Determination of Copper Price Expectations in the International Market: Some Important Variables

Salles, Andre Assis de and Magrath, Raphael Sebastian and Malheiros, Matheus Manzani

Polytechnic School, Federal University of Rio de Janeiro,
Polytechnic School, Federal University of Rio de Janeiro,
Polytechnic School, Federal University of Rio de Janeiro

13 February 2019

Determination of Copper Price Expectations in the International Market: Some Important Variables

Andre Assis de Salles
Polytechnic School, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
as@ufrj.br

Raphael Sebastian Magrath
Polytechnic School, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
sebastian@poli.ufrj.br

Matheus Malheiros
Polytechnic School, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
matheusmalheiros@poli.ufrj.br

Abstract

The purpose of this work is to identify variables that are relevant to the copper price setting in the international market. Thus statistical hypothesis tests and statistical tools that help to identify historical relevance and to measuring the intensity of the impact of each variable on the copper price on several time horizons was applied. At the end, a regression model that aims to assess the combined effect of the considered time series was estimated. The global industrial production and the aluminum price showed the greatest evidences of being relevant to the copper price. The results suggest that copper stocks, foreign exchange rates and crude oil price should also be considered.

Keywords: Copper Price. Cointegration. Causality. Impulse Response Function.

1. Introduction

Copper is the industrial metal with the highest financial volume and negotiations in the international commodity markets. Producers, consumers, and financial market players observes the copper market once copper price movements represent a relevant global economy leading indicator, important to identify macroeconomic trends. There is a relationship between the copper price and relevant indicators of global economic activity. Present in the production chain of industrial products, in particular in the electronic industry equipment, the transport sector computer industry, through the automotive vehicles production, the use of copper is very widespread among several sectors and industries, mainly due its durability, machinability and the ability to be molded with high precision and tolerance. It's copper thermal conductivity and resistance to extreme environments allows its use in heat exchangers, pressure vessels and tanks. Thus, the demand for copper means the production of items such as wires, rods, tubes, plates and ingots, which will later be used directly in a final application or in the production of another good. Currently, copper is also widely used in technology equipment such as cell phones and computers, and in products commonly found in modern homes, such as washing machines, refrigerators and air conditioners. Copper is traded among the various agents in the

market in several states, among others, refined, concentrated, blister and scrap. But the main demand, or the use, of copper for industry is the use of refined copper. And the reference price of copper traded in its various forms is refined copper. It is in this state that copper is traded in the various organized international markets, and the prices practiced in these markets are the benchmarks for establishing prices in refined copper negotiations, and in other states, in the world. The London Metal Exchange (LME) is the most traditional metal trading exchange in the world, with representative daily trading volumes, it is an important reference for the negotiations of metals for the world economy. Also worthy of mention is the Chicago Mercantile Exchange (CME), which is a major exchange for copper trading, and the Shanghai Futures Exchange (SHFE), which has become an important exchange in the global scenario, with growing Chinese participation In the metal markets in general.

As Amano and van Norden (1998) observes, knowing the swings in the price of copper is extremely important for countries whose economy depends heavily on imports or exports of copper, such as Chile. Copper price knowledge is also of extreme importance for the mining companies, responsible for removing the metal from the soil for later consumption. Understanding movements in the price of copper is also of extreme value to all market players: it allows producers to define hedging strategies, which consists of using future contracts in the financial market to curb their selling price ahead. In addition to the benefit of giving greater predictability of cash flow, this practice is even more valuable when prices fall to the maturity date of the contract used in the hedge as well as consumers who, in contrast, benefit most from the use of this strategy when prices rise. While to the speculators it allows the gain of profits in the financial market. Given the importance of this commodity for economic activity, many surveys related to the prices of copper practiced in the international market have been carried out by academics and direct participants of the international metal markets and, in particular, copper. Many studies and surveys have been developed to evaluate variables and indicators relevant to the formation of the price of copper in the international market. Among these works we can highlight the one of Cerda (2005), that of García-Cicco and Montero (2011), the one of Stürmer (2013b) and developed by Zhang et al. (2015).

This paper aims to identify and evaluate and quantify the relationship between the price of copper and some of the variables potentially relevant to the process of price formation of copper in the international market. Among these variables, we can relate: global production of refined copper, the price of oil in the international market, the price of aluminum in the international market, the global refined copper stock, the global industrial production variation and the exchange rate. In order to achieve these objectives, econometric procedures were implemented, such as: causality tests, cointegration tests and estimation of the global energy matrix crude oil is one of the most important energy sources in the world. It is an essential commodity in the global economy. Furthermore, petrochemical and chemistry industries produce or manufacture goods from petroleum, such as fertilizers used in agriculture for food industries. Despite the progress in development of renewable energy sources in recent decades, crude oil and their byproducts remain directly or indirectly present in people´s lives. Given this relevance, the crude oil price is an important variable for economic policy makers in national economies, where this commodity is the main source of energy as well as in worldwide economies.

This paper aims to identify and evaluate and quantify the relationship between the price of copper and some of the variables potentially relevant to the process of price formation of copper in the international market. Among these variables, we can relate: global production of refined copper, the price of oil in the international market, the price of aluminum in the international market, the global refined copper stock, the global industrial production variation and the exchange rate. To achieve these objectives, econometric procedures were implemented, such as: causality tests, cointegration tests and estimation of the impulse response function.

In addition to this introduction, this article presents in the next section details the methodological approach adopted for the development of the research. Following the presentation of the sample or data used, in Section 3. In Section 4 are related to analysis of the results obtained. And,

finally, Section 5 presents the conclusion and final comments of the paper.

2. The Literature Review - Some Related Work

Many studies and surveys have been developed to evaluate relevant variables and indicators for copper price formation in the international market. Among these works, some were selected here for a brief review of the literature presented in the following paragraphs.

With information from January 1994 to October 2003, Cerda (2005) sought to identify some relevant indicators as significant explanatory variables for the copper price in the international market. In addition to the copper production, the variables used were industrial production, the wholesale price index and the interest rate of the United States, South Korea, Brazil and countries in the Euro zone. While copper production shows us the behavior of the metal supply, industrial production was selected as a proxy for real demand and the wholesale price index indicates the variations in the relative copper price in each country. When the wholesale price index decreases, the relative copper price increases, causing a decrease in demand. Some comments on Cerda (2005) conclusions should be highlighted. Using regression analysis techniques, Cerda (2005) found statistical significance for industrial production, copper production, interest rate and wholesale price index (IPA) of selected countries. From the industrial production analysis and wholesale price index, one can not reject the hypothesis that these countries concentrate the positions of the international copper market, that is, no relevant economy was left out in the analysis. It should be noted that if the sample were updated this assertion would not be sustained. In recent years, China has emerged as the main agent in the copper market, as a major consumer and producer, in addition to other countries that have gained or lost importance in this market. Regarding copper production, Cerda (2005) limits its sample only to production in Chile, the most important producer of the period with a market share of 30%.

Another relevant work related to the topic was presented by García-Cicco and Montero (2011), which used models where the parameters change according to a Markov switching process. The data used in this analysis is that of the spot market price traded in the LME from 1975 to 2010. In the work of García-Cicco and Montero (2011) nine different cases are evaluated, where parameters such as variance, number of lags and constants vary or be static. García-Cicco and Montero (2011) conclude that the inclusion of an alternance regime in the variance is fundamental to explain the behavior of copper price. Models with such a feature outperform other models, including autoregressive conditional heteroskedasticity, such as the GARCH models. One extension of the study suggested by the authors is the combination of Markov switching processes with autoregressive vector models (VAR) considering other variables, such as stock level and exchange rate, to try to explain and predict the copper price better.

The Stürmer (2013b) work, which sought to identify factors responsible for long-term fluctuations in the prices of mineral commodities and, in particular, of copper, should be highlighted. Stürmer (2013b) analyzed the behavior of annual prices over a long period: from 1840 to 2010. Stürmer (2013b) evaluates only the global production of each commodity and the global income in each year, forming three different types of "shocks" in the commodity markets: supply shock or unexpected drops in copper production, such as mining strikes; shock of global activity demand or shocks in global demand for commodities, such as an unexpected growth in global activity; and other demand shocks, including all shocks are not related to the two previous categories, such as unexpected changes in stocks due to government strategic reserve programs. Using an autoregressive vector model or a VAR model, Stürmer (2013b) concludes that, in the case of copper, price fluctuations are mainly caused by demand shocks of global activity and also by supply shocks and other shocks demand shocks. That is, Stürmer (2013b) shows that changes in the level of global activity, overall copper production and the level of stocks are relevant factors in the formation of the copper price. Stürmer (2013b) makes an analysis of the impulse response function of these variables and stresses that a positive shock of demand for global activity causes a significant increase in the copper price, with the maximum effect

appearing within a year after the shock, and with relevant effects still lasting up to 10 years later while the positive supply shocks have a negative impact on copper price, relevant up to 4 years after the shock, with immediate maximum effect though on a somewhat smaller scale than the demand shocks of the global activity. Finally, other positive demand shocks also cause an immediate rise in copper price, with relevant effects lasting up to two years after the collision. Stürmer (2013b) shows the positive association between manufacturing production and demand for copper in several markets studied and using data from the period studied in 12 industrialized countries Stürmer (2013b) estimated an autoregressive distributed lag model or an ARDL model, and found the existence of long-term relationships involving copper and manufacturing activity.

Another work that deserves attention was developed by Zhang et al. (2015) on the effect of exchange rates on commodity prices with data from January 1996 to July 2015. Zhang et al. (2015) assess the existence of Granger's causality between the exchange rate and commodity prices in both directions over multiple time horizons. In the case of copper, they assess the causality between their price in the international market and the Chilean exchange rate, that is, Chilean peso per US$. In general, the effect of the commodity price on the exchange rate is more evident than the inverse relationship because a commodity can have great relevance for the economy of a country, such as through exports or the trade balance, while the relevance of a country to the market of a commodity is, in general, more limited. In the research by Zhang et al. (2015), however, the causality of the Chilean exchange rate on the copper price was very relevant, being the most relevant variable among those analyzed in the research. Zhang et al. (2015) also addresses an analysis of the impulse response function to measure the magnitude of the effect, as well as the intensity and duration, of the exchange rate on copper price, showing the existence of an immediate effect that lasts for approximately 5 days.

3. Methodological Approach

In order to achieve the objective of this work, the methodology used includes procedures that seek to evaluate, measure and understand the effects of several variables potentially relevant to the price of copper in the international market. In this sense, each variable considered important was analyzed and its relations with the price of copper evaluated. For each of these related variables, checks or hypothesis tests are performed which are presented below.

Initially, in order to characterize the time series involved in this work, the stationarity and normality assumptions were verified, respectively, through the Dickey Fuller Augmented and Jarque-Bera tests. Then, through the stationarity evaluation of the residual error of the linear combination of the involved variables, the hypotheses of cointegration between the selected variables and the copper price are verified. The determination of the cointegration between the time series the selected variables and the copper price is important per se for the estimation of autoregressive vector models or VAR models. Moreover, these models are fundamental for the study of the causality of variables and of regression models that represent the causal relationship of the selected variables with the price or variation of the price of copper in the international market. Finally, analyzes of the impulse-response function were carried out, in order to measure the possible causality verified in several time horizons, since the present study mainly tries to verify the temporal precedence between the variables

In the tests of the cointegration hypothesis between copper prices and the selected variables, the concept introduced by Engle and Granger (1987) was used and states that individually non-stationary variables may have stable long-term relationships, since a linear combination of these series has stochastic trends that cancel out, reaching stationarity. As observed by Gujarati (2004), two series are said to be cointegrated if they have a long-term relationship or equilibrium. If two variables are represented by their time series X_t and Y_t, for example, the combination that expresses the long term relationship can be obtained by linear regression between the variables, resulting in the following equation:

$$Z_t = \beta_1 + \beta_2 X_t + e_t \Rightarrow e_t = Z_t - \beta_1 - \beta_2 X_t \tag{1}$$

where X_t and Y_t are the values of two time series of the variables X and Y at time t, β_1 and β_2 are parameters and e_t is the residue or stochastic term. If the term et does not have a unit root, it can be said that the series are cointegrated, removing the possibility of a spurious regression and showing that there is a long-term relationship between the variables. The stationarity test used in this work was the unit root test of Dickey and Fuller Augmented, or ADF test. The null hypothesis of the Engle-Granger test states that the series are not cointegrated. In practice, two equations are estimated: in the first X_t is estimated as dependent variable and Y_t as independent variable, as presented in equation (1) above; and in the second X_t is estimated as dependent variable and Y_t as independent variable. If at least one of the two equations the coefficient of the dependent variable is statistically significant then the null hypothesis is rejected, that is, it can be said that the time series X_t and Y_t are cointegrated.

Autoregressive vector models, or VAR models, are commonly used in interrelated time series prediction systems and to analyze the dynamic impact of random perturbations on the system of variables. These are models in which a variable is explained by its own past value and by past values of the other endogenous variables of the model. In general, as Gujarati (2004) points out, there are no exogenous variables in the model. The VAR model with a lag, called VAR (1), can be represented by the following system of equations:

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + \beta_3 Z_{t-1} + \varepsilon_{1t} \tag{2}$$

$$Z_t = \beta_4 + \beta_5 Z_{t-1} + \beta_6 Y_{t-1} + \varepsilon_{2t} \tag{3}$$

where the time series X_t and Y_t are stationary. If this does not occur, the n differences of the variables are used until the differences from these variables become stationary, that is the time series are integrated until they become stationary. After this procedure, these variables are said integrated of order n, or $I(n)$. It is worth noting, however, that the use of large n can generate problems in small samples, since the estimation of the parameters of the VAR model will consume many degrees of freedom, as observed by Salles and Almeida (2017). In case the variables are cointegrated, the system must be altered in a way and consider such a long-term relationship. Thus, if two time series are integrated of order 1 and cointegrated, one has to obey the following relation:

$$X_t = \beta_1 + \beta_2 Y_t + \mu_t \tag{4}$$

where µt is stationary. Thus, a restricted case of VAR model, called vector model with error correction, or VEC model, consists of the system of equations presented below, in which all variables are stationary, as in the VAR model, and cointegrated - this system can be represented as follows:

$$\Delta X_t = \alpha_1 + \alpha_2 \mu_{t-1} + \varepsilon_{1t} \tag{5}$$

$$\Delta Y_t = \alpha_3 + \alpha_4 \mu_{t-1} + \varepsilon_{2t} \tag{6}$$

where Δ is the differentiation operator, that is, $\Delta X_t = X_t - X_{t-1}$. By replacing the residual term lagged, this system can be rewritten as follows:

$$X_t = \alpha_1 - (\alpha_2 \beta_2 - 1)X_{t-1} - \alpha_2 \beta_1 + \alpha_2 Y_{t-1}) + \varepsilon_{1t} \tag{7}$$

$$Y_t = \alpha_3 + (\alpha_4 + 1)Y_{t-1} - \alpha_4 \beta_1 - \alpha_4 \beta_2 X_{t-1}) + \varepsilon_{2t} \tag{8}$$

In the equations above, α_2 and α_4 are the error correction coefficients, since they indicate the magnitude of the response of the variables X_t and Y_t to a variation in the residual term μ_{t-1}. The coefficients must, in order to guarantee stability, satisfy the following constraints: $0 \leq \alpha_2 < 1$ and $-1 < \alpha_4 \leq 0$. For a positive

stochastic term, for example, ΔX_t will be positive and ΔY_t will be negative, thus restoring the equilibrium described by cointegration. In addition, the fact that the modules of these parameters are lower than the unit ensures that the model has no explosive behavior. For more details, one can draw on the work of Salles and Almeida (2017).

As observed in Gujarati (2004), although regression analysis indicates the dependence of a variable relative to other variables, this does not necessarily imply a causal relationship. In other words, the existence of a relationship between variables does not imply the existence of causality, or indicates the direction of influence between these variables. In this context Granger's causality test has become quite widespread in the econometric literature. The assumption from which Granger departs is that: the future cannot cause the past or the present; that is, if an event Y occurs after an event X, it is known that Y cannot cause X; at the same time, not necessarily X causes Y. Thus, given two time series X_t and Y_t, we are interested in whether there is a precedence relationship between them, or whether they occur simultaneously. This is the essence of Granger's causality test, which does not seek to identify a causal relationship in its sense of endogeneity. For further detail, one can draw on the work of Maddala (1992). Given the time series X_t and Y_t, the Granger causality test assumes that the relevant information for the prediction of the respective variables X and Y is contained only in the time series on these two variables. Thus, a series of stationary time X causes, in the sense of Granger, another stationary series Y if better statistically significant predictions of Y can be obtained by including lagged values of X to the lagged values of Y. The causality test statistic of Granger is an F test, where the null hypothesis states that there is no causality between the analyzed variables, that is, statistical evidence is required to conclude that there is causality, rejecting the null hypothesis. The test involves the estimation of the following autoregressive vector model:

$$X_t = \sum_{i=1}^{n} \beta_1 X_{t-i} + \sum_{j=1}^{n} \beta_2 Y_{t-j} + u_{1t} \tag{9}$$

$$Y_t = \sum_{i=1}^{n} \beta_3 Y_{t-i} + \sum_{j=1}^{n} \beta_4 X_{t-i} + u_{2t} \tag{10}$$

Equation (9) postulates that current values of X are related to past values of X itself as well as to lagged values of Y; Equation (10), on the other hand, postulated a similar behavior for the variable Y. Nothing prevents the variables X and Y from being represented in the form of growth rates since the economic or financial variables in general are not stationary in their levels. In this work, the logarithmic returns of the variables selected and analyzed were used. When the variables are cointegrated, the Granger causality test must be carried out in another way, incorporating possible long-term effects into a short-term analysis, which is the essence of cointegration. In this case, the set of equations related to the VEC model is used. The error correction mechanism is intrinsic to VEC models and verifies that lagged values of one variable can help explain the present values of another variable Y, even if past changes of Y are irrelevant. The intuition is that if the two variables are cointegrated, then part of the current change in X can be the result of corrective movements in Y so that the long-term equilibrium with the variable X is again reached. Since X and Y have a common tendency, causality must exist in at least one direction. There will be a causal relationship if the coefficient of error in the previous period, that is, μ_{t-1} is significant and / or if the coefficients of the term of each variable as an explanatory variable of the other are significant. It should be noted that the determination of the number of lags is essential in the study of the causality relation. Gujarati (2004) comments that such a study is highly sensitive to the number of lags used. In this work, the criterion used to determine the discrepancy was the exhaustive one, that is, the development of models with all the possible lags within a pre-established limit.

As for the impulse response function, the same as Granger's causality test indicates that there is a precedence relationship between two variables, it tells us nothing about the intensity of this effect, and how that intensity varies for different time horizons . To meet this need, you can use the impulse response function. Hill and Griffiths (2008) comment that studies of impulse response functions are intended to understand the effects of random shocks in time series. The use of impulse response

functions allows us to evaluate the impacts that a shock in a time series has on itself or other series. Basically, such functions help to understand the temporal effect that shocks on the explanatory variables of an autoregressive vector model has on the dependent variable (see Enders (2009)). As an illustration, as shown in Salles and Almeida (2017), let Y_t be a time series described by the following VAR model:

$$Y_t = \rho Y_{t-1} + v_t^y \tag{10}$$

where v_t^y is the residual term. Assuming zero initial value for this series, the effects on this series of a unitary shock at the initial time can be evaluated without additional shocks. In the specific case where $\rho = 1$, that is, a unit root process soon non-stationary, there is an "infinite memory" process, in which the effect of the initial shock would never be dissipated. Making an analogy with physics, this situation can be understood as a disturbance on a ball initially at rest on a frictionless table: the ball will be in motion indefinitely. In cases where $\rho < 1$, the variable will initially feel the effect of the shock, but will return to the null value after a certain period of time. The greater ρ, the longer the time needed to fully dissipate the effect of the initial shock. For the case of bivariate VAR model, we have the following equations:

$$X_t = \sum_{i=1}^{n} \alpha_i X_{t-i} + \sum_{j=1}^{n} \beta_j Y_{t-j} + v_t^x \tag{11}$$

$$Y_t = \sum_{k=1}^{n} \gamma_k X_{t-k} + \sum_{l=1}^{n} \delta_l Y_{t-l} + v_t^y \tag{12}$$

Salles and Almeida (2017) point out that there are two possible shocks, one in each variable. At each shock, two response functions are associated, one in each variable. We have, therefore, a total of four response functions related to the VAR model. This allows one to study the impact of the shock of one variable on the values of the variable itself or on the values of the other variable. In general, the output of impulse response function analysis is a plot that shows, in a main line, the estimated impact for each lag, wrapped by two red dotted lines that consist of the range of a standard deviation up and down. It is accepted that there is a statistically significant impact on a certain period if the interval between the red dotted lines does not contain the zero line at that point.

4. The Data – Sample Used

Copper is traded in several organized markets around the world and in general, in light of the market efficiency hypothesis their prices are related. And in moments of informational inefficiency these markets offer arbitrage operations opportunities. Thus, in general, price differences between the copper markets usually reflect the cost of transport between these markets and some other differential resulting from local supply and demand. As mentioned earlier, major exchanges for refined copper trading are in London, the London Metal Exchange – LME, in Chicago, the Chicago Mercantile Exchange – CME, and in Shanghai, the Shanghai Futures Exchange - SFE. Among these, LME can be considered the main one, given its tradition and large volume traded daily. Thus, the copper price information used in this survey was the price of copper traded on the spot market of LME, the price for immediate delivery of the metal that best represents the physical market situation of the metal, collected on the Bloomberg website. Beside the monthly copper price traded on the spot market in US$ per ton, the data used in this work were: total monthly production of refined copper, in millions of tons; monthly Brent spot price in US$ per barrel; monthly aluminum spot price, in US$ per ton; monthly observable copper stocks, in thousands of tons; exchange rate index, having as basis 100 the month of January 2009; and the monthly global industrial production.

The level of refined copper production is usually considered one of the most relevant variables for the formation of its price. A very high production can lead to an excess of refined copper in the market, which would generate negative pressures on its price. On the other rand, a very low production

can leave the market in deficit and this would cause positive pressures on the copper price. The global production data used in this work was collected in the work from International Copper Study Group (ICSG), an institute that compiles official copper production data from the governments of partner countries and estimates the production of countries that do not have official data. The data correspond to the total volume of refined copper produced monthly, in thousands of tons. The ratio of crude oil to copper is remarkable, and it occurs through the production costs of the metal. The higher the crude oil price, the greater the costs of the following: mining, given the fuel used in the machines and transportation through the use of trucks; refining, energy intensive process; and freight, given the maritime shipping of the metal to the consumer markets, notably China. Thus when the crude oil price increases, a rise in the price of copper is expected as a cost transmission mechanism. The data referring to the crude oil prices used in this research were the monthly average Brent oil price in US$ per barrel traded on spot market in London. This data was collected on the Bloomberg website. The aluminum price is another potentially relevant variable to the copper price since aluminum is a copper substitute metal in several applications. This way a lower aluminum price reduces the demand of copper, since it stimulates its use to the detriment of copper, exerting negative pressure on its price. Consequently, there is a positive relationship. Thus, as in the case of copper, the aluminum price time series of used in this work is the monthly average of the spot price traded in the LME. Refined copper stocks available in the world are also relevant in the analysis the metal price formation. This is because high stocks mean there is too much metal on the market, which should exert downward pressure on prices. Unfortunately, not all the copper stock available in the world is traceable. Stocks along the production chain, that is, in the hands of miners and manufacturers of copper products, for example, are difficult to measure accurately. In addition, strategic stocks of some governments are intentionally undisclosed. An example is the Chinese State Reserve Bureau (SRB), which is known to be active in metal markets such as copper, buying and selling in the Chinese market, when deemed convenient. However, the volume involved and the moment of performance are not disclosed. Thus, a global compilation of 4 of the world's leading stock data was used as the global stock data: copper stocks in the London Stock Exchange (LME), Chicago (CME) and Shanghai (SFE) warehouses; and stocks in so called bonded Chinese warehouses, which means copper available in specific areas where is understood that the metal has not yet entered the country and thus its warehousing is duty free. In this way, the monthly evolution of these copper stocks by locality in thousands of tons was used. Another variable of interest in the analysis of copper price behavior is the exchange rate. The impact of the exchange rate is twofold: purchasing power of the consuming countries and profitability of the producing countries. For a given price of copper in US$, a depreciated exchange rate makes the metal more expensive for import by consuming countries, discouraging consumption and, consequently, exerting downward pressure on the price of copper in dollars. On the other hand, for a given price of copper in US$, a depreciated exchange rate raises the revenue of producing countries that export the metal, stimulating supply and, consequently, exerting downward pressure on copper price in US$. Thus, the exchange rates of 4 countries / regions of extreme relevance for the copper market were: Novo Sol / US $ (Peru), Peso / US $ (Chile), Euro / US $ (Europe) and Yuan / US $ (China). An index was constructed from these four series of exchange rates that assigns equal weights to each of them, normalized to a basis 100 in January 2009. Finally, another relevant variable considered is the global industrial production. As previously mentioned, the demand for copper is closely related to the industrial activity, since, before being used in its final applications, the metal goes through some manufacturing process. Finally, another variable of relevance considered is the global industrial production. As previously mentioned, the demand for copper is closely related to the industrial activity, since, before being used in its final applications, the metal goes through some manufacturing process. Thus, the information on the industrial production used is a global compilation of industrial production weighted by the level of production of each country. This way, a global industrial production index was used, having 100 as basis. The weight of industrial production in each country or region was 64.5% for developed regions and 35.5% for emerging regions. Among the developed regions or countries the weights for the United States, Japan, Eurozone and other countries

were 21.6%, 10.3%, 18.8% and 13.7%, respectively. Meanwhile, among the emerging regions or countries, the weights for the Asian, Central and Eastern Europe, Latin American and African emerging countries, including the Middle Eastern countries were 18.2%, 2.7%, 6.6 % and 8.0%, respectively. This way the monthly series of global industrial production was used in this work.

Table 1, below, presents a statistical summary of the time series used in order to characterize the data used in this research. In addition to the statistical summary, this table presents the normality and stationarity tests of the time series used in this research. It can be observed that for usual levels of significance, up to 10%, none of the series can be considered stationary. Since the assumption of stationarity is fundamental for the continuation of the analysis, it is usual to make a transformation of the data, in order to make the series stationary. Thus, the logarithmic returns of the original variable X_t time series of were calculated according to the following formula:

$$Return\ X_t = \ln\left(\frac{X_t}{X_{t-1}}\right) \tag{13}$$

Table 1. Statistical Summary of Time Series Used

Variable Selected	Copper Price	Copper Production	Brent Oil Price	Aluminum Price	Copper Stocks	Exchange Rate	Industrial Production
Mean	7,027.18	1,706.91	87.18	1,949.10	1,033.56	93.18	117.55
Median	7,154.24	1,691.00	99.63	1,928.52	1018,01	91.05	118.99
Maximum	9,880.94	2,024.20	124.93	2,667.42	1,611.10	109,19	128.57
Minimum	3,260.36	1,395.70	37.72	1,338.06	667.13	86.94	98.76
Std Deviation	1,405.02	148.38	24.73	300.82	204.08	5.60	8.45
Skewness	-0.4294	0.2007	-0.4977	0.2546	0.6340	1.2572	-0.6765
Kurtosis	3.2595	2.0947	1.8655	2.7693	3.2573	3.6610	2.5525
JB test	2.7989	3.4326	7.9728	1.0936	5.8591	23.6552	7.1086
(p-value)	(0.2467)	(0.1797)	(0.0186)	(0.5788)	(0.0534)	(0.0000)	(0.0286)
ADF	-2.8744	-1.4158	-1.1466	-2.2839	-2.3318	-0.6674	-1.5350
(p-value)	(0.1762)	(0.8462)	(0.9141)	(0.4377)	(0.4116)	(0.9718)	(0.8137)
lags	1	24	1	0	12	2	0

Table 2. Statistical Summary of Variations Time Series Variables Used

Variable Selected	Copper Price	Copper Production	Brent Oil Price	Aluminum Price	Copper Stocks	Exchange Rate	Industrial Production
Mean	0.0042	0.0038	-0.0021	0.0006	0.0040	0.0011	0.0031
Median	0.0035	0.0066	0.0036	-0.0078	-0.0062	0.0007	0.0028
Maximum	0.1667	0.1148	0.1866	0.1431	0.2359	0.0277	0.0153
Minimum	-0.1218	-0.1089	-0.2599	-0.1238	-0.1361	-0.0234	-0.0077
Std Deviation	0.0561	0.0399	0.0780	0.0491	0.0720	0.0115	0.0041
Skewness	0.2533	-0.0485	-0.6483	0.3277	0.3920	0.1451	0.0954
Kurtosis	3.4570	4.0728	4.3420	3.0256	2.9487	2.6216	3.2202
JB test	1.6096	4.0128	12.0434	1.4875	2.1346	0.7864	0.2936
(p-value)	(0.4472)	(0.1308)	(0.0024)	(0.4753)	(0.3439)	(0.6749)	(0.8635)
ADF	-7.3778	-5.3953	-7.7961	-8.6813	-5.7280	-7.0027	-8.8367
(p-value)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
lags	0	0	0	0	0	0	0

Table 2, below, shows a statistical summary of returns or variations time series of the original information used in this work. It can be seen in Table 2 that the Jarque-Bera statistics and the probabilities associated with p-values indicate that the null hypothesis of normal distribution is rejected for all new series, at usual significance levels. As regards the stationarity tests, as expected, the use of the logarithmic returns of the series favored them to approach this property. All series are considered to be stationary, given the p value of the tests.

5. Analysis of Results Obtained

The first methodological procedure used in the analysis was the Engle-Granger cointegration test. Since all series of returns are stationary, the use of the Engle-Granger's cointegration test with these series will erroneously lead to the conclusion that eventually all are cointegrated with the copper price returns time series. Stationary series oscillate over time around an approximately constant mean, that is, it has no tendency and the variance is also approximately constant. The difference between two series with these properties will probably provide a third series with these same properties, thus characterizing, by definition, the existence of cointegration. This means that the Engle-Granger cointegration test must be applied to pairs of non-stationary series. Thus, the original time series not their logarithmic returns were used. The implications of the cointegration tests involving the original time series are also valid for the models that involve the logarithmic returns of these same time series, whose models will consider this relation existing in the context of the original time series. Table 8 below consolidates the results of the cointegration test for each series considered in the analysis, with the copper price time series written as Y_t.

Table 3. Cointegration Test Results

Dependent Variable	Copper Price			Variable X_t		
Variable X_t	ADF Test	(p-value)	lags	ADF Test	(p-value)	lags
Copper Prodution	-2.9077	0.1452	1	-0.3247	0.9745	11
Brent Price	-2.7422	0.1956	1	-1.4526	0.7813	2
Aluminum Price	-2.9082	0.1449	0	-2.5790	0.2555	0
Copper Stocks	-2.4427	0.3131	1	-2.8681	0.1564	2
Exchange Rate	-2.6980	0.2109	1	-0.6989	0.9446	1
Industrial Prodution	-1.9933	0.5346	1	-1.0830	0.8858	4

Table 3 shows that for usual significance levels, such as 10%, no time series should be considered cointegrated with the copper price. However, several of them have a slightly higher p value than this limit, thus making their complete rejection somewhat rigid. Given the sample size, complexity / market imperfections and even possibly data imperfection, reevaluating such conclusions would be adequate. It is important to keep in mind that a limit of 15%, for example, would lead to the conclusion that copper production series and aluminum price are cointegrated to the copper price time series. Raising this limit to 20% would add the Brent price time series and stocks to that list. Finally, a significance level of 25% would allow the conclusion that the exchange rate time series is also cointegrated with the copper price time series. As a conclusion, none of the time series present sufficient evidence to assert with certainty that there is cointegration with the copper price time series. At the same time, the evidence presented by the copper production, Brent price, aluminum price, stocks and exchange rate time series are not sufficiently weak to completely rule out the existence of

cointegration with the copper price time series. Only the industrial production time series seems to allow one to believe with greater conviction that it is not cointegrated with the copper prices series.

This analysis is valid for the purpose of an initial evaluation of the results obtained. However, for the following analysis of Granger's causality, an objective decision regarding the existence or non-existence of cointegration is necessary. This is due to the fact that the existence of cointegration requires the development of VEC models, while the lack of it allows the development of VAR models. In this case, the existence of cointegration is usually disregarded for situations in which the p value of the Engle-Granger cointegration test is greater than 10%. Assuming the existence of cointegration in a test whose p value is 20%, for example, means accepting a 20% probability that this assumption is wrong, which is a high value. Therefore, for the purposes of the Granger causality test, all time series analyzed are considered as non-integrated to the copper price series. Thus, there is no need to develop VEC models, and it is possible to use simply VAR models.

One of the requirements for the development of VAR models is that the time series involved are stationary. Since all series of logarithmic returns of the selected variables are stationary, they will be used without any problems. An important definition for the development of VAR models is the number of lags of the endogenous variables used. There are methods or criteria for defining the number of these lags, one of which is Akaike's criterion. However, in order to avoid the risk of disregarding any relevant models not indicated by this criterion, it was decided to develop models with all lags up a 12 months limit. Granger's causality tests were made for all lags up to this 12 month limit and the one that provided the lowest p value was chosen, that is, the one that presents the most statistical evidence of the existence of causality in the Granger sense. The Granger causality tests results are summarized in Table 4 below.

Table 4. Causality Test Results

Lags	Copper Prodution F (p value)	Crude Oil Price F (p value)	Aluminum Price F (p value)	Copper Inventorie F (p value)	Exchange Rate F (p value)	Industrial Prodution F (p value)
1	0.153 (0.696)	0.100 (0.752)	0.004 (0.949)	0.591 (0.444)	0.885 (0.350)	0.442 (0.508)
2	0.351 (0.705)	0.156 (0.856)	2.337 (0.104)	0.796 (0.455)	1.443 (0.243)	0.847 (0.433)
3	0.065 (0.978)	0.243 (0.866)	2.102 (0.107)	1.000 (0.398)	1.811 (0.153)	1.188 (0.321)
4	0.348 (0.844)	0.498 (0.797)	2.139 (0.085)	1.448 (0.227)	1.106 (0.361)	1.372 (0.253)
5	0.428 (0.827)	0.584 (0.712)	3.373 (0.009)	1.049 (0.397)	0.731 (0.603)	1.196 (0.321)
6	0.393 (0.881)	0.526 (0.786)	3.017 (0.012)	0.995 (0.436)	1.217 (0.309)	2.050 (0.072)
7	0.323 (0.941)	0.325 (0.940)	2.199 (0.047)	0.700 (0.672)	1.201 (0.316)	2.655 (0.018)
8	0.284 (0.969)	0.321 (0.955)	2.247 (0.037)	0.558 (0.808)	1.019 (0.432)	2.561 (0.018)
9	0.312 (0.968)	0.278 (0.978)	1.897 (0.072)	0.446 (0.904)	1.021 (0.435)	2.339 (0.026)
10	0.308 (0.976)	0.594 (0.812)	1.592 (0.135)	0.459 (0.909)	0.959 (0.489)	2.059 (0.045)
11	0.307 (0.981)	0.707 (0.726)	1.615 (0.124)	0.404 (0.948)	0.933 (0.517)	1.920 (0.060)
12	0.259 (0.993)	0.806 (0.643)	1.704 (0.097)	0.352 (0.973)	0.995 (0.468)	1.817 (0.074)

The values highlighted in Table 4 refer to the lag chosen for each series, which corresponds to the lowest p value. The exception is the production series, whose lowest p-value was observed with a 1 month lag, however, the 2 month lag was selected because, despite having a marginally higher p value, it contains a considerably higher F statistic. Curiously, the test indicates that there is no evidence to believe that copper production causes, in the Granger sense, the copper price for any of the lags tested. Among the possible explanations for this result, two stand out. The first is that there is a causal relationship. However, the copper price responds immediately to changes in the level of production.

Since in the present analysis a monthly average price was used and the existing production is a monthly data, a response time considerably less than one month is not properly captured. The fact that the p values of the models with lower lags, 1 and 2 months, are considerably lower than all others, even if they are still quite high, reinforces this hypothesis that the response time of copper price to production is small. Another possibility is that copper production does not present a temporal precedence relation with the price of the metal. Some analysts argue that it is the demand that is a driver for metal prices, not supply, the offer reacts according to the price and not the other way around. In order to evaluate the reasonableness of this assertion, the VAR model whose dependent variable is copper production was analyzed. Here, lags of up to 18 months were considered, since the reaction time of producers at the price is expected to be higher. The result obtained reinforces this second hypothesis. Figure 1, below, shows the impulse response function of this VAR model, and shows that there is a statistically significant positive impact of the copper price on the metal production after 8 months, period t constitutes a lag of t-1.
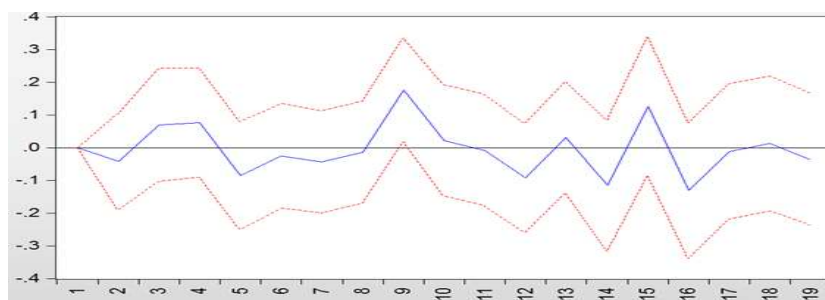

Figure 1. IRF – Copper Price on Copper Production

This 8 month period seems insufficient to indicate the time between the decision to invest in a new mine and the first copper extraction. However, it may be the period of time necessary for producers to raise their level of production by investing in improvements in the production process, that is, efficiency gains, increased working hours at the plant, increased capacity utilization and resumption of production in demining mines.

Regarding crude oil, the Brent price also showed no evidence that Granger cause the copper price. The expectation was that this causal relationship existed through production costs: when crude oil price rises, copper production costs rise and thus there would be a pass-through of this rise in costs to the final price. This transmission of the elevation of production cost to the final price is common in markets in which there is product differentiation. In the case of copper, the final price is given internationally as a mirror of the balance between supply and demand, that is, the producer has little or no influence on the price. Thus, a possible explanation for the absence of causality between crude oil and copper prices is that this transmission is not necessary or not possible: producers may try to pass on changes in their costs to the final price, but they can not because what determines the commodity price are the dynamics of supply and demand.

As for aluminum, Granger's causal relationship was identified for several lags, the most notable being that at 5 months. This temporal analysis will be evaluated in greater detail later, in the impulse response analysis. At this point, we can see that evidence of Granger's causal relationship between aluminum and copper prices is very strong: p value less than 1%.

In relation to stocks, the model that presented the best result was the one with 4 months lag. In this model, a p value of 22.7% was verified, which is a value that indicates that the relation with the price of copper is not negligible, but it is also not enough for the study to demonstrate with confidence that such a relationship exists. This suggests that the data used is not fully representative of global stocks. In fact, the ICSG estimates copper stocks on the London Metal Exchange (LME), Chicago

Mercantile Exchange (CME), Shanghai Futures Exchange (SHFE) and bonded Chinese warehouses to consist of approximately 60% of global copper stocks. The difficulty in accessing stocks outside these locations obfuscates the analysis of this variable.

Granger's causality of the exchange rate presented a p value of 15.3% for the 3 month lag model. Similar to the case of stocks, this value indicates that there is evidence that there is a causal relationship between the exchange rate and the price of copper, although not strong enough for this to be safely stated p values up to 10%. This result alone does not detract from Granger's causal relationship, however, it is valid to conclude that the inclusion of other relevant currencies in the exchange rate could possibly result in the greater acceptance of Granger's causality. It is noteworthy to observe that the construction of this simple exchange rate, containing only 4 coins with equal weights, presented a result of such relevance.

Finally, industrial production presented strong evidence that Granger-cause copper prices, with a value of 1.8% for the VAR model with a 7 month delay. This relationship was widely expected since industrial production is a good proxy for copper demand.

The interpretation of these VAR models is facilitated and complemented by impulse-response function analysis. While Granger's causality analysis only indicates whether or not there is a precedence relationship between the various variables under analysis and copper price, the impulse-response function analysis allows one to measure the intensity of the impact of one variable on the other, and how this intensity varies over time.

For all the variables under analysis, a period of 12 months was considered. In impulse response function plots, the solid blue line indicates the estimated impact on each time period, while dotted red lines consist of a standard deviation above and below this value. In general, the existence of the estimated impact in each period is accepted when the interval between the red dotted lines does not contain the zero line, that is, when the existence of that impact is accepted for variations of up to one deviation for each side. The value of each point represents the impact, in logarithmic return units, on the response variable in all cases, copper price caused by a change of one unit in the return of the impulse variable. In addition, the plots show that the period t corresponds to a lag 1 of t-1; this is because the instant named 1 in the plot corresponds in reality to the instant 0. Therefore, the plots contain 13 points, not 12.

Starting with copper production, Figure 2 below shows that there is no relevant impact on the copper price for any period considered, which is in line with the results suggested by the Granger causality analysis and the estimated VAR model, which does not attribute statistical relevance to the copper production variable, with any lag, as an explanatory factor for the price of the metal.
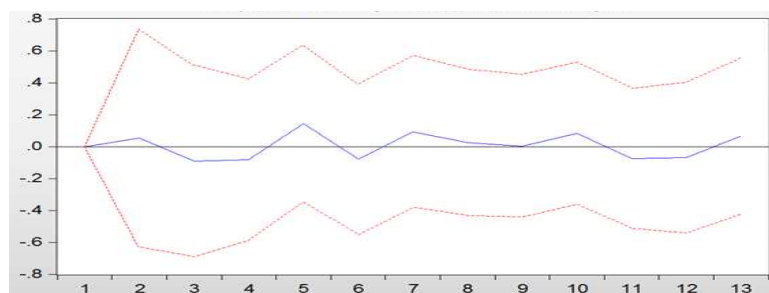

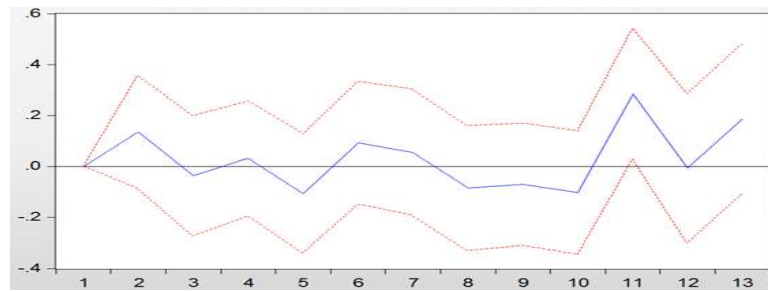**Figure 2. IRF – Copper Production on Copper Price.**

Figure 3. IRF – Brent Price on Copper Price

The impulse response function of the Brent price on the copper price, however, provided a different result than the one suggested by Granger's causality analysis. While the causality analysis suggested that there is no relevant relationship, the impulse response function plot indicates that there is a positive impact on the copper price with a 10 month lag, as in Figure 3, above. On the other hand, this is in line with the estimated VAR model in which the only statistically significant coefficient was the Brent price with a 10 month lag: with a p value of 0.5%. This suggests that, contrary to what Granger's causality test indicated, there is rather a positive effect of crude oil price on the copper price, that is, a rise in the Brent price generates a rise in copper price after 10 months . Possibly, it is the period of time necessary for the transmission of the change from cost to price, whose existence was denied by Granger's causality test. It is also worth noting that both the impulse-response function analysis and the previously estimated VAR model indicate a transmission of 29%, that is, a unit of variation in Brent's logarithmic return generates a variation of 0.29 unit 10 months later in the logarithmic return of the copper price.

Regarding the aluminum price, the analysis of the impulse response function indicates that there is a statistically significant impact on copper price after 5 months, which is in agreement with the result of the Granger causality analysis. However, Figure 4 and the previously estimated VAR model indicate that this effect is negative, contrary to what was initially expected. As mentioned previously, a positive relationship was expected due to the possible substitution effect between copper and aluminum in some applications: when the price of aluminum rises, a substitution movement is expected for copper, thus raising the price of the red metal. However, the impulse response analysis and the estimated VAR model indicate that the relationship is otherwise. Possibly, what defines producer decision-making between using aluminum or copper is not the price but the particular physicochemical properties of each metal. Although used in some similar applications, they are not perfect substitutes. An example of this is the energy grid: although both metals are used in their construction, each one is more advantageous for a certain application: UHV (ultrahigh voltage) transmission lines are aluminum intensive, while copper is more used in networks of distribution. Thus, positive changes in aluminum price would suggest, in this case, that energy grid investment is more focused on UHV transmission lines and, consequently, less focused on distribution networks. This would imply a lower demand for copper in the period ahead, which would exert negative pressure on its prices. Thus, positive changes in aluminum price would suggest, in this case, that energy grid investment is more focused on UHV transmission lines and, consequently, less focused on distribution networks. This would imply a lower demand for copper in the period ahead, which would exert negative pressure on its prices. Thus positive changes in aluminum price would suggest, in this case, that enegy grid investment is more focused on UHV transmission lines and, consequently, less focused on distribution networks. This would imply a lower demand for copper in the period ahead, which would exert negative pressure on its prices. It is also worth noting that the impulse response function analysis and the previously estimated VAR model agree that this negative effect of aluminum price on copper price after 5 months is approximately -0.5, that is, a positive change in one unit in the return of the aluminum subsequently generates a negative variation of -0.5 in the copper return. However, the VAR model suggests that there are also relevant impacts with other lags; still, the 5 month lag

demonstrated to be the most relevant.

In relation to stocks, the analysis of the impulse response function presents results similar to those of the Granger causality tests, that is, that there is no statistically significant impact on copper price as shown in Figure 5. However, the VAR model indicated that there is a statistically significant effect for lags of 2 and 3 months, the most relevant being lag of 3 months. Its effect is negative, as expected: a positive change in a stock return unit has a negative impact of -0.39 unit on copper price.

As in the Granger causality test, the impulse response function analysis suggests that there is no relevant impact of the exchange rate on copper price for the level of significance adopted, as shown in Figure 6. However, the VAR model suggests that the exchange rate has a significant impact on copper price with 3 month lag. The negative coefficient of -1.18 shows that this ratio is inverse as expected: when the exchange rate increases, it means that the dollar strengthened against the currencies participating in the index, a dollar becomes worth more units of those currencies. Depreciated currencies raise the profitability of copper-producing countries that export the metal and make copper more expensive for metal-importing countries, that is, it stimulates supply and shrinks demand, thereby exerting negative pressure on copper prices.
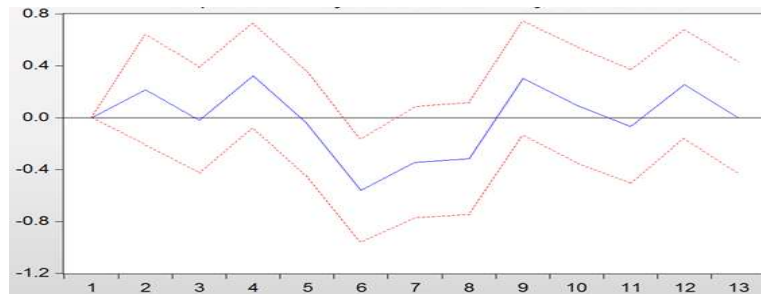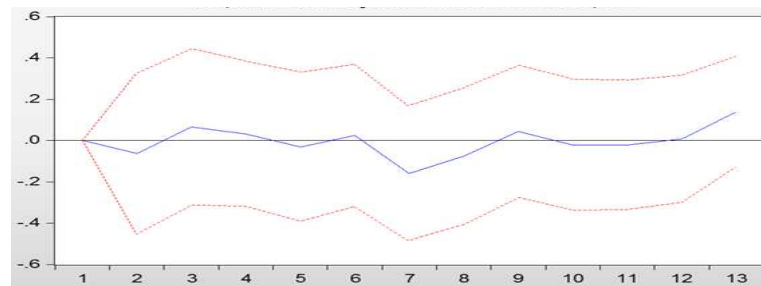

Figure 4. IRF – Aluminum Price on Copper Price

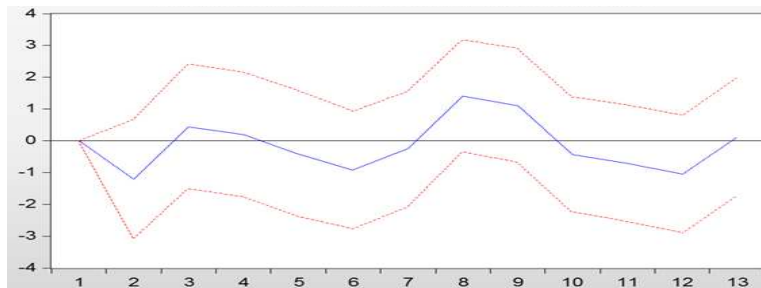**Figure 5. IRF – Copper Stocks on Copper Price**

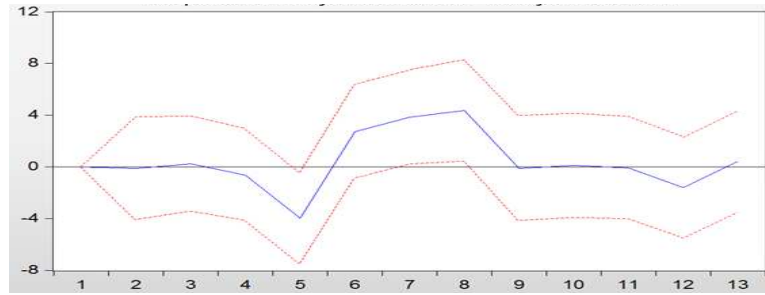**Figure 6. IRF – Exchange Rate on Copper Price**

Figure 7. IRF – Industrial Production on Copper Price

Finally, the analysis of the impulse-response function for industrial production resembles that of the Granger causality test, when the impact of greater relevance on copper prices occurs with 7 months of lag, as shown in Figure 7. As expected, this is a positive impact: increasing global industrial production suggests higher manufacturing of copper products, components and wires, hence a greater demand for refined metal, which would exert positive pressures on copper price. In addition, the impulse-response function presents results similar to those of the VAR model on the presence of other relevant lags: at 4 month lag, a negative initial impact on the copper price occurs, and then, with 6 month lag, there is a positive impact preceding the main effect, at the 7 month lag. It is worth noting the great intensity estimated for the impact of industrial production on the price of copper after 7 months: approximately 4 times higher than the variation in industrial production.

The previously performed analyzes indicated which variables are individually relevant for the formation of copper price, in what time horizon and with what intensity. It is interesting to evaluate how all the relevant variables identified jointly influence copper price. Relationships found between pairs of variables are possibly lost or altered when doing a global analysis involving several of them simultaneously. For this purpose, a multiple regression model was developed in which copper price is the dependent variable and the variables identified as relevant in the previous analyzes are the explanatory variables. In addition to allowing the joint effect of the relevant variables to be evaluated, the multiple regression model allows the determination of how much copper price is explained by these variables through a determination coefficient.

The starting point of the regression model was the results obtained previously in the analysis of the estimated VAR models. The first or preliminary estimated model considers the variables identified as relevant in the developed VAR models, that is, those that presented p value lower or very close to 10%, totaling 20 variables. Table 5 presents the main evaluation measures of the preliminary regression model obtained, while Table 6 presents the estimates of the parameters or coefficients of each model variable of the preliminary regression model. Table 5 shows the preliminary model that a moderate explanation coefficient of 43% was obtained. However, the adjusted explanation coefficient of only 20% indicates that the sample should be increased and / or the number of estimated parameters reduced. Table 6 confirms this by showing that some variables have lower statistical relevance with aggregation, which is indicated by the presence of high p values for the preliminary model.

Table 5. Regression Models Performance Measures

Preliminary Model		Final Model	
R-squared	0.4292	R-squared	0.3705
Adjusted R-squared	0.2009	Adjusted R-squared	0.2656
S.E. of Regression	0.0497	S.E. of Regression	0.0477
Sum Squared Resid	0.1236	Sum Squared Resid	0.1363
Akaike Criterion	- 2.9237	Akaike Criterion	- 3.1075
Schwarz Criterion	- 2.6576	Schwarz Criterion	- 2.7570

Thus, since the greatest interest lies in identifying the relevant variables, this model was modified so as to consider only statistically relevant variables, according to the criterion of p value lower or close to 10%. The procedure adopted consisted in performing successive regressions, always removing the variable with higher p value with each new or final regression model, until the first or preliminary model was reached in which all the variables met the stipulated criterion. The tables 5 and 7 present the values referring to the final model thus obtained.

Table 5 shows that, although the coefficient of determination was reduced, the adjusted coefficient of determination increased, which indicates that the explanatory power of the model, adjusted for the quantity of inputs, improved. The other measures of model evaluation also indicate that the final model is better than the first one: the standard error was reduced and the Akaike and Schwarz criteria. The exception was due to the sum of square residuals which increased. This, however, was expected since the reduction of R^2 necessarily implies that there was an increase in the sum of square residues. That is, both should be understood as one criterion. These results suggest that, in fact, the last estimated model or the final model should be maintained.

Table 6. Preliminary Regression Results - Estimates of Parameters

	Paremeter	Estimate	Stat t	p value
1	Constant term	-0.0224	-1.7116	0.0932
2	Brent Price $_{t-5}$	-0.0835	-0,7520	0.4556
3	Brent Price $_{t-10}$	-0.0577	-0.5786	0,5655
4	Aluminum Price $_{t-1}$	-0.4527	-1.9081	0.0621
5	Aluminum Price $_{t-3}$	-0.0189	0.0858	0.9320
6	Aluminum Price $_{t-5}$	0.0588	0.2575	0.7978
7	Aluminum Price $_{t-6}$	0.3940	1.8132	0.0758
8	Aluminum Price $_{t-7}$	0.3181	1.5131	0.1365
9	Aluminum Price $_{t-11}$	-0.5007	-2.3161	0.0247
10	Stocks $_{t-2}$	0.1337	0.9877	0.3280
11	Stocks $_{t-3}$	0.0257	0.2002	0.8412
12	Industrial Production $_{t-4}$	3.4682	1.5528	0.1268
13	Industrial Production $_{t-5}$	2.8223	1.4088	0.1651
14	Industrial Production $_{t-6}$	1.2735	0.6556	0.5151
15	Industrial Production $_{t-7}$	1.7080	0.8807	0.3827
16	CopperPrice $_{t-1}$	-0.4829	2.4905	0.0161
17	Copper Price $_{t-3}$	-0.1927	-0.8640	0.3917
18	Copper Price $_{t-5}$	0.1588	0.7048	0.4842
19	Copper Price $_{t-6}$	-0.6923	-3.1423	0.0028
20	Copper Price $_{t-7}$	-0.0435	-0.2069	0.8370
21	Copper Price $_{t-11}$	0.3554	1.5463	0.1283

Table 7. Final Regression Results - Estimates of Parameters

	Paremeter	Estimate	Stat t	p value
1	Constant term	-0.0219	-2.2385	0.0289
2	Aluminum Price $_{t-1}$	-0.4501	-2.2557	0.0277
3	Aluminum Price $_{t-6}$	0.2871	1.5960	0.1157
4	Aluminum Price $_{t-7}$	0.2393	1.8955	0.0628
5	Aluminum Price $_{t-11}$	-0.2823	-1.9418	0.0569
6	Stocks $_{t-2}$	0.1453	1.7028	0.0938
7	Industrial Production $_{t-4}$	4.6688	2.3689	0.0211
8	Industrial Production $_{t-5}$	4.3138	2.8536	0.0059
9	Copper Price $_{t-1}$	0.5160	3.0085	0.0038
10	Copper Price $_{t-3}$	-0.2254	-1.7065	0.0931
11	Copper Price $_{t-6}$	-0.5302	-2.8372	0.0062

Table 7 presents the parameters of the final model and its respective p values. Comparing the return of the copper price estimated by the model with the observed real values in many situations is a good way of verify whether there were price rise or fall, once it is a price variable, the market makers can have future expectations. For this type of situation, an analysis of the 71 points of the model shows that this model is extremely satisfactory: the model hits the side, that is, if the return is positive or negative, in 52 points, that is, in 73% of the cases, a high hit rate.

Analyzing the final regression model obtained, it is worth noting that industrial production is the variable with the greatest impact on copper price as the extremely high level of its coefficients was obtained. This result is consistent with all the previous analyzes, which indicate that industrial production has a strong positive impact on copper prices with a few months of lag. More than that, industrial production coefficients that are much larger than all others suggest that it is the predominant variable that "dictates" the copper price movements while the performance of the others functions as "adjustment", that is, softening or amplifying movements fundamentally generated in response to industrial production. One thing that reinforces this analysis is the fact that industrial production has shown consistency in terms of its relevance in all the various analyzes. The analysis of the aggregate variables, through the regression model, did not alter the relations observed in the analyzes conducted individually with the industrial production. The same can not be said for the other variables. The Brent oil price, for example, was relevant in the developed VAR model and the impulse response function analysis, however it was discarded in the regression model. On the other hand, the aluminum price was also relevant in all analyzes, but the estimate of its impact varied significantly among them. Signals of the parameters estimates of the VAR model, for example, were reversed in the regression model, indicating an effect opposite to that previously suggested. That is, the inclusion of variables makes the measurement of the effects of aluminum price on copper price more "nebulous", different from that of industrial production. In any case, it should be noted that aluminum price was considered relevant in all analyzes. The lagged copper price itself was also relevant in the regression model.

6. Conclusion and Final Comments

The main objective of this work was to identify variables relevant to the formation of copper prices in the international market. Thus, candidate variables were selected and several tests and analyzes were carried out in order to verify the existence of relevance or statistical significance of each of these variables. In order to achieve the objectives of this work, we performed cointegration tests,

constructed autoregressive vector models, analyzed impulse response functions and estimated linear regression models. Among the relevant variables for the international copper price, industrial production was the most relevant variable, presenting evidence in all the tests carried out that its impact on the copper price is significant. The aluminum price and the lagged copper itself were also relevant in the various analyzes carried out, although their respective impacts are much weaker and more irregular than the global industrial production. The variables stock, crude oil price and exchange rate were relevant in some analyzes, but in others this did not occur. Thus, it is not possible to say with certainty that these variables are determinant in the copper price, although they should not be discarded in other analyzes or future research. Given the hypotheses verified, the copper production variable did not present any evidence to infer that this variable is relevant for the pricing of copper in the international market.

It should be noted that the results obtained here are linked to the sample used, and therefore, the time series of the stock and exchange rate variables, which were obtained with limitations, have results that should be viewed with caveats. Regarding the copper production variable, the result obtained indicates that this variable is caused by the price while the reciprocal does not happen. Although this research did not have the main purpose of constructing a regression model to explain the copper price, the estimated model presented a result regarding the direction of copper price variation that deserves to be highlighted: the model showed the direction of the movement in copper prices in 73% of cases. This result can be very useful for market participants, for example in the decision-making process of hedge strategies by producers and in the determination of speculative operations with copper prices. The statistical significance of the estimated parameters of the model and the low coefficient of determination obtained indicate that, although relevant, the variables selected are not able to fully explain the fluctuations in copper price. This suggests the existence of other variables not considered in this work that are also relevant for determining the copper price in the international market. Thus other variables can be considered in future work on the theme, such as interest rates, level of scrap use, stock exchange performance and inflation rates.

The results obtained in this work can be used as a basis for the development of other studies that seek to study copper price behavior and its perspectives. In addition to verifying if other variables are relevant for copper pricing, other methodologies and samples that may contribute to the clarification of the topic discussed here should be verified in future works.

References

Cerda, R. (2005) Market power and primary commodity prices: the case of copper. Department of Economics, Pontificia Universidad Católica de Chile.

Enders, W. (2009) Applied Econometric Time Series. 3th Ed, John Wiley & Sons, New York.

Engle, R. F. and Granger, C. W. (1987) Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, **55**, 251-276. http://dx.doi.org/10.2307/1913236

García-Cicco, J. and Montero, R. (2011) Modeling copper price: a regime-switching approach. Documento de Trabajo No. 32 de la Escuela de Economía "Francisco Valsecchi" de la Facultad de Ciencias Económicas de la Universidad Católica Argentina.

Gujarati, D. (2004) Basic Econometrics. 4th Ed, McGraw-Hill Companies, New York.

Hill, C. R and Griffiths, E.W. (2008) Principles of Econometrics. 4th Ed., John Wiley & Sons, New York.

International Copper Study Group (2014) The World Copper Factbook 2014. Lisbon.

Maddala, G. (1992) Introduction to Econometrics. 2nd Edition, Macmillan Publishing Company, New York.

Salles, A. A., Almeida, P. H. A. (2017) The Crude Oil Price Influence on the Brazilian Industrial Production. *Open Journal of Business and Management*, **5**, pp. 401-414.

Stürmer (a), M. (2013) 150 years of boom and bust: what drives mineral commodity prices? Institute for International Economic Policy (IIW) – University of Bonn.

Stürmer (b), M. (2013) Industrialization and the Demand for Mineral Commodities. Bonn Graduate School of Economics, Department of Economics, University of Bonn.

Zhang, H., Dufour, J. and Galbraith, J. (2015) Exchange rates and commodity prices: measuring causality at multiple horizons. CIRANO - Scientific Publications 2013s-39.