PIOTR FIJAŁKOWSKI*

Models of dependence in stock exchange quotations in the example of quotes of copper and stock of KGHM Polska Miedź SA

Key words: raw materials prices, stock exchange, correlation, econometric modelling, prediction

Summary: This paper describes the structure and testing of econometric models of dependence in copper quotes and quotes of KGHM Polska Miedź SA stock on the basis of 6-month data from the London Metal Exchange (LME) and from the Warsaw Stock Exchange (WSE). The reasons for application of such models come from the high correlations between copper quotes in LME in US dollars (x) and KGHM quotes in Polish Zloty (s) in WSE (about 0.96), which confirms the natural hypothesis of the dependency of these values.

The linear, power and exponential models are tested in the paper. This choice came from the naturalness of the linear model, the possibility of reduction of the power model and of the exponential model to the linear with variable replacement, frequent practical verifiability and lack of theoretical grounds for application of other models. The obtained linear model comes in the form: $y = 0.0303x - 120.4991$, whereas the alternative power model is: $y = 8.1816 \cdot 10^{-6} x^{1.8378}$. The exponential model and the linear, power and exponential models which take into consideration "one day delay" proved to be worse.

Contrary to the expectations of the author, taking into consideration exchange rate fluctuations in Dollar against Zloty and conversion of copper prices into Zloty at the current rate did not improve the data for the model, as the corresponding correlation ratio (about 0.94) is slightly lower than in the original version. Therefore, the design of the corresponding models proved to be useless.

These types of models may be used in predicting one value on the basis of forecast quotes of another. In the described case, the model may be the basis for predicting the quotes of KGHM Polska Miedź SA on the basis of the predicted quotes of copper.

* Piotr Fijałkowski, PhD—senior lecturer, Institute of Architecture and Urban Planning, Podhalańska State Higher Vocational School in Nowy Targ.

1. Introduction

The objective of this paper is designing and testing three models of dependence between two variables: linear, power and exponential, for the actual data, in the form of stock exchange quotes of copper and stock of KGHM Polska Miedź SA taken from a 6-month period.

2. Quotes

The data for this paper come from the website (1). The following table summarises copper quotes in the London Metal Exchange (LME), with 3-month delivery periods, in USD/t (variable x) and the KGHM closing price in the Warsaw Stock Exchange (WSE) in PLN (variable y) within a 6-month period from mid-September 2010 to mid-March 2011. This summary has omitted Saturdays and Sundays, that is the days of the week during which both stock exchange facilities are off, with further days omitted: 01 Nov 2010, 11 Nov 2010, 24 Dec 2010 and 06 Jan 2011, the days on which the WSE was closed. For 15 Sep 2010, the table takes into account only copper quotes, due to the fact that this day is taken into account only in the time-shift model.

Further columns give average NBP exchange rates of Dollar against Zloty (variable x_1) and copper quotes translated into PLN (variable x_2).

Date	x (USD/t)	y (PLN)	x_1 (PLN/USD)	x_2 (PLN/t)
15.03.2011	8991.0	168.6	2.9284	26329.24
14.03.2011	9235.0	172.5	2.8825	26619.89
11.03.2011	9073.0	170.5	2.9212	26504.05
10.03.2011	9210.0	169.0	2.8849	26569.93
09.03.2011	9629.0	175.8	2.8644	27581.31
08.03.2011	9431.0	176.3	2.8549	26924.56
07.03.2011	9854.0	181.4	2.8372	27957.77
04.03.2011	9980.0	182.0	2.8646	28588.71
03.03.2011	9910.0	183.9	2.8704	28445.66
02.03.2011	9830.0	177.8	2.8843	28352.67
01.03.2011	9920.0	175.1	2.8643	28413.86
28.02.2011	9845.0	175.5	2.8765	28319.14
25.02.2011	9675.0	168.1	2.8770	27834.98
24.02.2011	9440.0	163.5	2.8941	27320.30
23.02.2011	9512.0	164.0	2.8868	27459.24
22.02.2011	9650.5	164.5	2.9198	28177.53
21.02.2011	9815.0	168.5	2.8755	28223.03
18.02.2011	9800.0	166.5	2.8803	28226.94

17.02.2011	9790.0	165.5	2.8790	28185.41
16.02.2011	9905.0	171.0	2.8864	28589.79
15.02.2011	10110.0	175.0	2.9159	29479.75
14.02.2011	10124.0	177.0	2.9219	29581.32
11.02.2011	9905.5	177.8	2.9116	28840.85
10.02.2011	9865.0	173.7	2.8795	28406.27
09.02.2011	9973.0	177.9	2.8581	28503.83
08.02.2011	9925.0	175.3	2.8471	28257.47
07.02.2011	10120.0	176.2	2.8438	28779.26
04.02.2011	9978.0	175.9	2.8666	28602.93
03.02.2011	9941.0	173.3	2.8424	28256.30
02.02.2011	9920.0	174.0	2.8230	28004.16
01.02.2011	9810.0	174.9	2.8468	27927.11
31.01.2011	9680.0	168.2	2.8845	27921.96
28.01.2011	9585.0	170.5	2.8501	27318.21
27.01.2011	9471.0	173.9	2.8508	26999.93
26.01.2011	9370.0	167.0	2.8280	26498.36
25.01.2011	9316.5	167.0	2.8558	26606.06
24.01.2011	9480.0	173.5	2.8561	27075.83
21.01.2011	9455.0	175.0	2.8779	27210.54
20.01.2011	9465.0	175.0	2.8856	27312.20
19.01.2011	9756.0	181.5	2.8879	28174.35
18.01.2011	9715.0	187.5	2.8848	28025.83
17.01.2011	9651.0	176.8	2.9181	28162.58
14.01.2011	9572.0	175.5	2.9048	27804.75
13.01.2011	9620.0	180.0	2.9353	28237.59
12.01.2011	9618.0	178.2	2.9466	28340.40
11.01.2011	9472.0	167.0	3.0065	28477.57
10.01.2011	9360.5	161.4	3.0268	28332.36
07.01.2011	9390.0	162.5	2.9818	27999.10
05.01.2011	9434.5	166.5	2.9476	27809.13
04.01.2011	9715.5	171.0	2.9415	28578.14
03.01.2011	9665.0	170.0	2.9822	28822.96
31.12.2010	9665.0	173.0	2.9641	28648.03
30.12.2010	9514.5	167.9	2.9979	28523.52
29.12.2010	9430.0	166.0	3.0383	28651.17
28.12.2010	9332.0	163.1	3.0112	28100.52
27.12.2010	9332.0	162.5	3.0215	28196.64
23.12.2010	9241.0	163.0	3.0323	28021.48
22.12.2010	9359.0	163.2	3.0322	28378.36
21.12.2010	9353.5	162.9	3.0370	28406.58
20.12.2010	9206.0	157.9	3.0396	27982.56
17.12.2010	9072.5	155.0	2.9840	27072.34

16.12.2010	9020.0	154.0	3.0125	27172.75
15.12.2010	9060.0	159.4	2.9978	27160.07
14.12.2010	9208.5	160.9	2.9708	27356.61
13.12.2010	9141.0	163.0	3.0445	27829.77
10.12.2010	9069.0	159.9	3.0458	27622.36
09.12.2010	9011.0	157.6	3.0508	27490.76
08.12.2010	8849.0	159.5	3.0558	27040.77
07.12.2010	8986.0	159.0	2.9955	26917.56
06.12.2010	8729.0	151.9	3.0057	26236.76
03.12.2010	8707.0	144.9	3.0172	26270.76
02.12.2010	8670.5	141.2	3.0282	26256.01
01.12.2010	8499.0	139.1	3.0753	26136.97
30.11.2010	8286.0	134.0	3.1308	25941.81
29.11.2010	8260.0	135.3	3.0441	25144.27
26.11.2010	8235.0	137.0	3.0363	25003.93
25.11.2010	8251.0	137.3	2.9881	24654.81
24.11.2010	8245.0	138.1	2.9808	24576.70
23.11.2010	8120.0	134.7	2.9020	23564.24
22.11.2010	8376.0	138.6	2.8591	23947.82
19.11.2010	8391.5	137.7	2.8749	24124.72
18.11.2010	8360.0	139.0	2.8871	24136.16
17.11.2010	8110.0	134.1	2.9217	23694.99
16.11.2010	8470.0	134.0	2.8937	24509.64
15.11.2010	8565.0	137.2	2.8915	24765.70
12.11.2010	8681.5	139.0	2.8865	25059.15
10.11.2010	8720.0	136.9	2.8220	24607.84
09.11.2010	8835.0	144.0	2.8278	24983.61
08.11.2010	8680.0	134.1	2.8178	24458.50
05.11.2010	8710.5	134.5	2.7648	24082.79
04.11.2010	8539.0	133.7	2.7449	23438.70
03.11.2010	8405.0	128.0	2.8010	23542.41
02.11.2010	8419.0	131.3	2.8297	23823.24
29.10.2010	8330.0	128.0	2.8873	24051.21
28.10.2010	8340.0	126.1	2.8728	23959.15
27.10.2010	8345.0	126.0	2.8567	23839.16
26.10.2010	8474.5	129.0	2.8207	23904.02
25.10.2010	8502.0	130.5	2.8119	23906.77
22.10.2010	8325.0	125.5	2.8616	23822.82
21.10.2010	8424.0	125.9	2.8193	23749.78
20.10.2010	8272.0	124.5	2.8714	23752.22
19.10.2010	8290.0	125.1	2.8268	23434.17
18.10.2010	8344.5	131.9	2.8297	23612.43
15.10.2010	8371.0	131.5	2.7717	23201.90

14.10.2010	8415.5	131.5	2.7697	23308.41
13.10.2010	8410.5	131.5	2.8327	23824.42
12.10.2010	8271.0	126.0	2.8802	23822.13
11.10.2010	8326.0	123.8	2.8525	23749.92
08.10.2010	8115.0	122.0	2.8660	23257.59
07.10.2010	8286.0	125.1	2.8401	23533.07
06.10.2010	8239.0	126.5	2.8422	23416.89
05.10.2010	8140.0	123.1	2.8838	23474.13
04.10.2010	8120.0	119.5	2.8922	23484.66
01.10.2010	8132.0	119.4	2.8772	23397.39
30.09.2010	8055.0	117.3	2.9250	23560.88
29.09.2010	8030.0	117.5	2.9227	23469.28
28.09.2010	7861.0	113.8	2.9645	23303.93
27.09.2010	7940.0	115.7	2.9425	23363.45
24.09.2010	7905.5	115.9	2.9704	23482.50
23.09.2010	7885.0	114.5	2.9853	23539.09
22.09.2010	7710.0	115.0	2.9636	22849.36
21.09.2010	7696.0	115.7	3.0053	23128.79
20.09.2010	7752.0	115.6	3.0156	23376.93
17.09.2010	7770.0	115.8	3.0079	23371.38
16.09.2010	7685.0	115.9	3.0131	23155.67
15.09.2010	7610.5	–	–	–

3. Models of dependence

The hypothesis of high correlation in copper quotes and quotes of KGHM stock is quite natural. Due to the fact that there are many other factors which may affect price of stock, and do not necessarily affect copper quotes (and vice versa), it is not obvious that the dependence between these two values should be very strong. It does not have to be of linear nature. The three models of this dependency are presented here: linear, power and exponential. Despite the naturalness of the linear model in this case, the situation of faster or slower increase/ decrease in optimism depending on the increase/ decrease in copper quotes is not easily excluded. The choice of the power and exponential models (not of the linear one) comes from the simplicity of these models, frequent practical verifiability, the possibility of reduction to the linear model with variable replacement and (obviously enough) lack of theoretical grounds for application of other models.

Due to the fact that the effect of delay in the reaction of the stock market to the quotes in the raw materials market may be expected in the described case of dependency, we will examine the linear, power and exponential models with the one-day delay taken into account.

All the calculations were done with Microsoft Excel. The detailed record of the calculations is omitted here, as they may be easily verified by the reader.

3.1. The linear model

The design of the linear model starts with determination of the coefficient of correlation between the variables *x* and *y*, described, for example, in (2):

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}},$$

which is the measure of the linear relationship between *x* and *y*.

For the given data, the result shall be:

$$r = 0.9604,$$

which is surprisingly high due to the fact that other factors should also affect the stock price, completely unrelated to copper quotes.

The linear model shall be designed as a model of linear regression described, for example, in (2), that is as an equation of linear relationship:

$$y = ax + b,$$

with the parameters *a* and *b* adjusted so that the sum of squares

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

of the differences between the actual values y_i and the theoretical values, that is

$$\hat{y}_i = ax_i + b,$$

was minimised. The following known formulae are obtained from the above condition, which allow calculation of the parameters of the equation:

$$a = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2},$$

$$b = \bar{y} - a\bar{x}.$$

With the above formulae, we shall determine the linear model for our data:

$$y = 0.0303x - 120.4991.$$

We shall select the convergence ratio as a measure of correspondence of the actual and theoretical values:

$$\varnothing^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2},$$

whose values close to 0 indicate good consistence of the actual and theoretical data.

The value is obtained for these data:

$$ø^2 = 0.0777,$$

which indicates good consistence of the actual and theoretical data.

3.2. The power model

The power model means the hypothetical dependence in the form of:

$$y = bx^a.$$

Taking logarithms for both sides of the above equality, the power model may be converted to the linear model:

$$\ln y = a\ln x + \ln b.$$

With $b' = \ln b$, $x' = \ln x$, $y' = \ln y$, the equality results:

$$y' = ax' + b',$$

which may be regarded as conversion of the power model for the variables x and y to the linear model of the variables x' and y'. When values are defined for the new variables, the correlation ratio for them may be determined along with the linear regression equation and the convergence ratio:

$$r = 0.9614,$$

$$y' = 1.8378x' - 11.7136,$$

$$ø^2 = 0.0757.$$

The comparison of the value of the correlation ratio and convergence with the corresponding coefficients for the previous model leads to the conclusion that the power model is slightly better for our values, at least theoretically. In practice, the linear model may be selected due to simplicity of calculations if (as is the case here) the parameters r and $ø^2$ are similar for both models.

After determination of $b = e^{b'}$, the form of the power model is to be determined:

$$y' = 8.1816 \cdot 10^{-6} x^{1.8378}.$$

3.3. The exponential model

The exponential model means the hypothetical dependence in the form of:

$$y = ba^x.$$

Taking logarithms for both sides of the above equality, the power model may be converted to the linear model:

$$\ln y = x\ln a + \ln b.$$

With $b' = \ln b$, $x' = \ln x$, $y' = \ln y$, the equality results:

$$y' = ax' + b',$$

which may be regarded as conversion of the exponential model for the variables *x* and *y* to the linear model of the variables *x* and *y'*. When values are defined for the new variables, the correlation ratio for them may be determined along with the linear regression equation and the convergence ratio:

$$r = 0.9580,$$

$$y' = 0.000206x - 3.163799,$$

$$\varnothing^2 = 0.0895.$$

The comparison of the values of the correlation ratio and convergence with the corresponding coefficients for the previous models leads to the conclusion that the exponential model is slightly worse for our values. After determination of $b = e^{b'}$, $a = e^{a'}$, the form of the exponential model may be determined:

$$y = 23.660322 \cdot 1.000206^x.$$

4. "Delay" models

In expectation of the delay effect in the reaction of the KGHM stock market to the quotes in the market copper, we are examining the linear, power and exponential models which take into account this delay. It would be difficult to expect any delay by more than one quote, thus we shall examine the dependence between the copper quote for the given day and the KGHM stock quote for the next day (or, in case of days off, for the first session after this time). We shall use the same variable markings for convenience.

In order to leave the calculations clear, the data for several days when the London LME was open and the Warsaw WSE was closed are not corrected here. It should not have any significance for the results. With these assumptions, the following models are obtained:

the linear model, the power model, the exponential model.

4.1. The linear model

The *x* and *y* correlation ratio, which takes into account the delay, is $r = 0.9540$, and the model comes in the form of the equation:

$$y = 0.0296x - 114.3115.$$

The convergence ratio is

$$\varnothing^2 = 0.0898.$$

4.2. The power model

The correlation ratio for the variables $x' = \ln x$, $y' = \ln y$ with the delay taken into account is $r = 0.9555$, and the model with the new variables comes in the form:

$$y' = 1.7959x' - 1.3303.$$

The convergence ratio is

$$\varnothing^2 = 0.0870.$$

The model for the variables x and y comes in the form:

$$y = 1.20039 \cdot 10^{-5} x^{1.7959}.$$

4.3. The exponential model

The correlation ratio for the variables x and $y' = \text{h}\, y$ with the delay taken into account is $r = 0.9521$, and the model comes in the form of the equation:

$$y' = 0.000201x - 3.205105.$$

The convergence ratio is

$$\varnothing^2 = 0.0935.$$

The model for the variables x and y comes in the form:

$$y = 24.658097 \cdot 1.000201^x.$$

Comparing the coefficients of correlation and the coefficients of convergence for the models without delay and with delay, it may be found out that the models which take the delay into consideration are slightly less accurate than those without delay. This could prove fast reaction of the market to significant information.

It may seem that the linear, power and exponential models are mutually exclusive, which means that when one is good, the others are not. However, our case proved that all three models are practically good. This paradox may be explained with relatively small variability of the data; for data with higher variability the theoretical values determined from them would differ more significantly.

5. Exchange rate fluctuations

The above models do not take into consideration changes in the Dollar to Zloty exchange rates. It may seem that elimination of exchange rate fluctuations by translating LME copper quotes from SD/t into PLN/t should increase the correlation between the quotes of copper and KGHM. However, this is not true. x_1 in the above table stands for the average Dollar exchange rate in NBP for the given day, and

x_2 stands for copper quotes in LME calculated at this exchange rate into Polish Zloty. The correlation ratio for LME copper quotes expressed in PLN and KGHM quotes is 0.9405 and is unexpectedly slightly lower than the corresponding coefficient for copper quotes not converted into Polish Zloty (0.9604). The strange thing is that the Dollar exchange rate and the KGHM quotes are practically not correlated (the correlation ratio for x_1 and y is 0.0086).

6. Application

The models of the type presented here may be used to predict one value on the basis of predictions of another value. In the described case, due to the good quality of the model measured with the convergence coefficient, the predictions of KGHM quotes may be obtained whose accuracy depends on the accuracy of predictions of copper quotes.

Bibliography

1. *Bizzone: Bliżej biznesu*. [online, accessed: 2011-03-16]: http://www.rynek.bizzone.pl/.
2. Aczel A. D. 2000. *Statystyka w zarządzaniu*. Transl. by Z. Czerwiński, W. Latusek. Warszawa: Wydawnictwa Naukowe PWN. ISBN 83-01-13001-6.
3. Chiang A. C. 1994. *Podstawy ekonomii matematycznej*. Transl. by E. M. Syczewska. Warszawa: Państwowe Wydawnictwo Ekonomiczne. ISBN 83-208-0942-8.
4. Januszkiewicz W. (ed.). 1991. *Giełdy w gospodarce światowej*. Warszawa: Państwowe Wydawnictwo Ekonomiczne. ISBN 83-208-0860-X.
5. Welfe A. 2009. *Ekonometria*. Warszawa: Państwowe Wydawnictwo Ekonomiczne. ISBN 978-83-208-1767-6.

Modele zależności notowań giełdowych na przykładzie notowań miedzi i akcji KGHM Polska Miedź SA

Streszczenie: Niniejsza praca opisuje konstrukcję i testowanie ekonometrycznych modeli zależności notowań miedzi i akcji KGHM Polska Miedź SA na podstawie półrocznych danych z London Metal Exchange (LME) i Giełdy Papierów Wartościowych w Warszawie (GPW). Uzasadnieniem celowości takich modeli jest wyznaczony wysoki współczynnik korelacji notowań miedzi na LME w dolarach amerykańskich (x) i notowań KGHM w złotych polskich (y) na GPW (około 0,96), potwierdzający naturalną hipotezę o związku między tymi wielkościami.

W pracy testowany jest wariant liniowy, potęgowy i wykładniczy modelu. Uzasadnieniem takiego wyboru jest naturalność modelu liniowego, sprowadzalność modelu potęgowego i wykładniczego do liniowego przez zamianę zmiennych, częste sprawdzanie się w praktyce oraz brak teoretycznych podstaw do stosowania innych modeli. Wyznaczony model liniowy ma postać:

$y = 0,0303x - 120,4991$, natomiast alternatywny potęgowy: $y = 8,1816 \cdot 10^{-6} x^{1,8378}$. Model wykładniczy oraz modele liniowy, potęgowy i wykładniczy uwzględniające opóźnienie czasowe o jedno notowanie okazały się słabsze. Wbrew oczekiwaniom autora uwzględnienie zmian kursowych dolara względem złotego i przeliczenie cen miedzi na złote według aktualnego kursu nie poprawiło danych do modelu, gdyż odpowiedni współczynnik korelacji (około 0,94) jest nieco niższy niż w wersji pierwotnej. W związku z tym konstrukcja odpowiednich modeli okazała się niecelowa.

Zastosowaniem tego typu modeli może być prognozowanie wartości jednej wielkości na podstawie prognoz drugiej. W opisanym przypadku model może być punktem wyjścia do prognozowania notowań KGHM Polska Miedź SA na podstawie prognoz notowań miedzi.

S ł o w a k l u c z o w e: ceny surowców, giełda papierów wartościowych, korelacja, modelowanie ekonometryczne, prognozowanie