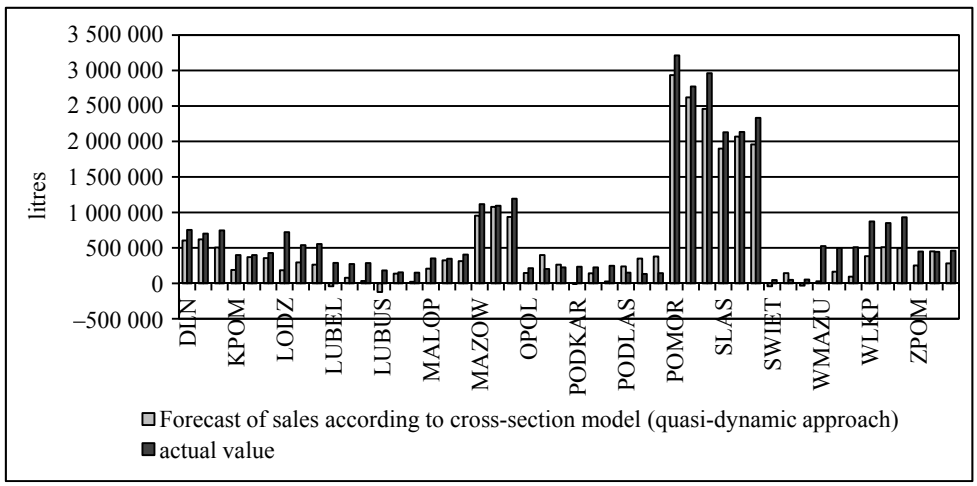

Figure 9. Forecast of the unleaded gasoline sale according to cross-section models including structural parameters variation in time

Source: own study.

The collected data also allows to conduct the combined analysis in time and in space. Because of the fact that dependent variable varies in time and space, cross-section-time-series dynamic model with individual effects in the following form was used:[8]

$$y_{it} = b_{0,it} + b_{1,it}x_{1,it} + b_{2,it}y_{i,t-1} + v_i + u_{it} \tag{1.4}$$

where: x_{1it} – retail gasoline price in region i in time period t,
$y_{i,t-1}$ – gasoline sales in region i in time period $t-1$,
v_i – individual effect for region i[9].

The above presented model was estimated with the OLS method. The detailed forecast was produced based on previous assumptions about average retail prices and by including individual effects. The results are presented in Table 6 and in Figure 10. The obtained results indicate a slightly lower average forecast accuracy than in the second approach when the individual models for each region were used. The average ex-post error is 13.2% and forecasts are overestimated. The least accurate sale forecast was determined for: Świętokrzyski, Zachodniopomorski and Opolski.

[8] The average gross wages appeared to be insignificant and were removed from the model.

[9] Factors changing in space, such as the number of petrol stations, highways, proximity to the border, are the reason for introducing individual effects for the region.

Table 6. Detailed forecast based on model (1.4) – cross-section-time-series approach

5 Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error	Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error
DLN	Aug. 2008	752 743.5	317.3	0.0%	PODKAR	Aug. 2008	230 494.6	4 000.5	1.7%
	Sep. 2008	767 568.7	–66 207.3	–9.4%		Sep. 2008	227 847.2	–2 977.5	–1.3%
	Oct. 2008	769 090.8	–22 417.3	–3.0%		Oct. 2008	220 275.5	28 448.2	11.4%
KPOM	Aug. 2008	399 956.7	–412.1	–0.1%	PODLAS	Aug. 2008	146 839.3	4 909.4	3.2%
	Sep. 2008	403 608.7	–4 512.2	–1.1%		Sep. 2008	146 168.2	–13 111.8	–9.9%
	Oct. 2008	396 713.9	32 560.4	7.6%		Oct. 2008	133 477.2	11 264.8	7.8%
LODZ	Aug. 2008	607 256.2	113 903.6	15.8%	POMOR	Aug. 2008	2 881 218.3	331 509.9	10.3%
	Sep. 2008	629 378.1	–91 128.4	–16.9%		Sep. 2008	2 854 399.3	–78 606.0	–2.8%
	Oct. 2008	634 434.4	–79 300.3	–14.3%		Oct. 2008	2 830 375.7	131 284.5	4.4%
LUBEL	Aug. 2008	285 966.4	1 630.8	0.6%	SLAS	Aug. 2008	2 104 976.2	23 456.8	1.1%
	Sep. 2008	293 924.9	–21 213.8	–7.8%		Sep. 2008	2 099 355.6	33 651.7	1.6%
	Oct. 2008	290 272.1	–4 972.7	–1.7%		Oct. 2008	2 086 343.8	246 566.6	10.6%
LUBUS	Aug. 2008	196 164.8	–14 962.9	–8.3%	SWIET	Aug. 2008	68 564.4	–20 888.0	–43.8%
	Sep. 2008	200 453.1	–44 345.9	–28.4%		Sep. 2008	75 849.2	–27 203.3	–55.9%
	Oct. 2008	196 967.5	–45 808.6	–30.3%		Oct. 2008	76 486.5	–23 247.8	–43.7%
MALOP	Aug. 2008	381 372.8	–30 901.0	–8.8%	WMAZU	Aug. 2008	538 830.9	–12 794.7	–2.4%
	Sep. 2008	396 703.9	–48 772.8	–14.0%		Sep. 2008	553 036.3	–55 827.0	–11.2%
	Oct. 2008	121 918.5	283 410.0	69.9%		Oct. 2008	554 709.5	–44 534.1	–8.7%
MAZOW	Aug. 2008	1 081 978.2	35 043.0	3.1%	WLKP	Aug. 2008	897 390.5	–23 881.2	–2.7%
	Sep. 2008	1 099 361.4	–4 430.6	–0.4%		Sep. 2008	896 379.4	–45 851.9	–5.4%
	Oct. 2008	1 103 251.7	90 869.8	7.6%		Oct. 2008	885 647.5	45 070.2	4.8%
OPOL	Aug. 2008	258 810.3	–45 837.8	–21.5%	ZPOM	Aug. 2008	587 842.5	–138 509.2	–30.8%
	Sep. 2008	268 162.8	–64 830.0	–31.9%		Sep. 2008	611 664.6	–167 476.9	–37.7%
	Oct. 2008	267 781.8	–42 983.5	–19.1%		Oct. 2008	622 089.1	–162 675.6	–35.4%

Source: own study.

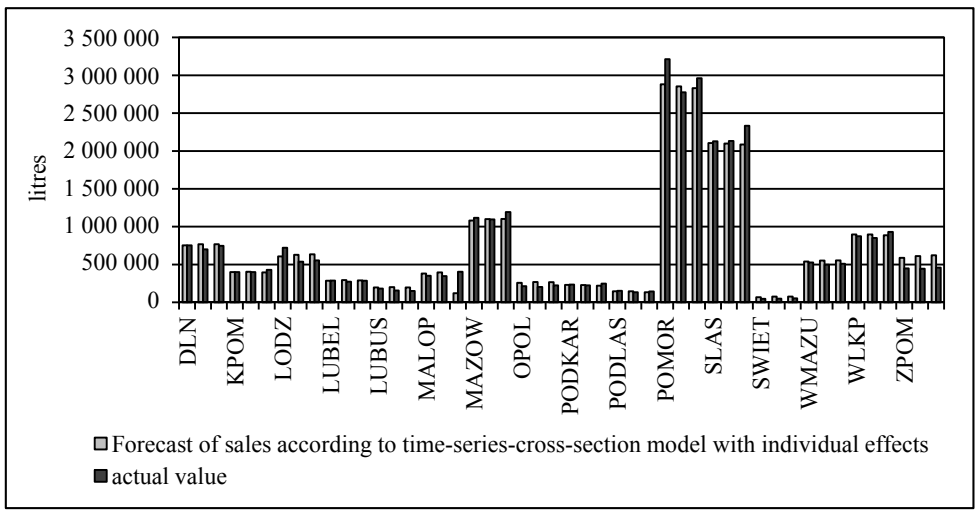

Figure 10. Forecast of unleaded gasoline sales according to time-series-cross-section model including individual effects.

To increase accuracy, the combined forecast of two predictions, which were considered the best, was created. The components of the combined forecast were the results obtained using the second approach (individual models for each region) and cross-section-time-series model. Weights were assigned on the basis of ex-post errors for those approaches [Dittmann 2008]. The results are presented in Table 7 and in

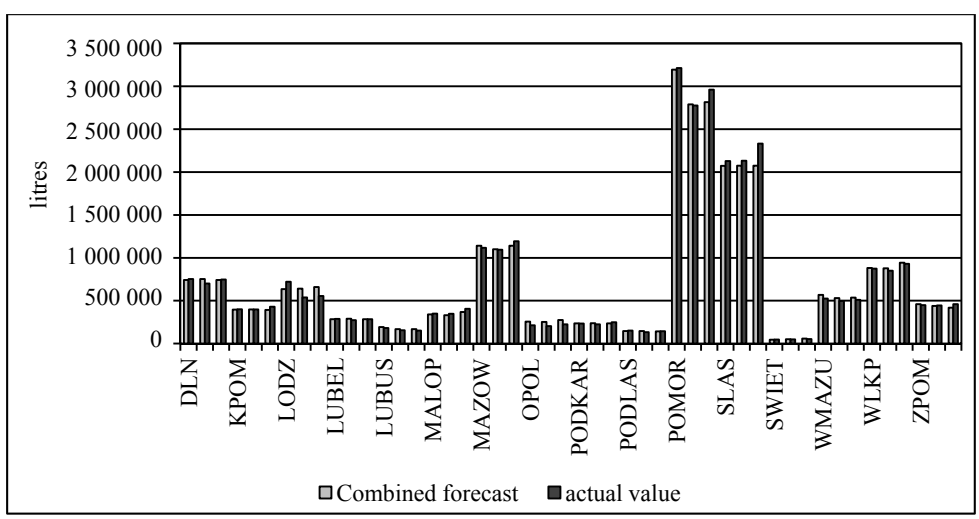

Figure 11. Combined forecast of gasoline sales

Source: own source.

Table 7. Detailed forecast based on combined prediction

6 Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error	Region	Forecasting period	Forecast value (litres)	Ex-post error	Relative ex-post error
DLN	Aug. 2008	742 501.5	10 559.3	1.4%	PODKAR	Aug. 2008	235 696.1	–1 200.9	–0.5%
	Sep. 2008	752 165.7	–50 804.3	–7.2%		Sep. 2008	237 517.3	–12 647.6	–5.6%
	Oct. 2008	740 883.1	5 790.3	0.8%		Oct. 2008	235 158.0	13 565.8	5.5%
KPOM	Aug. 2008	397 122.3	2 422.2	0.6%	PODLAS	Aug. 2008	145 540.8	6 207.9	4.1%
	Sep. 2008	398 081.4	1 015.1	0.3%		Sep. 2008	145 830.5	–12 774.2	–9.6%
	Oct. 2008	391 817.0	37 457.2	8.7%		Oct. 2008	141 727.8	3 014.2	2.1%
LODZ	Aug. 2008	635 733.3	85 426.5	11.8%	POMOR	Aug. 2008	3 194 209.8	18 518.4	0.6%
	Sep. 2008	641 018.2	–102 768.5	–19.1%		Sep. 2008	2 788 598.2	–12 804.9	–0.5%
	Oct. 2008	658 659.4	–103 525.2	–18.6%		Oct. 2008	2 815 684.8	145 975.4	4.9%
LUBEL	Aug. 2008	283 949.2	3 648.0	1.3%	SLAS	Aug. 2008	2 071 704.1	56 729.0	2.7%
	Sep. 2008	289 315.6	–16 604.5	–6.1%		Sep. 2008	2 073 823.4	59 183.9	2.8%
	Oct. 2008	283 722.8	1 576.6	0.6%		Oct. 2008	2 074 308.0	258 602.5	11.1%
LUBUS	Aug. 2008	192 172.4	–10 970.6	–6.1%	SWIET	Aug. 2008	44 823.2	2 853.2	6.0%
	Sep. 2008	169 055.4	–12 948.2	–8.3%		Sep. 2008	50 556.5	–1 910.5	–3.9%
	Oct. 2008	168 590.9	–17 432.0	–11.5%		Oct. 2008	57 884.9	–4 646.3	–8.7%
MALOP	Aug. 2008	340 832.4	9 639.4	2.8%	WMAZU	Aug. 2008	567 913.7	–41 877.6	–8.0%
	Sep. 2008	331 554.7	16 376.4	4.7%		Sep. 2008	531 280.6	–34 071.3	–6.9%
	Oct. 2008	367 846.8	37 481.8	9.2%		Oct. 2008	536 466.7	–26 291.2	–5.2%
MAZOW	Aug. 2008	1 141 427.1	–24 405.8	–2.2%	WLKP	Aug. 2008	881 731.2	–8 221.9	–0.9%
	Sep. 2008	1 099 968.2	–5 037.4	–0.5%		Sep. 2008	878 557.4	–28 029.8	–3.3%
	Oct. 2008	1 141 551.6	52 570.0	4.4%		Oct. 2008	942 664.3	–11 946.6	–1.3%
OPOL	Aug. 2008	257 039.7	–44 067.2	–20.7%	ZPOM	Aug. 2008	460 214.9	–10 881.6	–2.4%
	Sep. 2008	253 504.6	–50 171.8	–24.7%		Sep. 2008	439 606.9	4 580.7	1.0%
	Oct. 2008	274 576.8	–49 778.5	–22.1%		Oct. 2008	418 449.0	40 964.5	8.9%

Source: own source.

Table 8. Comparison of accuracy measures for set forecasts

Prediction methodology	Mean forecast bias	Relative forecast bias	Average ex-post error (liters)	Relative ex-post error for all objects
Time trend model (1.2)	33 360	4.93%	112 152	15.79%
Individual models (second approach)	5 404	0.77%	62 766	8.83%
Cross-section model (1.3) (static approach)	–82 343	10.4%	200 145	28.2%
Cross-section model (1.3) (quasi-dynamic approach)	170 478	31.57%	250 097	35.20%
Cross-section-time-series model with individual effects	–682	–0.10%	93 757	13.20%
Combined forecast	3 715	0.5%	55 547	7.8%

Source: own study.

Figure 11. They confirm the correctness of the decision of the combined prediction. As a result of the combination the underestimated forecast was achieved with the ex-post error of 7.8%. The detailed list of accuracy measures for all created forecasts is presented in Table 8.

4. Conclusion

Forecasting the method of sales based on time-trend model enabled to achieve the average squared error of 15.8% for all regions and forecasting periods. Its advantage is that it is computationally simple and can be set on the basis of one model. Its disadvantage, however, is that it does not take into account possible changes in geographical sales structure. A solution could be an adjustment of the total sales indices for each object on the basis of the company's strategic plans, which would allow for more accurate forecasts. In the regional breakdown, based on time-series data (the second approach) one versatile econometric model for the sale of fuel cannot be applied due to the fact that the analyzed objects are heterogeneous. The sales in various provinces depends on varied factors. Therefore, for the purpose of forecasting the best fitted models should be chosen. This approach, although very laborious, gives satisfactory results in terms of an accurate forecast. The advantage of this approach is that spatially customized models enable to improve forecast accuracy. In the analyzed case several impermissible forecasts were obtained that undoubtedly affected the final accuracy of the forecast. Further research in this area should consist of searching for models that better reflect the regularity occurring in these regions. The disadvantage of the approach is its laborious intensity. Monitoring of sales in geographical terms, taking into account the business strategy and changes

in individual objects, would probably improve the accuracy of predictions using this method, thus increasing its practical usefulness.

The approach to constructing cross-section models for homologous periods was used in the static analysis based on the cross-section data. Given the very large variation in the dependent variable in space (variation coefficients in the regional breakdown for the homologous periods amounted to: 1.109 for August 2007, 1.0346 for September 2007 and 1.0167 for October 2007), the obtained forecasts are not distinguished with high accuracy. The analysis of the results, however, allows to conclude that after adjustment to certain objects, the forecast accuracy could be higher. For provinces where the sale of fuel is lower than average, some weightings downward coefficients for sales forecasts should be introduced. The advantage of this method is that it does not require complex calculations. On the other hand the disadvantage is that it does not take into account the diversity of endogenous variable in terms of geography, which reduces the accuracy of the forecasts. For the prediction of phenomena in such highly diverse objects, this method should rather not be used. In the quasi-dynamic analysis, based on cross-section data and models for all time periods, the average squared error of 35% was achieved. This is caused by the high variation of sales in space as well as the averaging and prediction of structural parameters. This also leads to the fact that it is possible to obtain negative values in provinces where there is the lowest sales of fuel, which does not let the application of this approach to the above example. The advantage of this method is an ability to determine trends in structural parameters. This method, however, is labour intensive because of the need for estimating both spatial models for each unit of time as well as time-trend models of structural parameters. For long time series the workload ratio for the effect is too high. Due to the large variation of the dependent variable, time-series-cross-section analysis was conducted in accordance with individual effects. In this approach the average squared error of 13.2% and the lowest forecast bias were achieved. The advantage of this method is that it takes into account changes both in time and space, but is rather computationally complex. This method is remarkable in supporting the increasingly common use of computer packages with the option for a panel data estimation. A combined forecast, constructed on the basis of the two most accurate predictions, allowed to determine the average underestimated forecasts, with an error of slightly below 8%. The combination of forecasts taking into account the most accurate approaches is becoming an increasingly popular method of forecasting, because it usually allows to receive the most accurate forecasts. There is a possibility, however, of obtaining estimates characterized by less accuracy than the forecasts included in the composition [Dittmann 2008].

Due to the fact that ex-post measures have been used for the evaluation of the forecasts,the above analysis can be a source to search for the best method of forecasting of sales for the company. For the above example the most accurate approaches proved to be: individual models for each region (the second approach), cross-section-time-series approach with individual effects and the combined forecast.

To use these methods in practice, constant monitoring of the obtained forecasts as well as constant updating of the models are recommended, taking into account changes in the structure of sales and marketing strategy.

Literature

Baltagi B.D., *Econometric Analysis of Panel Data*, Wiley&Sons Ltd., Chichester 2008.
Banfi S., Filippini M., Hunt L.C., *Fuel tourism in border regions: The case of Switzerland*, "Energy Economics" 2005, Vol. 27.
Beck N., *Time-Series-Cross-Section Data: What have we learnt in the past few years?*, "Annual Review of Political Sciences" 2001, Vol. 4.
Bello A., Cavero S., *The Spanish retail petroleum market: New patterns of competition since the liberalization of the industry*, "Energy Policy" 2008, Vol. 36.
Dées S., Karadeloglou P., Kaufmann R., Sánchez M., *Modelling the world oil market: Assessment of a quarterly econometric model*, "Energy Policy" 2007, Vol. 35.
Dittmann P., *Prognozowanie w przedsiębiorstwie. Metody i ich zastosowania*, Oficyna WoltersKluwer, Kraków 2008.
Ekonometria przestrzenna, ed. A. Zeliaś, PWE, Warszawa 1991.
Kufel T., *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETL*, PWN, Warszawa 2007.
Sobiechowska-Ziegert A., *Ekonometryczna analiza popytu na paliwa płynne w Polsce*, „Gospodarka w Praktyce i teorii" 2005, nr 1.
Wooldridge J.M., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge 2002.

WYKORZYSTANIE DANYCH PRZESTRZENNO-CZASOWYCH DO PROGNOZOWANIA SPRZEDAŻY W PRZEDSIĘBIORSTWIE – PRZYKŁAD RYNKU PALIW PŁYNNYCH

Streszczenie: W artykule wskazano możliwości wykorzystania danych przestrzenno-czasowych do prognozowania sprzedaży w przedsiębiorstwie. Przedstawiono różne podejścia do prognozowania ekonometrycznego przy użyciu tego typu danych. Wyznaczono krótkookresowe prognozy sprzedaży benzyny bezołowiowej Pb95 w przekroju województw oraz dokonano oceny ich jakości przy użyciu mierników ex-post. Dwie najdokładniejsze metody prognozowania wykorzystano do zbudowania prognozy kombinowanej. W podsumowaniu wskazano wady i zalety poszczególnych podejść oraz warunki ich zastosowania praktycznego.