

Rafał Kasperowicz "The Causality Relationships between Energy Variables and Sold Industrial Output in Polish Economy", *Journal of International Studies*, Vol. 4, No 1, 2011, pp. 69-80.

The Causality Relationships between Energy Variables and Sold Industrial Output in Polish Economy

PhD Rafał Kasperowicz

Department of Microeconomics

Faculty of Management

Poznań University of Economics

rafal.kasperowicz@ue.poznan.pl

Abstract. Considering that energy is a very important source of industrial production, because production does not exist without energy, there are some significant relations between the use of energy and industrial production. The aim of this paper is to identify the causality relationships between the sold industrial output of Poland and variables describing energy consumption or energy sources. Another aim is to estimate a SIO (sold industrial output) equation. According to it the time-series underwent adjustment procedure Census X11 and Hodrick-Prescott's filter. In this way fluctuations of time series were obtained. The procedure of seeking for the relationships between SIO and variables describing energy use was conducted by Granger causality analysis. Series selected by the causality relationship test were used to estimate SIO equation.

Keywords: business cycles, energy, industrial fluctuations, Granger causality, Poland.

JEL classification: E32, R33, P2.

Introduction

There is a debate among economists about the relationship between energy consumption and output fluctuations. Considering that energy is a very important source of industrial production, because production does not exist without energy, there are probably some significant relations between energy use and the industrial production. Most of the literature on the subject investigates the relations between energy consumption and output changes by testing for the causality between the variables. There are publications where it was found that causality runs from output to energy use (Cheng, Tin, 1997; Masih, Masih, 1996). There are some opposite cases where the causality runs from energy use to output (Asafu-Adjaye, 2000; Masih, Masih, 1996). Finally, there are cases where causality runs in both directions (Stern, 1993; Stern, 2000).

A visible lack of energy use research in Poland and a big economy dependence on energy, which is a important source of production, generates the need for study of relations between energy consumption and industrial production. Understanding the mechanisms is crucial in order to guarantee the continuity and stability of the economy. Therefore it should be a priority to provide new research in this field of the market behavior.

Scientific Papers
Journal of
International
Studies
© Foundation of
International
Studies, 2011
© CSR, 2011

Submitted: August, 2011
1st revision: September, 2011
Accepted: September, 2011

The aim of this paper is to identify the causality relationships between sold industrial output of Poland and variables describing energy consumption or energy sources. Another aim is to estimate a SIO (sold industrial output) equation where the energy variables will serve as explanatory variables.

Methodology

It is supposed that the time series contain the random component, the seasonal component, the cyclical component, and the trend. To undertake the analysis all the time series must undergo the procedure of disaggregation. First, seasonal component and the random component must be eliminated; second, the time series must be detrended to obtain the pure cyclical movements of the data. According to that the time-series underwent adjustment procedure of seasonal correction Census X11 in Statistica software. This procedure (Evans, 2003, pp. 212-215):

- makes the adjustment taking into consideration the number of business days;
- makes the adjustment of extreme values – most of real time-series contain weird observations, which means extreme fluctuations caused by rare events;
- makes numerous corrections – corrections due to weird observations, extreme observations and different number of business days, which can be used more than once to achieve better and better estimators;
- calculates also the percentage change from month to month of random term and trend-cycle term – when the length of time-series grows, then the share of long-term fluctuations in overall volatility grows, yet we expect random fluctuations to be on the same level.

This way Henderson's curves which reflect trends and seasonal fluctuations at the same time were obtained. The next step was to state the values of trends. The conducted research had a short-term character therefore the author decided to use stochastic trend. Kydland and Prescott in 1990 offered the method of stochastic trend estimation, which was called Hodrick-Prescott's filter (Kydland, Prescott, 1990, pp. 8-9).

This method involves defining cyclical output y_t^c as current output y_t reduced by trend output y_t^g , with trend output being a weighted average of past, current and future observations (King, Rebelo, 2000):

$$y_t^c = y_t - y_t^g = y_t - \sum_{j=-J}^J a_j y_{t-j}$$

The HP-filter computes a stochastic trend by minimizing the sum of squared deviations of a time series from its trend subject to the constraint that the sum of the squared second differences is not too large (Pedersen, 2002). The HP filter is derived by solving the following minimization problem (Mills, 2003):

$$\left\{ y_t^g \right\}_{t=0}^{\infty} \min \sum_{t=1}^{\infty} \left\{ (y_t - y_t^g)^2 + \lambda [(y_{t+1}^g - y_t^g) - (y_t^g - y_{t-1}^g)]^2 \right\}$$

where λ is a smoothing parameter that penalizes variation in the growth component. In EViews, the default parameters for λ are 100 for annual data, 1600 for three months ahead data, and 14400 for monthly data.

The next step was to test the existence of unit root in the estimated fluctuations. This can be tested by the use of the ADF model, which is concerned with the estimate of α . In the following equation, we test the null hypothesis of $\alpha = 0$ against the alternative hypothesis of α

< 0 (John, Nelson, Reetu, 2007, p. 66):

$$\Delta y_t = u + \beta t + \alpha y_{t-1} + \sum_{i=1}^k c_i \Delta y_{t-1} + \varepsilon_t$$

where Δ denotes the first difference, y_t is the time series being tested, t is the time trend variable, and k is the number of lags which are added to the model to ensure that the residuals, ε_t are white noise. To estimate the optimal lag length of time series Akaike Information Criterion (AIC) was used. The rejection of the null hypothesis indicates the time series is stationary.

To establish the relationship between variables the Error-Corrected Granger Causality Test was used. Granger introduced the concept of causality for stationary series in which information about X is expected to affect the conditional distribution of the future values of Y , given the “dependent” variable (Y) and X the “explanatory” variable. The Granger test for causality relies on the estimation of the bivariate auto-regressive models. To test for causality from X to Y , the following model is used (Bennett, 2002):

$$\Delta X_t = \sum_{i=1}^p \delta \Delta X_{t-1} + \sum_{j=1}^q \theta_j \Delta Y_{t-j} + \mu_t$$

where μ_t is white noise, p is the order of the lag for X and q is the order of the lag for Y . The null hypothesis that X does not Granger-cause Y is that $\theta_j = 0$ for $j=1, 2, \dots, q$. Thus, a rejection of the null hypothesis indicates that X Granger-causes Y .

The next step was the estimation of the regression models using OLS. The general regression model can be written as follows (Evans, 2003, pp. 68-70):

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + \varepsilon_i$$

where Y is the dependend variable, the X_k are the independent variables, β_1 is the constant term, or intercept, of the equation, the other β are the parameter estimates for each of the X terms, ε_i is the error term, and there are i observations.

Data and time of research

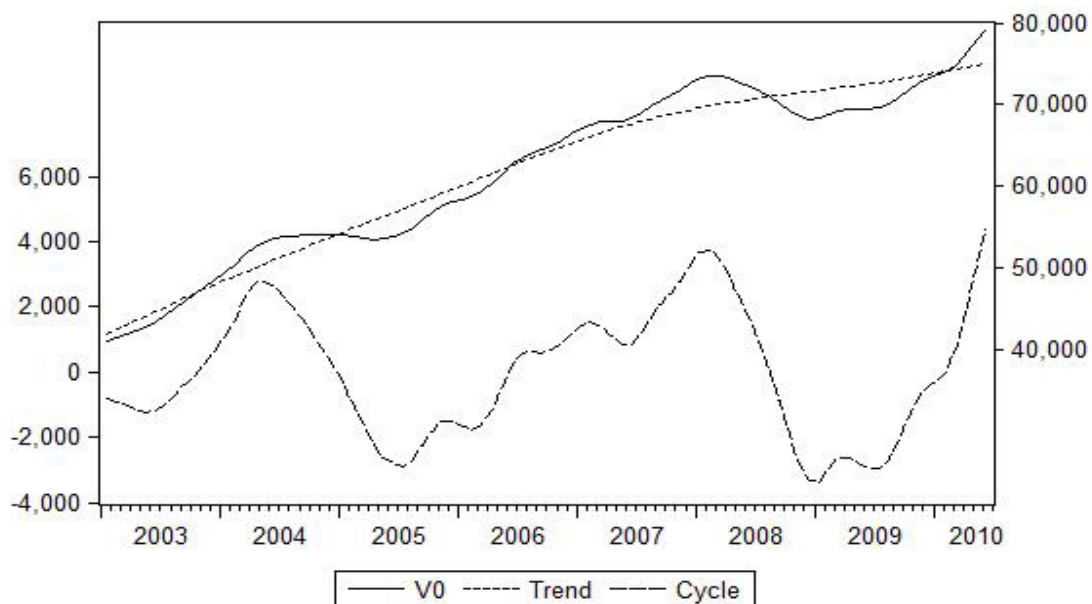
Time series used in calculations were obtained from Eurostat and from the publications of Central Statistical Office of Poland. Total sold industrial output in actual prices (SIO) was regarded the reference variable. The research was carried out from January 2003 till October 2010. The research interval was dictated by free access available databases. *Table 1* includes the description of the data.

Table 1. Interpretation of analysed variables

Sold Industrial Output	V0
Total Energy Production (TOE)	V1
Crude Oil Production (thousand tons)	V2
Crude Oil Imports (thousand tons)	V3
Gas Production (TJ)	V4
Gas Imports (TJ)	V5
Electricity Production (GWH)	V6
Mining of coal and lignite; extraction of peat	V7
Manufacture of Coke and Refined Petroleum Products	V8

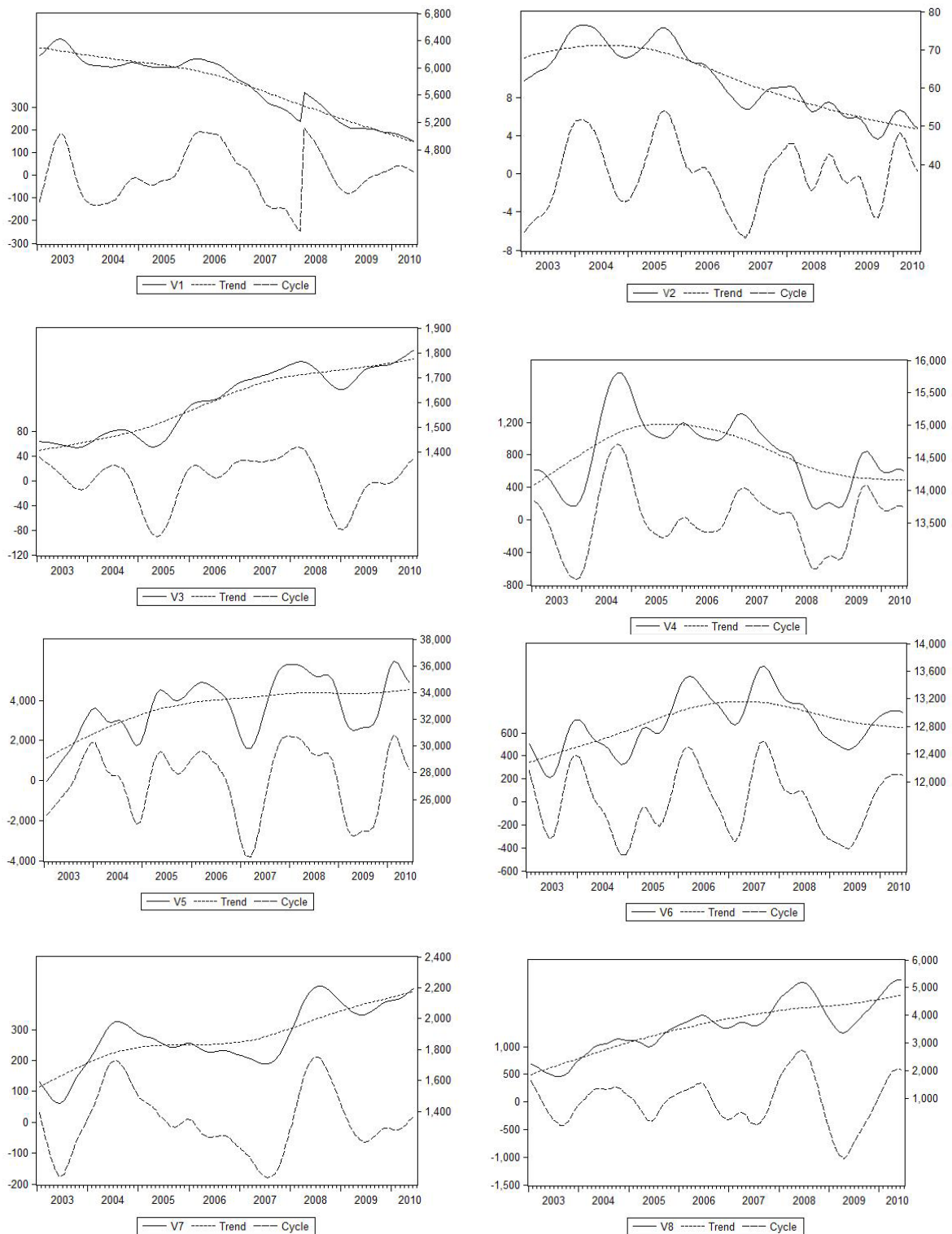
Empirical analysis

At the beginning of the empirical analysis, the selected variables underwent the decomposition procedure according to the adopted methodology. First, the time series were purified from the random and the seasonal fluctuations. Then, the stochastic HP-trends were obtained. The results given by the preliminary calculations allowed to estimate the cyclical behaviour of the data. The results are presented in graphs.



Graph 1. Sold Industrial Output (dependent variable)

Source: Own calculations.



Graph 2. The changes of independent variables
 Source: Own calculations.

The search for relationships between estimated aggregates reflecting percentage deviations from stochastic trend line was preceded by a stationarity analysis. In case of economical studies time series are regarded as a stochastic process (Charemza, Deadman,

1997, p. 107). Stochastic processes are mostly non-stationary (Cieślak, 2001, pp. 200-2001). Before conducting any further analysis the used time series were examined by ADF (Augmented Dickey-Fuller) to find the existence of a unit root. The test is needed because the applied statistical-econometrical methods assume stationarity of analysed time series (Auerbach, 1982, p. 590).

Table 2. The results of ADF test

Variable name & Lag Length	Value of ADF statistics with a constant
V0 Lag Length: 3	-4,52
V1 Lag Length: 0	-2,67
V2 Lag Length: 6	-2.87
V3 Lag Length: 10	-2.99
V4 Lag Length: 8	-3.21
V5 Lag Length: 10	-3.93
V6 Lag Length: 11	-3.41
V7 Lag Length: 6	-2.63
V8 Lag Length: 6	-4.77
Critical value	
1% critical value	-3.51
5% critical value	-2.89
10% critical value	-2.58

Source: Own calculations

Stationarity analysis of ADF test was conducted taking into consideration periods of delay given by AIC criterion in a system with constant. The results of the test are shown in *Table 2*.

In case of time series V0, V5 and V8 the zero hypothesis (H0) can be rejected and assumed their stationarity in 99% confidence interval. In case of V3, V4 and V6 the zero hypothesis (H0) can be rejected and assumed their stationarity in 95% confidence interval. The last time series V1, V2 and V7 can be considered as stationary at the level of 90% confidence interval.

Assuming stationarity of variables allowed conducting further research, the aim of which was to state the econometric causality between analysed time series, and in the next step – estimation of short-term regression.

When examining Grangers causality, it was assumed that the longest existing significant dependencies between explanatory variables and reference variable are up to 8 months lag at 5% confidence interval. The results reached are presented in *Table 3*.

Table 3. The results of Granger's causality test

Null Hypothesis:	Obs	F-Statistic	Prob.
V1 does not Granger Cause V0	82	2.58581	0.0162
V0 does not Granger Cause V1		1.47389	0.1841
V2 does not Granger Cause V0	82	5.09454	6.E-05
V0 does not Granger Cause V2		1.12641	0.3578
V3 does not Granger Cause V0	82	2.23492	0.0358
V0 does not Granger Cause V3		2.35223	0.0275
V4 does not Granger Cause V0	89	3.02341	0.0856
V0 does not Granger Cause V4		4.00796	0.0484
V5 does not Granger Cause V0	82	5.52002	2.E-05
V0 does not Granger Cause V5		0.84580	0.5663
V6 does not Granger Cause V0	87	4.33854	0.0069
V0 does not Granger Cause V6		5.54400	0.0017
V7 does not Granger Cause V0	82	2.67859	0.0131
V0 does not Granger Cause V7		2.01663	0.0581
V8 does not Granger Cause V0	82	2.50068	0.0196
V0 does not Granger Cause V8		0.90730	0.5161

Source: Own calculations

The results show that in case of all variables the Granger's test H0 can be rejected. It means that the selected data show the econometric causality against explanatory variable, so the data can be a reason of industrial output changes.

To isolate the most important advances of particular variables (Granger's significance), for each of them a regression analysis using OLS was carried out. The results presenting the most significant advances of explanatory variables along with their determination coefficient are presented in *Table 4*.

Table 4. The most significant time advances of explanatory variables through regression analysis

Explanatory variable name	Time shift in months	R ²	Adjusted R ²
V1	0	0,063150	0,051440
V2	7, 8	0,160486	0,139233
V3	1, 2, 7	0,791600	0,783585
V4	-	-	-
V5	7, 8	0,184275	0,163624
V6	1, 2, 3	0,419402	0,398416
V7	6, 7, 8	0,761979	0,752824
V8	0, 8	0,372817	0,356939

Source: Own calculations

On the basis of explanatory variables that were finally obtained, which reflect cycle

fluctuations of the data, about 40 multiple regressions, which are various combinations of depended variables presented in *Table 4*, were estimated. By the OLS estimation the procedure of eliminating insignificant variables, that were correlated with other independent variables, was considered. The evaluation criterion of regression models was based on determination coefficient. The most significant equations are presented in *Table 5*. The final Sold Industrial Output equation is presented in *Table 6*.

Table 5. Selected regression equations

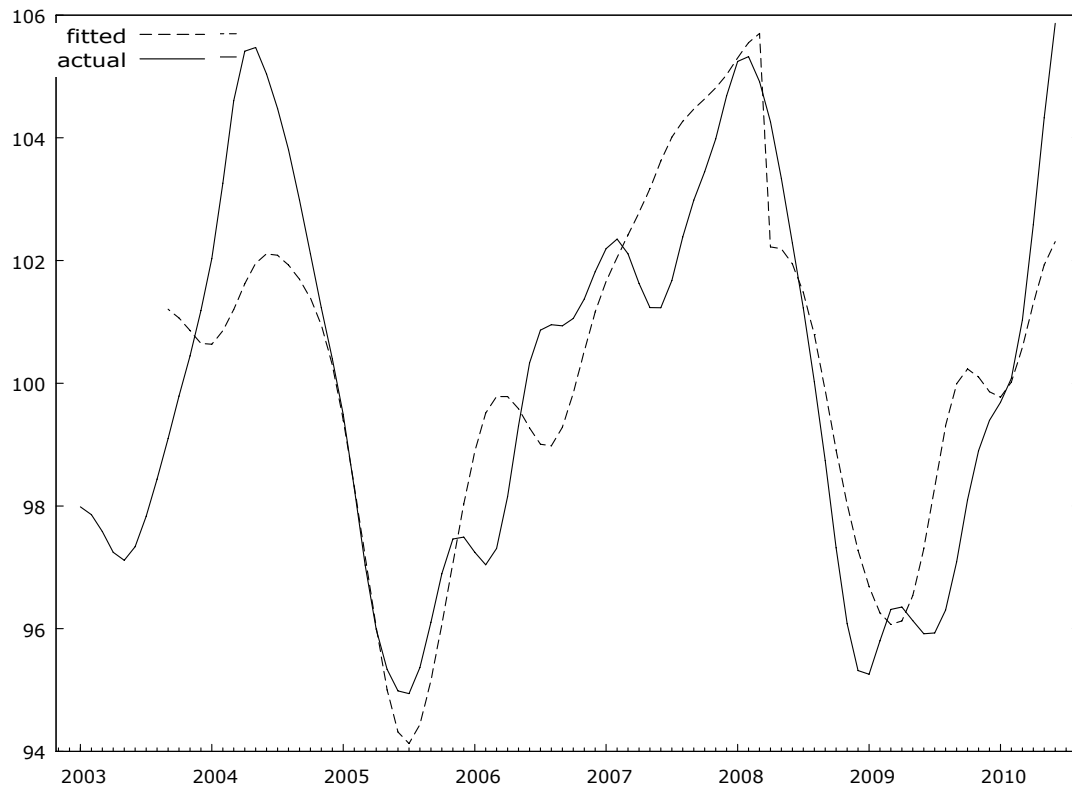
Equation 28: $V0_t = 57,370 - 0,550654V1_t + 1,02341V3_{t-2} - 0,0435849V8_{t-8}$			
Mean dependent var	100,1974	S.D. dependent var	3,178556
Sum squared resid	209,3175	S.E. of regression	1,638157
R-squared	0,744223	Adjusted R-squared	0,734386
F(3, 78)	75,65117	P-value(F)	4,97e-23
Equation 31: $V0_t = 42,7335 - 0,547364V1_t + 1,12322V3_{t-1}$			
Mean dependent var	100,1974	S.D. dependent var	3,178556
Sum squared resid	208,8414	S.E. of regression	1,625903
R-squared	0,744805	Adjusted R-squared	0,738344
F(2, 79)	115,2837	P-value(F)	3,73e-24
Equation 36: $V0_t = 40,1818 - 0,409276V1_t - 0,109060V2_{t-8} + 1,11958V3_{t-2}$			
Mean dependent var	100,1974	S.D. dependent var	3,178556
Sum squared resid	205,5172	S.E. of regression	1,623218
R-squared	0,748867	Adjusted R-squared	0,739208
F(3, 78)	77,53089	P-value(F)	2,44e-23

Source: Own calculations

Table 6. The SIO equation

$SIO_t = 40,1818 - 0,409276V1_t - 0,109060V2_{t-8} + 1,11958V3_{t-2}$			
Mean dependent var	100,1974	S.D. dependent var	3,178556
Sum squared resid	205,5172	S.E. of regression	1,623218
R-squared	0,748867	Adjusted R-squared	0,739208
F(3, 78)	77,53089	P-value(F)	2,44e-23
Log-likelihood	-154,0242	Akaike criterion	316,0484
Schwarz criterion	325,6752	Hannan-Quinn	319,9134
		DW Statistic	0,223074

Source: Own calculations



Graph 3. The fluctuations of monthly-calculated SIO in Poland

Source: Own calculations

The adjusted determination coefficient is (0,74), which allows to state that estimated equation in 74% explains the volatility of sold industrial output fluctuations. *F*-Snedecor's statistics is higher than critical value, which proves the significance of presented regression. The DW statistic isn't good, it shows that the residuals are autocorrelated. To verify the proper characteristics of estimated regression model an *ex post* analysis was carried out. Conducting this analysis needs some shortening of time series – last 21 monthly observations of independent series were deleted, whereby on that basis a SIO equation was calculated on shorter time series. That way an *ex post* SIO equation was obtained:

Table 7. The *ex post* SIO equation

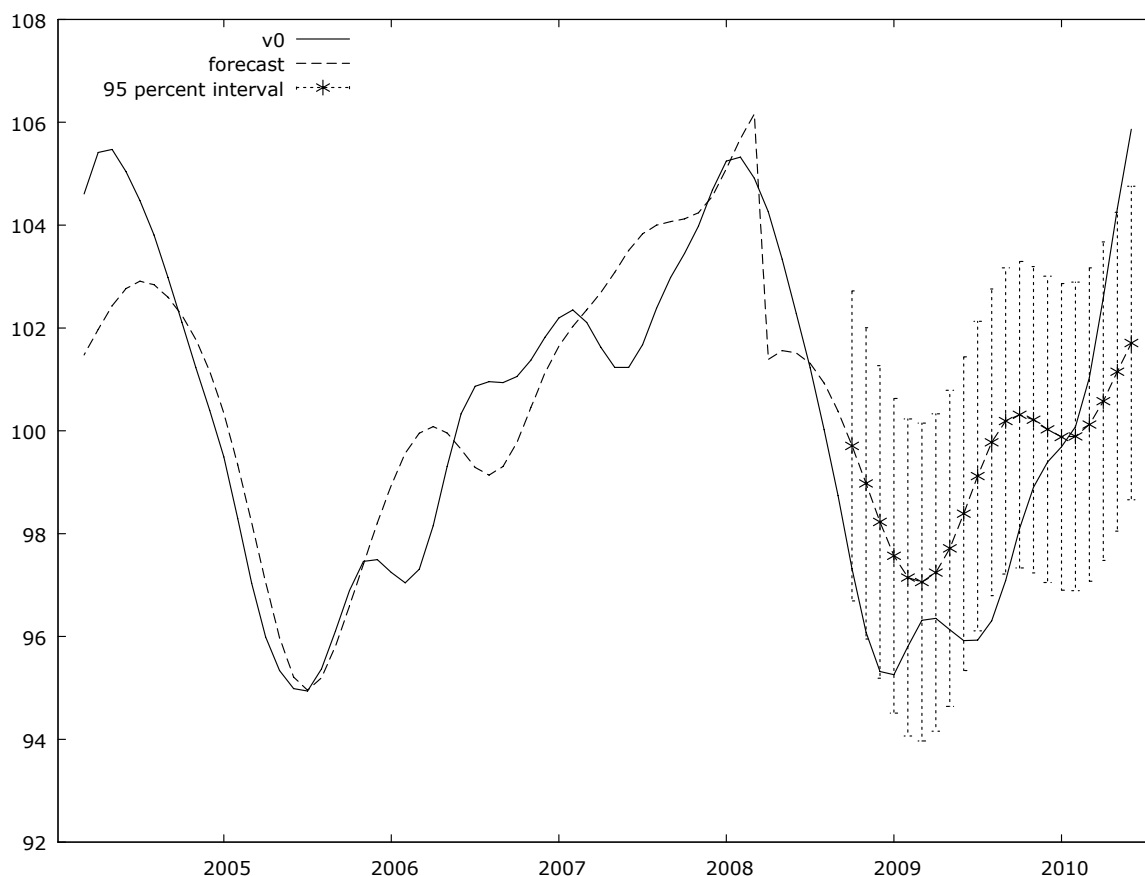
$SIO_{ex\ post} = 40,1818 - 0,409276V1_t - 0,109060V2_{t-8} + 1,11958V3_{t-2}$			
Mean dependent var	100.8588	S.D. dependent var	2.968264
Sum squared resid	123.7127	S.E. of regression	1.473227
R-squared	0.765977	Adjusted R-squared	0.753660
F(3, 57)	62.18874	P-value(F)	5.67e-18
Log-likelihood	-108.1214	Akaike criterion	224.2429
Schwarz criterion	232.6864	Hannan-Quinn	227.5519
		DW Statistic	0.333332

Source: Own calculations

Table 8. The evaluation statistics given by the ex-post equation

Mean Error	-1.0099
Mean Squared Error	5.2791
Root Mean Squared Error	2.2976
Mean Absolute Error	2.0047
Mean Percentage Error	-1.0851
Mean Absolute Percentage Error	2.0396
Theil's U	2.7552
Bias proportion, UM	0.19321
Regression proportion, UR	0.20892
Disturbance proportion, UD	0.59787

Source: Own calculations



Graph 4. The fluctuations of monthly-calculated ex-post SIO equation

Source: Own calculations

The reached *ex post* equation can properly map the fluctuations of sold industrial output. A comparison of time run of evaluated *ex post* equation with actual values is presented by graph 4. An analysis of the values given by the *ex post* equation (Table 8) proves oscillation consistency of *ex post* equation with real values. It confirms that estimated regression can certainly reproduce the direction of SIO fluctuations and the achieved regression equation is satisfying characterised.

Conclusions

The conducted research proved that there are significant causality relationships between analyzed time series and we can state that in the case of Poland the causality is running from the total energy consumption to sold industrial output. The results given by Granger Causality test have proved that the crude oil production, the gas imports and the production of coke and refined petroleum products are running their causality in the same way as total energy consumption, from the variable to sold industrial output. In case of the changes of the crude oil imports, the electricity production and the mining of coal and lignite, the causality relationships are running in both directions from the variables to sold industrial output and also from sold industrial output to the variables. Only in one case the causality relation was found to be running only in direction from sold industrial output to a variable, and the variable is the gas production.

The causality analysis proved that by using the selected variables it is possible to describe fluctuations of sold industrial output in Poland. The implementation of the elimination procedure of insignificant variables and the estimation of multiple regression helped to determine economic values, referring to energy, the fluctuations of which in the most significant way influenced the fluctuations of sold industrial output. They are the following variables:

- $V1_t$ – fluctuations of total energy production;
- $V2_{t-8}$ – fluctuations of production with a 8-months lead;
- $V3_{t-2}$ – fluctuations of crude oil imports with a 2-months lead.

The estimated equation accounts for 74% of the volatility of sold industrial output fluctuations in Poland. Moreover, on the basis of the conducted research we can state that the Polish sold industrial growth is energy efficient. When we compare the diagrams showing the long-term trends of the total energy production and the sold industrial output it becomes clear, that the changes of these variables are not cointegrated, they are going in different directions. It is positive verified by the coefficient of total energy production in the sold industrial output equation. The conducted energy efficiency of the Polish economy encourages the author to make further research on the interactions between energy use and economy.

References

- Asafu-Adjaye, J., (2000), *The relationship between energy consumption, energy prices, and economic growth: time series evidence from Asian developing countries*, Energy Econ. 22.
- Auerbach, A. J. (1982), *The Index of Leading Indicators: „Measurment without theory”*, Thirty-five Years Later, The Review of Economics and Statistics No. 82.
- Bennett, A. (2002), *Closed-End Country Found Discounts and Systematic UK and US Market movements: Co-integration and Error Corrected Granger Causality Tests*, Massey University.
- Boehm, E. A. (2001), *The Contribution of Economic Indicator Analysis to Understanding and Forecasting Business Cycles*, Melbourne Institute of Applied Economic and Social Research, Working Paper No. 17/01.
- Boehm, E. A., Summers, P. M. (1999), *Analyzing and Forecasting Business Cycles with the Aid of Economic Indicators*, Melbourne Institute of Applied Economic and Social Research, Working Paper No. 18/99.
- Charemza, W., Deadman, D. (1997), *Nowa Ekonometria*, PWE.
- Cheng, S.B., Tin, W.L. (1997), *An investigation of co-integration and casuality between energy consumption and economic activity in Taiwan, Province of China*. Energy

- Econ. 19.
- Cieślak, M., Red. Nauk. (2001), *Prognozowanie gospodarcze. Metody i zastosowanie*, PWN.
- Diedold, F. X., Rudebusch, G. D. (1989), *Scoring the Leading Indicators*, Journal of Business, Vol. 62, No. 3.
- Evans, M. K. (2003), *Practical Business Forecasting*, Blackwell Publishers.
- Granger, C. W. J. (2001a), *Investigating Casual Relations by Econometric Models and Cross-Spectral Methods*, (Econometrica, 37, 1969), w: E. Ghysels, N. R. Swanson, M. W. Watson, Essays in Econometrics, Collected Papers of Clive W. J. Granger, Volume II: *Causality, Integration and Cointegration, and Long Memory*, Cambridge University Press.
- Granger, C. W. J. (2001b), *Testing for Causality: A Personal Viewpoint*, (Journal of Economic Dynamics and Control, 2, 1980), w: E. Ghysels, N. R. Swanson, M. W. Watson, Essays in Econometrics, Collected Papers of Clive W. J. Granger, Volume II: *Causality, Integration and Cointegration, and Long Memory*, Cambridge University Press.
- Hendry, D. F., Mizon, G. E. (1999), *The Pervasiveness of Granger Causality in Econometrics*, w: R. F. Engle, H. White, *Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive W. J. Granger*, Oxford University Press.
- John, G., Nelson, P., Reetu V. (2007), *Unit Root Tests and Structural Breaks: A Survey with Applications*, REVISTA DE METODOS CUANTITATIVOS PARA LA ECONOMIA Y LA EMPRESA (3).
- King, R.G., Rebelo, S.T. (2000), *Resuscitating Real Business Cycles*, NBER Working Paper.
- Kruszka, M. (2002), *Wyodrębnianie wahań cyklicznych, maszynopis powielony*, AE Poznań.
- Kydland, F. E., Prescott, E. C. (1990), *Business Cycles: Real Facts and Monetary Myth*, Federal Reserve Bank of Minneapolis Quarterly Review.
- Lucas, Jr. R. E. (1995), *Understanding Business Cycles*, in F. E. Kydland (ed), *Business Cycle Theory, The International Library of Critical Writings in Economics*, Edward Elgar Publishing Company.
- Marcellino, M. (2004), *Leading Indicators: What Have We Learned?*, IEP-Bocconi University, IGIER and CEPR.
- Masih, A., Masih, R. (1996), *Energy consumption and real income temporal causality, results for a multi-country study based on cointegration and error-correction techniques*, Energy Econ. 18.
- McGuckin, R. H., Ozyildirim, A., Zarnowitz, V. (2001), *The Composite Index of Leading Economic Indicators: How to Make It More Timely*, NBER Working Paper No. 8430.
- Mills, T.C. (2003), *Modeling Trends and Cycles in Economic Time Series*, Loughborough University.
- Pedersen, T. M. (2002), *Alternative Linear and Non-Linear Detrending Techniques: A Comparative analysis based on Euro-Zone Data*, Copenhagen: Ministry of Economic and Business Affairs.
- Stern, D. (1993), *Energy use and economic growth in the USA, a multivariate approach*. Energy Econ. 15.
- Stern, D. (2000), *A multivariate cointegration analysis of the role of energy in the US macroeconomy* 22.