Equilibrium
2 (5) 2010
ISSN 1689-765X

Rafał Kasperowicz

Identification of Industrial Cycle Leading Indicators Using Causality Test

Key words: *business cycle, industrial fluctuations, leading indicators, leading index, Granger causality*

Abstract: The biggest business activity fluctuation analysts' attention is focused on leading indicators. It is due to their utility in forecasting resulting form their properties. Leading indicators are aggregates describing a part of economy (e.g. sector, branch) and, therefore, they also partly anticipate new behaviours of the whole of the economy. The first aim of the paper is to identify industrial business cycle leading indicators in Poland. The second aim is to estimate a leading index of cyclical fluctuations of industry. When identifying the fluctuations, first one has to purify the time-series of incidental and seasonal fluctuations. Then, the time-series underwent the adjustment procedure Census X11 and Hodrick-Prescott's filter. This is the way in which the cyclical fluctuations of the time-series were obtained. Seeking variables determining leading indicators of the reference variable was conducted on the basis of Granger causality analysis. Series selected in that way were used to create a forecasting econometric model (leading index).

Introduction

The division of business indicators into simultaneous, leading or retarded was established in 1930s by A. F. Burns and W. C. Mitchell, researchers from the National Bureau of Economic Research (NBER). An economic variable, which could be used as a business indicator should fulfil adequate criteria. Such an index should fulfil certain criteria, namely it should (Boehm, Summers, 1999, p. 27–28; Boehm, 2001, p. 13–14):

– be a significant economic variable;

– be appropriate from the point of view of statistics;

– be subject to no doubts,

– all the time indicate cohesion of dependency with lower and higher turning points of a business cycle, which means that the indicator in all researched cycles should be convergent with economic fluctuations, exceed them or to be retarded;

– in all the researched time show the cohesion of dependency with times of a decrease and an increase of the referred value;

– to be purified of unseasonable fluctuations;

– be regularly available (without any problems), best monthly, at least quarterly.

The biggest business activity fluctuation analysts' attention is focused on leading indicators. It is due to their utility in forecasting resulting form their properties. Leading indicators are aggregates describing a part of economy (e.g. sector, branch) and, therefore, they also partly anticipate new behaviours of the whole of the economy. They are very vulnerable to changes in the economic ambience and they are sensitive to most of market changes (Rekowski, 2003, p. 27; Klein, Moore, 1982, p. 2, 25–27; Diedold, Rudebusch, 1989, p. 370). Moreover, leading indicators can forecast not only turning points of the cycle but also all other changes (monthly or quarterly) which take place in the economy (Vaccara, Zarnowitz, 1978, p. 24).

Methodology

When identifying the indicators, first, one has to purify the time-series of incidental and seasonal fluctuations. Then the time-series underwent an adjustment procedure of seasonal correction – Census X11 with the use of the Statistica software. This procedure (Evans, 2003, p. 212–215):

– makes the adjustment taking into consideration the number of business days;

– makes the adjustment of extreme values – most of real time-series contains 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 was of a short–term character, therefore, the author decided to use stochastic trend. Kydland and Prescott in 1990 offered a method of stochastic trend estimation, which was named 'Hodrick-Prescott's filter'. This method allows calculating trend values due to minimisation of the sum of squares of time-series deviations of its trend but the way that the sum of squares of trend second differences was not too big (Kydland, Prescott, 1990, p. 8–9). Then the obtained curve is relatively ‘smooth’ because violent changes in the trend run are eliminated.

The need of trend determining is due to a research character, which focuses on defining leading indicators of an industrial cycle. These indicators reflect seasonal fluctuations of particular economic magnitudes. Seasonal fluctuations can be described as business fluctuations, after R. E. Lucas, and they are a set of movements around the trend estimated for the real GDP (gross domestic product) of a specified country (Lucas, 1995, p. 87). Obviously such a definition will be reserved only for business oscillations. However, nothing impedes the way to conceive the essence of seasonal fluctuations for different figures but then the trend will be searched in another series than GDP variable (Kruszka, 2002, pp. 1–2).

Eventually, after filtering we receive a series of values indicating a long-term trend which is changeable in time. Dividing empirical values of Henderson’s curves by the implementation arising from a use of HP we get (after multiplying by 100) a series showing percentage deviations form the trend line, which means cyclical fluctuations.

Seeking variables determining leading indicators of reference variable (sold industrial output) was conducted on the basis of an econometric causality analysis. Clive W. J. Granger presented most commonly used definition of causality in econometrics. This theory refers only to stochastic variables and bases on the assumption that future events cannot be the reason of the past ones (Granger, 2001a, p. 36). It is assumed that the variable x_t is the reason (Granger’s) of variable y_t only if the current values of y can be described and forecast with higher accuracy when using the past values of x than without using them in unchanged information set (Granger, 2001b, pp. 49–50; Hendry, Mizon, 1999, pp. 103–105).

Conducting Granger’s causality test enabled us to identify a maximum anticipation of explanatory variables regarded as the reason of reference variable. Granger’s test was conducted on the assumption of short-term dependencies between the variables, which meant the period not longer than a year. In connection with that the test of explanatory variables series significance in relation to reference values was done for a lead of 1, 2, 3, ..., 11, 12 months. Finally, Granger’s test indicated the maximal significant lead of explanatory variables treated as a reason. Indicating a significant lead of four periods suggests the purposefulness of including the four series describing with the lead of 1, 2, 3 and 4 periods in the explained variable regression. Using multiple regressions along with the iterative (step by step) procedure allowed limiting the set of explanatory variables leading n times to the only one – the most significant – lead of particular explanatory variable.

Series selected in that way, significantly exceeding the reference series, constitute leading indicators and can be used to create forecasting econometric models.

Data and time of research

The data used in calculations were obtained from the publications of The Central Statistical Office (Główny Urząd Statystyczny). The total sold industrial output in actual prices (SIO) was regarded as the reference variable. Yet explanatory variables are sold as industrial output in actual prices according to Polish Activities Classification (PAC).

The research was carried out in the period commencing in January 1994 and ending in February 2009. The assumption is that analysed time series consist of the following: incidental fluctuations, seasonal, cycle and long-term trend. Table 1 includes the description of the data.

Table 1. Interpretation of analysed variables

Sold industrial output according to PAC	Variable name
Mining of coal and lignite; extraction of peat	CL
Manufacture of food products and beverages	FP
Manufacture of textiles	T
Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	W
Manufacture of pulp, paper and paper products	PP
Manufacture of coke, refined petroleum	CP
Manufacture of chemicals and chemical products	CH
Manufacture of rubber and plastic products	R
Manufacture of other non-metallic mineral products	NM
Manufacture of basic metals	M
Manufacture of fabricated metal products, except machinery and equipment	MP
Manufacture of machinery and equipment n. e. c.	ME
Manufacture of electrical machinery and apparatus n. e. c.	MA
Manufacture of motor vehicles, trailers and semi-trailers	MV
Manufacture of other transport equipment	TE
Manufacture of furniture; manufacturing n. e. c.	F
General data	
Total sold industrial output	SIO

Source: own calculations.

Empirical analysis

It should be clearly stated that starting searching for relationships between especially prepared aggregates reflecting percentage deviations from stochastic trend line according to PAC was preceded by a stationarity analysis. In the case of economic studies, time series are regarded as a stochastic process (Charemza, Deadman, 1997, p. 107), yet stochastic processes are mostly non-stationary (Cieślak, 2001, p. 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. It is due to the fact that the used statistical-econometrical methods assume stationarity of an analysed time series (Auerbach, 1982, p. 590).

The stationarity analysis of ADF test was conducted with the consideration of four periods of delay in a system with a constant. The results of the test are shown in table 2.

Table 2. The results of ADF test

Variable name	Value of ADF statistics with a constant
CL	-4.664964
FP	-5.991903
T	-6.716203
W	-6.205471
PP	-6.895842
CP	-5.128629
CH	-6.156549
R	-4.538582
NM	-5.415199
M	-5.304896
MP	-4.382498
ME	-7.147920
MA	-4.764205
MV	-3.650036
TE	-3.221561
F	-5.209532
SIO	-5.816737
Critical value	
5% critical value	-2.8778

Source: own calculations.

In all the analysed time series the zero hypothesis (H0) can be rejected and their stationarity in 95% confidence interval can be assumed on the condition that in time series there is only the constant.

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

When examining Granger's causality it was assumed that the longest existing significant dependencies between explanatory variables and reference variable are up to 12 months of lead at 5% confidence interval. The results reached are shown in table 3.

The analysis of the results allows affirming that in the case of three variables (FP – Manufacture of food products and beverages; PP – Manufacture of pulp, paper and paper products; MP – Manufacture of fabricated metal products, except machinery and equipment) the H0 cannot be rejected from Granger's test and, hence, they do not show the econometric causality against explanatory variable. The rest of variables appeared to be significant so they can be a reason of cyclical fluctuations of industrial output.

Table 3. The results of Granger's causality test

Explanatory variable name	The value of a lead in a month (number of leads which should be included in the regression of explained variable)	Probability of successive leads of explanatory variables
CL	3	0.00477
FP	-	-
T	5	0.01977
W	12	0.03848
PP	-	-
CP	12	0.01102
CH	9	0.04223
R	4	0.00553
NM	9	0.04248
M	9	0.04628
MP	-	-
ME	12	0.00481
MA	12	0.00419
MV	6	0.03209
TE	4	0.04573
F	12	0.00465

Source: own calculations.

To isolate the most important advances of particular variables (Granger's significance) for each of them an iteration, after the analysis of regression, was carried out. The results presenting the most significant advances of explanatory variables along with their determination coefficient are shown in table 4.

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

Explanatory variable name	Time shift in months	R^2	Adjusted R^2
CL	1, 3	0.258	0.250
T	2	0.506	0.503
W	1, 7	0.684	0.680
CP	1, 2	0.588	0.584
CH	2	0.432	0.429
R	1, 3, 4	0.651	0.645
NM	2, 3, 5	0.626	0.620
M	1, 2	0.794	0.791
ME	2	0.380	0.376
MA	1	0.601	0.598
MV	1	0.616	0.614
TE	1, 4	0.287	0.278
F	1	0.439	0.436

Source: own calculations.

On the basis of explanatory variables that were obtained eventually, which reflect cycle fluctuations of sold industrial output due to PAC in actual prices, a number of multiple regressions which are various combinations of dependent variables (presented in table 4) were estimated. The evaluation criterion of regression models was based on the determination coefficient. As a result of the procedures of eliminating insignificant variables that were correlated with other independent variables the most significant leading indicators were selected which create an index exceeding business fluctuations of sold industrial output – *SIO*:

$$\boldsymbol{SIO_t = 17.964 + 0{,}068MV_{t-1} + 0{,}387R_{t-1} + 0{,}077M_{t-2} + 0{,}154ME_{t-2} + 0{,}131CL_{t-1}}$$

Mean dependent var 99.98 S.D. dependent var 3.30
Sum squared resid 367.38 S.E. of regression 1.45
R-squared 0.812 Adjusted R-squared 0.806
F(5, 174) 150.242 P-value(F) 3,36e-61

Table 5. The parameters of the SIO index

	coefficient	std. error	t-ratio	p-value
const	17.9646	5.25263	3.420	0.0008
CL	0.131270	0.0308697	4.252	3.45e-05
R	0.387470	0.0412472	9.394	3.26e-017
M	0.0773939	0.0171558	4.511	1.18e-05
ME	0.154324	0.0359785	4.289	2.97e-05
MV	0.0677920	0.0172721	3.925	0.0001

Source: own calculations.

where:

SIO – an index exceeding cycle fluctuations of sold industrial output in Poland;

MV_{t-1} – cycle fluctuations of sold industrial output in branch – manufacture of motor vehicles, trailers and semi-trailers in actual prices with a 1-month lead;

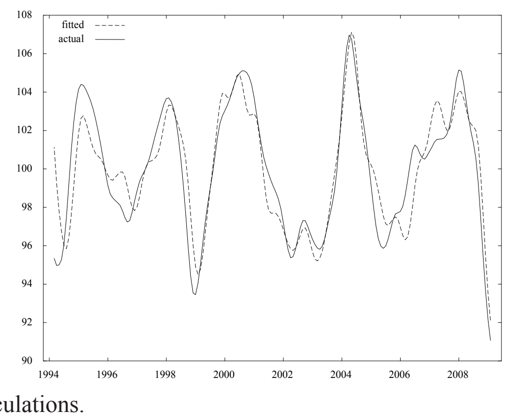
R_{t-1} – cycle fluctuations of sold industrial output in branch – manufacture of rubber and plastic products in actual prices with a 1-month lead;

M_{t-2} – cycle fluctuations of sold industrial output in branch – manufacture of basic metals in actual prices with a 2-month lead;

ME_{t-2} – cycle fluctuations of sold industrial output in branch – manufacture of machinery and equipment n. e. c. in actual prices with a 2-month lead;

CL_{t-1} – cycle fluctuations of sold industrial output in branch – mining of coal and lignite; extraction of peat in actual prices with a 1-month lead.

Diagram 1. Cycle fluctuations of monthly-calculated sold industrial output in Poland in 1994–2009



Source: own calculations.

Adjusted determination coefficient is (0.806), which allows us to state that estimated index (*SIO*) in 81% explains the volatility of cycle fluctuations of sold industrial output. *F*-Snedecor's statistics is higher than the critical value which proves the significance of presented regression. Estimated index can forecast 1-month direction of cycle fluctuations of sold industrial output in current prices in Poland. Diagram 1 presents the run of estimated index along with its actual values.

To verify the forecasting abilities of regression model an *ex post* analysis was carried out. Conducting this analysis needs some shortening of time series – the last twelve-month observation of independent series were deleted, whereby on that basis an *SIO* index was calculated on shorter time series. That way the *ex post* index was obtained is as follows:

$$\boldsymbol{SIO_{ex\ post} = 20{,}18 + 0{,}06\ MV_{t-1} + 0{,}40\ R_{t-1} + 0{,}086\ M_{t-2} + 0{,}1\ ME_{t-2} + 0{,}15\ CL_{t-1}}$$

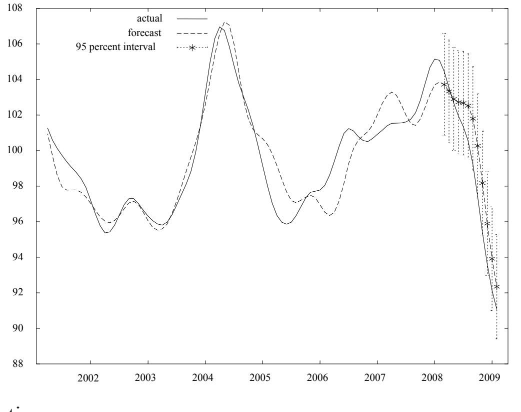
Mean dependent var 100.084 S.D. dependent var 3.186
Sum squared resid 336.99 S.E. of regression 1.442
R-squared 0.801 Adjusted R-squared 0.795
F(5, 162) 130.654 P-value(F) 5.78e-55

Table 6. The parameters of the SIO $_{ex\ post}$ index

	Coefficient	Std. error	t-ratio	p-value
const	20.1838	5.32116	3.793	0.0002
CL	0.149922	0.031474	4.763	4.20E-06
R	0.397979	0.041294	9.638	1.17E-17
M	0.086438	0.017347	4.983	1.60E-06
ME	0.103669	0.038481	2.694	0.0078
MV	0.059344	0.017454	3.4	0.0008

Source: own calculations.

Diagram 2. Cycle fluctuations of monthly-calculated sold industrial output in Poland in 1994–2009 – *ex post* index forecast



Source: own calculations.

Table 7. Values of the forecasts given by the SIO $_{ex\ post}$ index

Obs	actual	prediction	std. error	95% interval
2008:03:00	104.4506	103.711166	1.460906	100.826291 - 106.596041
2008:04:00	103.5321	103.32855	1.463879	100.437805 - 106.219296
2008:05:00	102.6622	102.90096	1.466819	100.004409 - 105.797512
2008:06:00	101.9295	102.720434	1.471556	99.814530 - 105.626337
2008:07:00	101.3384	102.671117	1.477399	99.753673 - 105.588561
2008:08:00	100.5443	102.522771	1.483239	99.593795 - 105.451747
2008:09:00	99.22406	101.798675	1.485695	98.864850 - 104.732500
2008:10:00	97.40441	100.268319	1.483944	97.337951 - 103.198687
2008:11:00	95.38673	98.171865	1.479557	95.250160 - 101.093570
2008:12:00	93.60831	95.895513	1.476295	92.980250 - 98.810775
2009:01:00	92.13192	93.917616	1.477038	91.000886 - 96.834347
2009:02:00	91.06656	92.345377	1.48296	89.416953 - 95.273802
For 95% confidence intervals, t(162, 0.025) = 1.975				

Source: own calculations.

Table 7. Forecast evaluation statistics

Mean Error	-1.4144
Mean Squared Error	3.3004
Root Mean Squared Error	1.8167
Mean Absolute Error	1.5716
Mean Percentage Error	-1.469
Mean Absolute Percentage Error	1.6197
Theil's U	1.4383
Bias proportion, UM	0.6062
Regression proportion, UR	0.071268
Disturbance proportion, UD	0.32254

Source: own calculations.

The reached *ex post* index can forecast 1-month ahead deviation of sold industrial output. A comparison of time run of evaluated *ex post* index with actual values of forecast variable is presented by diagram 2. An analysis of the values given by the *ex post* index (table 7-8) proves oscillation consistency of *ex post* index with real values. Consequently, it shows that estimated regression model can certainly indicate the direction of changes which are about to occur in cycle fluctuations of sold industrial output. The achieved model of leading indicators proves satisfying forecasting characteristics.

Conclusions

The conducted research proved that business fluctuations of sold industrial output in PAC branches in actual prices – except for three branches (FP - Manufacture of food products and beverages; PP - Manufacture of pulp, paper and paper products; MP - Manufacture of fabricated metal products, except machinery and equipment) – indicate causality according to cycle fluctuations of sold industrial output in actual prices in Poland. It means that by using these variables one can describe cycle fluctuations in the industry. The implementation of the elimination procedure of insignificant variables and the estimation of multiple regression allowed determining economic values in actual prices. Business fluctuations in the most significant way influenced future fluctuations of sold industrial output. These are following leading indicators:

MV_{t-1} – cycle fluctuations of sold industrial output in branch – manufacture of motor vehicles, trailers and semi-trailers in actual prices with a 1-month lead;

R_{t-1} – cycle fluctuations of sold industrial output in branch – manufacture of rubber and plastic products in actual prices with a 1-month lead;

M_{t-2} – cycle fluctuations of sold industrial output in branch – manufacture of basic metals in actual prices with a 2-months lead;

ME_{t-2} – cycle fluctuations of sold industrial output in branch – manufacture of machinery and equipment n. e. c. in actual prices with a 2-months lead;

CL_{t-1} – cycle fluctuations of sold industrial output in branch – mining of coal and lignite; extraction of peat in actual prices with a 1-month lead.

Estimated regression explains 81% of business fluctuation volatility of sold industrial output. Given the results of the test of forecast abilities of the estimated index, we can say that it describes the examined phenomenon adequately.

On the basis of the conducted research we can state that cycle fluctuations in the industry in Poland are incident to cycle fluctuations of the industry branches. The increase of business activity in the sales of MV, R, M, ME and CL branches determines the future growth of the entire industry.

References

Auerbach A. J., (1982), The Index of Leading Indicators: *'Measurement without theory'*, Thirty-five Years Later, "The Review of Economics and Statistics" No. 82.

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.

Cieślak M., scietific ed., (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), [in:] 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), [in:] 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*, [in:] R. F. Engle, H. White, *Cointegration, Causality, and Forecasting. A Festschrift in Honour of Clive W. J. Granger*, Oxford University Press.
Klein P. A., Moore G. H., (1982), *The Leading Indicator Approach to Economic Forecasting – Retrospect and Prospect*, NBER Working Paper no. 941.
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 Mineapolis Quarterly Reviev".
Lucas R. E. Jr., (1995), *Understanding Business Cycles*, [in:] F. E. Kydland (ed.), Business CycleTheory, 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.
McGuckin R. H., Ozyildrim A., Zarnowitz V., (2001), *The Composite Index of Leading Economic Indicators: How to Make It More Timely*, NBER Working Paper No. 8430.
Rekowski M., scientific ed., (2003), *Wskaźniki wyprzedzające jako metoda prognozowania koniunktury w Polsce*, Akademia Ekonomiczna w Poznaniu.
Vaccara B. N., Zarnowitz V., (1978), *Forecasting with the Index of Leading Indicators*, University of Chicago and NBER, Working Paper No. 244.
Zarnowitz V., Boschan Ch., (1975), *Cyclical Indicators: An Evaluation and New Leading Indexes*, Progress Report.

Identyfikacja wskaźników wyprzedzających wahania koniunkturalne przemysłu za pomocą testu przyczynowości

Słowa kluczowe: *cykl koniunkturalny, cykl przemysłowy, wskaźniki wyprzedzające, indeks wyprzedzający, przyczynowość wg Grangera*

Abstrakt: Celem niniejszego opracowania jest identyfikacja wskaźników wyprzedzających wahania koniunkturalne przemysłu. Wskaźniki wyprzedzające to agregaty ekonomiczne opisujące część rynku – sektor bądź gałąź – z tego też powodu mogą antycypować zmiany w całej gospodarce. Aby zidentyfikować wskaźniki wyprzedzające, w pierwszej kolejności określono wahania referencyjne, którymi w prezentowanym artykule, były wahania cykliczne przemysłu. Następnie oszacowano wahania cykliczne agregatów

ekonomicznych będących zmiennymi zależnymi. Estymacji wahań cyklicznych dokonano za pomocą procedur ekonometrycznych – Census X11 (usuwa wahania przypadkowe i sezonowe); filtr Hodrick'a-Prescott'a (usuwanie trendu stochastycznego). Otrzymane szeregi opisujące wahania cykliczne porównano ze sobą stosując test przyczynowości Grangera. Na tej podstawie wyłoniono wskaźniki wyprzedzające wahania koniunkturalne przemysłu, z których oszacowano złożony indeks wyprzedzający (równanie regresji wielorakiej).