• FINANSE I PRAWO FINANSOWE •
• JOURNAL OF FINANCE AND FINANCIAL LAW •

Wrzesień/ September 20025 • vol. 3(47): 129–157

COPE Member since 2021 JM15102

Funding information: Not applicable. **Conflicts of interests:** None. **Ethical considerations:** The Author assures of no violations of publication ethics and takes full responsibility for the content of the publication.

GPT MODELS OR ECONOMETRIC MODELS: A COMPARATIVE ANALYSIS OF GOLD PRICE DETERMINANTS

Jan Krawczyk*

https://doi.org/10.18778/2391-6478.3.47.08

GPT MODELS OR ECONOMETRIC MODELS: A COMPARATIVE ANALYSIS OF GOLD PRICE DETERMINANTS

ABSTRACT

The purpose of the article. The main objective of this article is to compare the results of data analysis regarding gold prices and their determinants using two approaches: a classical econometric model and Microsoft Copilot, which integrates advanced artificial intelligence technologies, including the GPT-4 language model (Generative Pre-trained Transformer 4). The secondary objective is to identify, based on the existing literature, the main factors influencing fluctuations in gold prices. These include: the price of crude oil, the USD/EUR exchange rate, the S&P 500 index, and the Consumer Price Index (CPI) in the United States.

Methodology. The empirical study involves determining the descriptive statistics of the analyzed variables, the correlation matrix, and estimating the structural parameters of the model explaining the gold price.

Results of the research. The best results were obtained for the logarithmic returns of the analyzed variables. In line with the stated hypotheses, there is a negative relationship between the gold price and changes in the S&P 500 index, a negative relationship between the gold price and changes in the US$/EUR exchange rate, and a positive relationship between the gold price and the CPI. The study shows that, during the analyzed period (02.2004 –11.2023), changes in crude oil prices did not have a statistically significant impact on gold price changes. To obtain data analysis results using Microsoft Copilot, a "chat" session was conducted. The responses provided the following information: proposed determinants of gold prices, a list of scientific articles, and R code to perform the auto.arima procedure. A comparison was made between the model incorporating economic theory-based factors and the model from the auto.arima procedure suggested by Microsoft Copilot. Based on the conducted study, it can be concluded that the model incorporating both autoregressive factors and other gold price determinants better explains the analyzed variable.

Keywords: gold price, econometric model, gold price determinants, GPT.

JLL Class: C53, C58.

* Graduate of the Technical High School of Łodz, Wroclaw University of Science and Technology, participant in "Talented at School, Brilliant at the University" programme at the University of Lodz, 7th edition, https://orcid.org/0009-0008-9424-9865
E-mail: jankraw9@gmail.com

Introduction

In the 1930s, econometric models began to be used for the quantitative analysis of economic phenomena. Econometrics is a field of social science that combines statistics, mathematics, and economics. The 20th and 21st centuries have seen a dynamic development of econometric methods and a growing interest among economists in quantitative analyses. The use of appropriate tools for modeling economic variables significantly reduces forecasting errors, and estimating the structural parameters of economic models answers the question of how economies actually function. Additionally, statistical tests help economists assess whether there is a statistically significant relationship between the analyzed variables. Econometric models are now widely used for analyses in both businesses and corporations, as well as in government administrative units. They are applied in both microeconomic and macroeconomic analyses (e.g., to study the impact of government policy on the economy).

The 21st century has also seen the development of tools in the field of artificial intelligence, which utilize advanced machine learning methods, leading to significant changes in the global economy. Artificial intelligence (AI) and machine learning are now key technologies influencing changes across many sectors of the economy, from finance to manufacturing, retail, healthcare, and many others. New opportunities are opening up for businesses to improve operations, mainly through cost reduction and minimizing human error.

The purpose of this paper is twofold. First, it aims to compare the results of data analysis related to gold prices and their determinants using two approaches: classical econometric modeling and Microsoft Copilot, a tool that incorporates artificial intelligence technologies including the GPT-4 language model developed by OpenAI. The second objective is to identify, based on a review of the literature, the key economic factors influencing gold price fluctuations. These include the price of crude oil, the USD/EUR exchange rate, the S&P 500 index, and the Consumer Price Index (CPI) in the United States.

Importance of Gold in the Economy

Gold is a precious metal that can be distinguished by three main macroeconomic functions:

- monetary function,
- industrial function,
- investment function.

Monetary Function

The monetary function of gold occurs when it is used as a medium of exchange, a unit of account, or a store of value. The gold standard is a system in which the value of currency is directly linked to gold. In practice, this means that for each unit of currency, a specified amount of gold can be obtained. This system was popular at the turn of the 19th and 20th centuries until World War I. One of the main advantages of the gold standard was its stability. Prices did not change too drastically, providing market participants with a sense of certainty. This system was based on three fundamental principles: the government setting the gold parity, the convertibility of currency into gold, and the strict link between the issuance of money and gold reserves. However, the system also had its drawbacks. Due to the fixed exchange rate, it hindered adjustments to changes in the competitiveness of individual countries. Additionally, the gold standard exacerbated the effects of the Great Depression. Countries that abandoned the gold standard and devalued their currencies recovered from the depression more quickly.

Industrial Function

In electronics, gold is an excellent conductor and resistant to corrosion, making it valuable in the production of devices. In dentistry, gold is durable and biocompatible, making it an effective material for dental crowns and bridges. In the space industry, gold is used in spacecraft to protect against infrared radiation and stabilize temperatures. In art and jewelry, gold is prized for its color and malleability, often used for jewelry making and gilding in art. Therefore, changes in gold prices impact production costs and the aggregate level of GDP in the economy.

Investment Function

Gold plays an important role as an investment instrument. Its attractiveness to investors stems from several characteristics. First, gold is a very effective tool for diversifying an investment portfolio. This means that when other instruments (such as stocks or bonds) are not yielding profits, the value of gold can offset losses. Second, gold is perceived as a hedge against inflation. In times of currency devaluation, the price of gold typically rises as investors seek to protect their assets. Gold is also attractive in times of economic uncertainty. During recessions, the price of gold usually increases. During the global financial crisis of 2007–2008, the value of the S&P 500 index fell by more than 50%, while the price of gold rose by several dozen percent (see Figure 1, daily data). However, gold does not always gain value immediately following the outbreak of an unexpected global threat, as market participants may be determined to sell gold for cash. In the initial phase of the COVID-19 pandemic, in March 2020, a sharp decline in gold prices was recorded. When comparing prices from March 2020 to May 2022, it increased by nearly 20% (see Figure 2, daily data). Ongoing military conflicts, such as the Russia–Ukraine war and instability in the Middle East, also lead to significant increases in gold prices. Although on February 24, 2022, when Russia invaded Ukraine, the price of gold slightly declined, a few days later, after sanctions were imposed on Russia (e.g., the exclusion of certain important officials from the SWIFT system), gold spectacularly gained value (see Figure 3, daily data). Therefore, investing in gold requires a thorough analysis of the factors that influence its price.

Figure 1

Gold price (per troy ounce, USD) from 2007 to 2008

Source: Bankier.pl

Figure 2

Gold price (per troy ounce, USD) during the approximate duration of the COVID-19 pandemic

Source: Bankier.pl

Figure 3

Gold price (per troy ounce, USD) during the initial days of the Ukraine war

Source: Bankier.pl

Thus, gold fulfills fundamental macroeconomic functions. The sharp fluctuations in the price of gold are most often related to extreme geopolitical or catastrophic events.

Modeling and Determinants of Gold Price

In the literature, three main approaches can be distinguished for modeling and forecasting gold prices. First, the impact of macroeconomic factors on gold prices can be described using causal econometric models. Bukowski (2016) identified the main determinants of gold prices, which include: the price of crude oil, the US$/EUR exchange rate, the S&P 500 index, and the yield on 10-year U.S. Treasury bonds. In empirical analyses, the first differences of the logarithmic variables were used. The following conclusions were formulated in this study:

- the relationship between the price of gold and the US$/EUR exchange rate is negative,
- the relationship between the price of gold and the price of Brent crude oil is positive,
- the relationship between the price of gold and changes in the S&P 500 index is negative,
- the relationship between the price of gold and the yield on 10-year U.S. Treasury bonds is negative.

The depreciation of the dollar makes gold cheaper for investors outside the USA, which increases demand for gold and raises its price in US dollars (Levin &Wright, 2006). Crude oil has replaced diesel as an energy source for heating during gold mining. Consequently, an increase in oil prices leads to higher extraction costs, resulting in a higher price of gold (Polyus). Conversely, during economically unstable periods, investors may prefer gold (a safe asset) over stocks, which are associated with higher risk (Puci et al., 2022).

The impact of other independent variables on gold prices has also been studied. Choong et al. (2012) included the price of silver among the determinants of gold. They found a positive relationship between the variables, which contradicted their hypothesis that silver is a substitute for gold. The study also confirmed, in line with the hypothesis, a positive relationship between the price of gold and the Consumer Price Index (CPI), which measures inflation and affects consumer behavior.

Second, a popular method for predicting the price of gold is ARIMA models. Guha and Bandyopadhyay (2016) demonstrated that for a monthly sample from 2003 to 2014, among six different time series model forms, the ARIMA(1,1,1) model proved to be the best. They also noted the inefficiency of ARIMA models in the case of sudden, sharp changes in economic conditions. Yang (2019) selected ARIMA(3,2,1) as the best model based on empirical results. Setyowibowo et al. (2021) used a hybrid ARIMA-GARCH model, which is particularly effective for high-frequency data.

Third, rapid advancements in artificial intelligence have enabled quantitative analyses of gold prices using machine learning methods. Makalala and Li (2021) showed that the Support Vector Machine (SVM) method is more effective than ARIMA models. Zhang and Ci (2020) applied Deep Belief Network (DBN) models, comparing them with ARIMA models, traditional BP neural networks, and genetic algorithms for optimizing BP neural networks. The DBN model proved to be the best, achieving the lowest Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE).

ARIMA models and machine learning models require the use of advanced econometric methods and tools, and thus, are not applied in this study. This paper attempts to identify the determinants of gold prices using causal econometric models.

Based on the presented literature, the following model will be analyzed in the empirical part:

$$Gold\_price = f(Crude\_oil\_price, S\&P500, US\$/EUR, CPI)$$

where: $Gold\_price$ – the price of gold (per troy ounce, which corresponds to a weight of 31.1035 grams); $Crude\_oil\_price$ – the price of crude oil (per barrel, USD, averaged from Brent, WTI, and Dubai); $S\&P500$ – the S&P 500 index; $US\$/EUR$ – the US Dollar to Euro exchange rate; CPI – the Consumer Price Index (CPI).

According to the literature and economic theory, the aim of the first part of the study is to verify the following research hypotheses:

1. An increase in the US$/EUR exchange rate causes a decrease in the price of gold.
2. An increase in the S&P 500 index causes a decrease in the price of gold.
3. An increase in the Consumer Price Index (CPI) causes an increase in the price of gold.
4. An increase in the price of crude oil causes an increase in the price of gold.

Methodology

The most common method for estimating econometric models is the method of least squares (OLS). Let's consider a linear regression model as follows (Maddala, 2001):

$$Y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \ldots + \beta_k x_{k,t} + \varepsilon_t \quad (1)$$

where: Y_t – the dependent variable; $x_1, x_2, \ldots, x_k -$ the independent variables; β_0 – the intercept; $\beta_1, \beta_2, \ldots, \beta_k$ – the regression coefficients, the structural parameters of the model; $\varepsilon -$ the error term.

The Ordinary Least Squares (OLS) method for linear regression involves finding the values of coefficients β that minimize the sum of the squared differences between the actual values of the dependent variable and the values predicted by the model. The estimator of the Ordinary Least Squares (OLS) method is defined by the formula (Maddala, 2001):

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (2)$$

The estimator given by equation (2) is BLUE (Best Linear Unbiased Estimator), which means that it is consistent, unbiased, and the most efficient among the class of linear estimators for the vector β – Gauss-Markov Theorem (Greene, 2002). According to the assumptions of the linear regression model, the error term should have the following properties:

1. The expected value of the error term is 0.
2. The variance of the error term is identical for all observations -–homoskedasticity.
3. The covariance between two error terms from different periods is 0 – there is no autocorrelation.
4. The error term has a normal distribution.

Failure to meet the above assumptions can lead to non-spherical error terms. In such cases, the OLS estimator ceases to be BLUE. Therefore, any estimation of econometric model parameters requires verification of the stochastic properties of the error term. For this purpose, the following statistical tests can be used (Verbeek, 2012):

– **Test for normality of residuals**

The Doornik-Hansen test is used to verify the normality of the random component distribution. The set of hypotheses is as follows:

H_0: the random component follows a normal distribution,

H_1: the random component does not follow a normal distribution.

Its statistic is defined by the formula:

$$DH = z_1^2 + z_2^2 \tag{3}$$

where: z_1 – the transformed skewness coefficient; z_2– the transformed kurtosis coefficient. The statistic (3) follows a chi-squared χ^2. If the DH statistic is less than or equal to the critical value, we do not have grounds to reject the null hypothesis.

– **Lagrange multiplier test (LM test)**

The LM test is a method for checking the presence of autocorrelation, which involves correlating current values of the random component with past values. In the first stage, the parameters of model (1) are estimated, and then a model is estimated where the residual component from model (1) is the dependent variable. As explanatory variables, the lagged residuals up to order P are additionally included:

$$\varepsilon_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \ldots + \beta_k x_{k,t} + \beta_{k+1}\varepsilon_{t-1} + \ldots + \beta_{k+P,t}\varepsilon_{t-P} + \eta_t \tag{4}$$

The null hypothesis of the LM test is the absence of autocorrelation up to order P:

$$H_0: \beta_{k+1} = \ldots = \beta_{k+P} = 0 \tag{5}$$

The test statistic is:

$$LM = nR^2 \tag{6}$$

where: n – the sample size in model (4); R^2– the coefficient of determination of model (4).

The LM statistic follows a chi-squared χ^2 distribution with P degrees of freedom. If the LM value is less than or equal to the critical value from the χ^2, we do not have grounds to reject the null hypothesis.

– **White's test**

White's test for heteroskedasticity is a method for checking the presence of heteroskedasticity in the error term. The dependent variable is the squared residuals from model (1):

$$\varepsilon_i^2 = \beta_0 + \beta_1 x_{1,i} + \cdots + \beta_k x_{k,i} + \beta_{k+1} x_{1,i}^2 + \cdots + \beta_{k+k} x_{k,i}^2 + \beta_{k+k+1} x_{1,i} x_{2,i} + \\ \ldots + \beta_{k+k+S} x_{k-1,i} x_{k,i} + \eta_i \tag{7}$$

The null hypothesis of White's test is the homoskedasticity of the error term:

$$H_0: \sigma_i^2 = \sigma^2 \tag{8}$$

and the test statistic is defined by the formula:

$$LM = nR^2 \tag{9}$$

The statistic (9) follows a chi-squared χ^2 with M degrees of freedom (the number of all independent variables in the test regression). If the LM value is less than or equal to the critical value from the χ^2 distribution, we do not have grounds to reject the null hypothesis.

The statistical tests discussed above are a basic set of econometric tools used to assess the properties of the error term in the estimated model. Good stochastic properties of the model therefore imply no autocorrelation, homoskedasticity, and a normal distribution of the error term.

The next step in the substantive and statistical evaluation of the model is to examine the significance of the explanatory variables' impact on the dependent variable and to assess the fit to the empirical data. We conduct the significance test separately for each parameter according to the following scheme:

$H_0\colon \beta_k = 0$ – the explanatory variable x_k does not significantly affect the dependent variable Y (10)

$H_1\colon \beta_k \neq 0$ – the explanatory variable x_k significantly affects the dependent variable Y (11)

The null hypothesis is typically tested at a significance level of 0.05 with a two-tailed rejection region (which follows from the form of the alternative hypothesis). The test statistic in this test is given by:

$$t_{\widehat{\beta}_k} = \frac{\widehat{\beta}_k}{S_{\widehat{\beta}_k}} \tag{12}$$

where: $\hat{\beta}_k$ – the estimated regression coefficient for the k-th variable; $S_{\widehat{\beta}_k}$ – the standard error of the estimator β_k.

The statistic (12) follows a Student's t-distribution with $n - k$ degrees of freedom, where n – the number of observations, k – the number of estimated parameters.

The coefficient of determination (R^2) indicates the proportion of the variability in the dependent variable that is explained by the variability in the explanatory variable (Chudy-Hyski, 2006). The R^2 value ranges from 0 to 1. The closer R^2 is to 1, the greater the proportion of variability in the dependent variable explained by the model (Borkowski et al., 2007). The coefficient of determination allows for the assessment of which of the analyzed models has the best fit to the empirical data.

Empirical Results

Data

The analyzed sample includes monthly data for the period from February 2004 to November 2023. All figures were generated using Gretl. The data sources for the respective variables in the model are as follows:

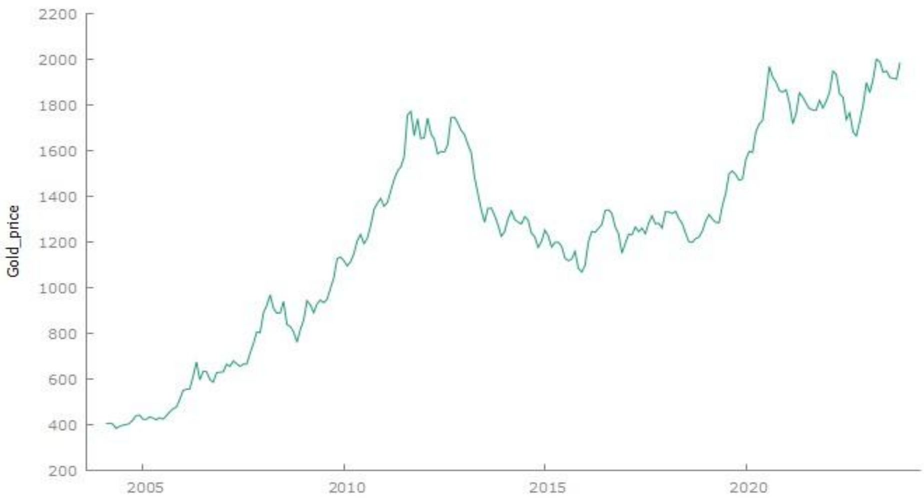
- gold prices (per troy ounce, equivalent to 31.1035 grams) – World Gold Council database (see Figure 4),

- crude oil prices (per barrel, USD, average of Brent, WTI, and Dubai) – World Bank database (see Figure 5),
- S&P 500 Index – Stooq database (see Figure 6),
- US$/EUR exchange rate - Investing.com database (see Figure 7),
- CPI (Consumer Price Index) in the United States, 1982–1984=1 – Federal Reserve Bank of St. Louis database (see Figure 8).

The descriptive statistics of the variables are provided in the appendix.

Figure 4

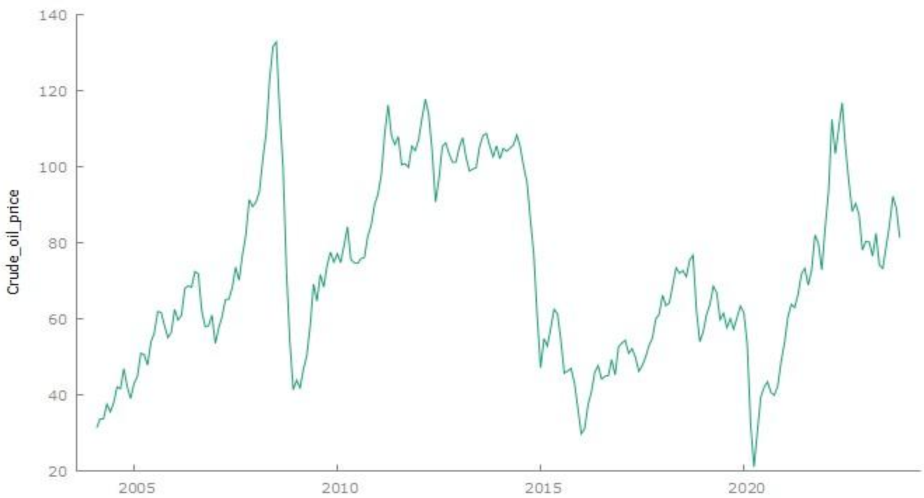
Gold price (per troy ounce, USD)



Source: Gretl software.

Figure 5

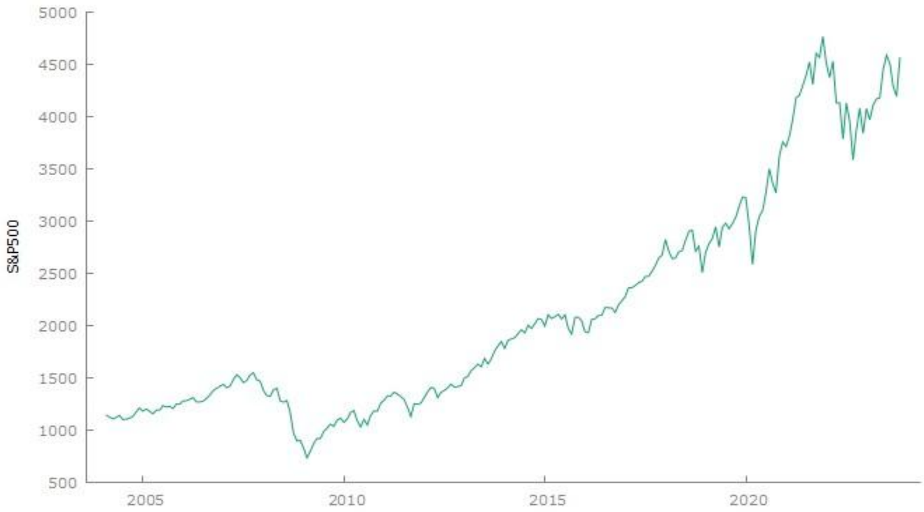
Crude oil price (per barrel, USD, average of Brent, WTI, and Dubai)



Source: Gretl software.

Figure 6

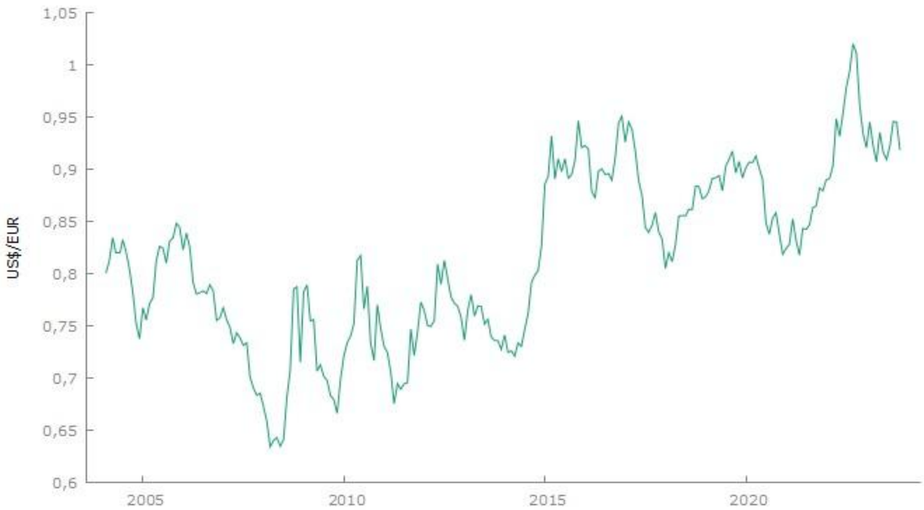
S&P500 Index



Source: Gretl software.

Figure 7

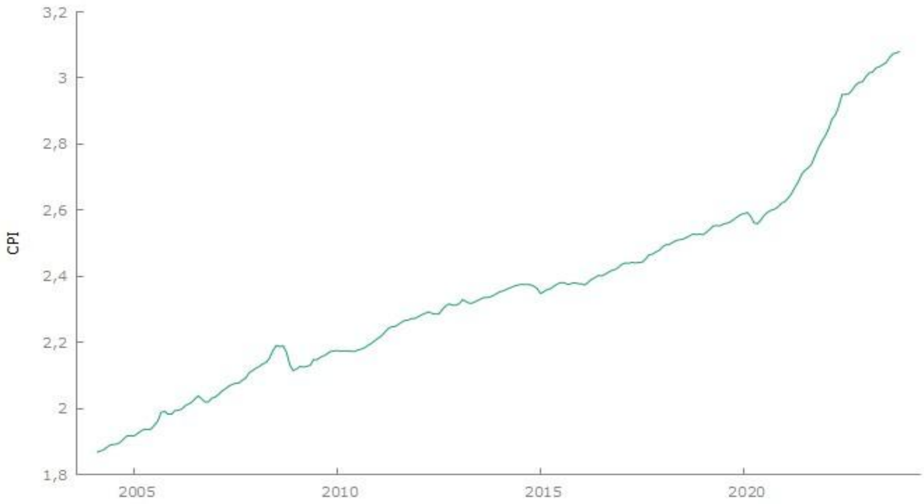
US$/EUR exchange rate



Source: Gretl software.

Figure 8

CPI



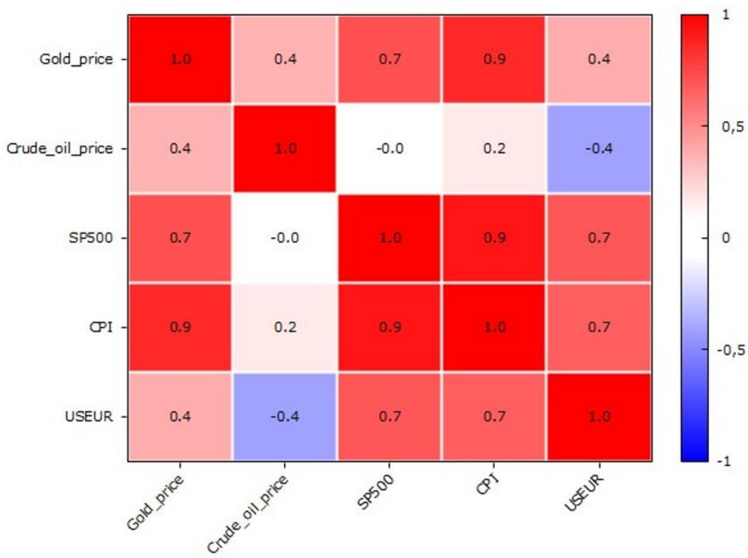
Source: Gretl software.

Correlation Matrix

The graphs below present correlation matrices (see Matrix 1, Matrix 2) – matrices where the elements represent the correlation coefficients for the respective pairs of variables. Matrices were generated using Gretl.

Matrix 1

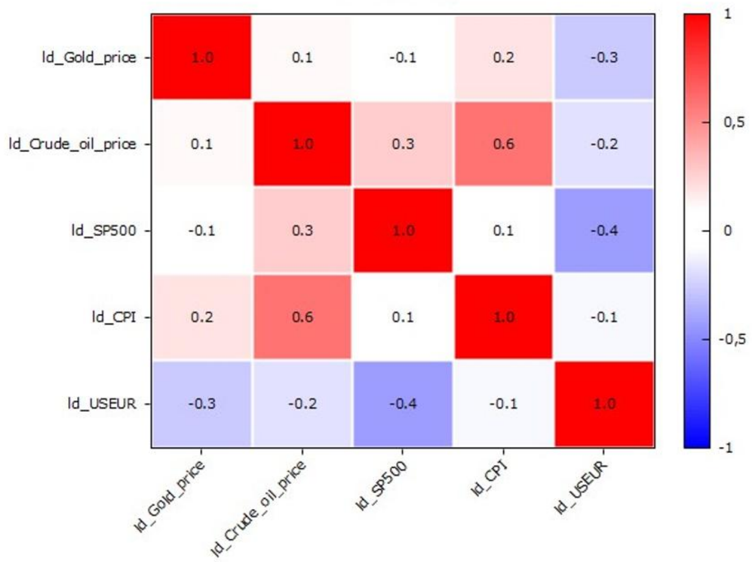
Correlation matrix for variable levels



Source: Gretl software.

Matrix 2

Correlation matrix for logarithmic increments of variables



Source: Gretl software.

For the levels of variables, the highest correlation is observed between the gold price and the CPI index, while the weakest correlations are between the gold price and the oil price, as well as between the gold price and the US$/EUR exchange rate. For the logarithmic increments of variables (see Matrix 2), the signs of the relationships between the gold price and other variables are consistent with the research hypotheses (oil price and CPI index positively impact the gold price, while the S&P 500 index and US$/EUR exchange rate negatively impact it). The weakest correlations are between the gold price and the S&P 500 index, as well as between the gold price and the oil price.

Gold Price Model

In the first step of the empirical section, the structural parameters of the model were estimated:

$$\ln Gold\_price_t = \beta_0 + \beta_1 \ln Crude\_oil\_price_t + \beta_2 \ln S\&P500_t + \beta_3 \ln US\$/EUR_t + +\beta_4 \ln CPI_t + \varepsilon_t \quad (13)$$

Based on the estimation results (see Table 1), it can be observed that the oil price does not have a statistically significant impact on the gold price ($|t - Statistic| \leq 2$), which contradicts the research hypothesis. For the remaining explanatory variables – the CPI index, US$/EUR exchange rate, and S&P 500 index – the results are consistent with the hypothesized relationships, and there is a statistically significant impact on the dependent variable. The CPI index has a positive effect on the gold price, while the S&P 500 index and the US$/EUR exchange rate have a negative impact on the gold price. An increase in the U.S. Consumer Price Index (1983=100) by 1% leads to a 6% increase in the gold price, ceteris paribus, whereas a 1% increase in the US$/EUR exchange rate results in a decrease in the gold price by approximately 1.2%, and a 1% rise in the S&P 500 index causes a 0.5% decrease in the gold price, ceteris paribus.

Table 1

OLS estimation results, observations used: 2004:02–2023:11 (N = 238)

Dependent variable (Y): l_Gold_price				
	Coefficient	**Standard error**	**t-Statistic**	**p-value**
Const	5,93119	0,390155	15,20	<0,0001
l_CPI	5,46213	0,270593	20,19	<0,0001
l_Crude_oil_price	−0,0314360	0,0481458	−0,6529	0,5144
l_USEUR	−1,18899	0,202854	−5,861	<0,0001
l_SP500	−0,484438	0,0652202	−7,428	<0,0001

Source: Gretl software.

Table 2

Selected fit measures and statistics for model (13)

Mean of dependent variable	7,037260	**Standard deviation of dependent variable**	0,445108
Sum of squared residuals	7,373719	**Standard error of residuals**	0,177896
R-squared	0,842961	**Adjusted R-squared**	0,840265
Residual autocorrelation – rho1	0,981882	**Durbin-Watson statistic**	0,052570

Source: Gretl software.

To examine the stochastic properties of the error term, the following tests were conducted: White's test for heteroskedasticity, the LM test for autocorrelation of the error term, and the Doornik-Hansen test for normality of the error term distribution (see Table 3).

Table 3

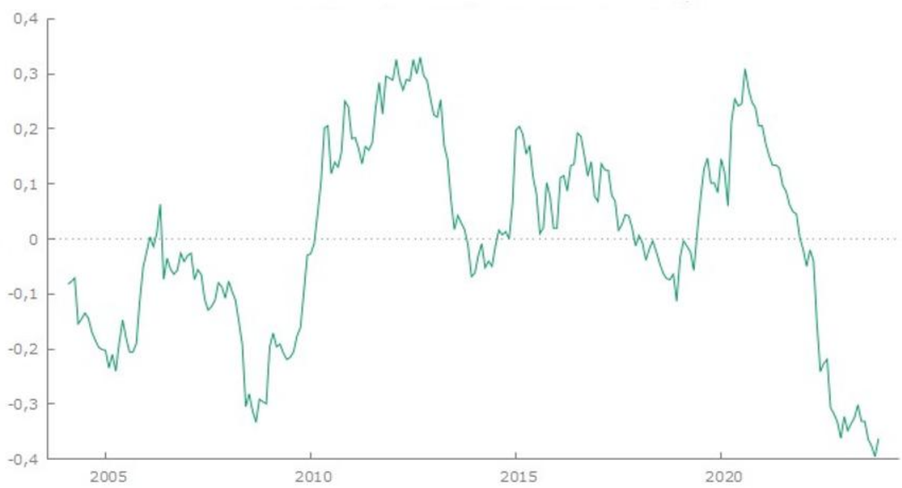
Testing the stochastic properties of the model

Test	Null hypothesis	Test statistic	p-value
White's test	`No heteroskedasticity in residuals	168, 058	1,68951e-28
Doornik-Hansen test	Error term is normally distributed	8,98924	0,0111689
LM test for 12-th order autocorrelation	No autocorrelation in error term	372,356	3,09014e-139

Source: Gretl software.

Figure 9

Residual plot of the error term for model (13)



Source: Gretl software.

The error term exhibits heteroskedasticity ($p - value \leq 0{,}05,$ so there is a reason to reject the null hypothesis, H_0). The distribution of the error term is not normal ($p - value \leq 0{,}05$, so we reject H_0 in favor of the alternative hypothesis). There is autocorrelation in the error term ($p - value \leq 0{,}05$, so there is a reason to reject H_0). The empirical residuals do not constitute white noise (see Figure 9).

To eliminate the presence of autocorrelation, a lagged dependent variable was introduced as an explanatory variable:

$$\ln Gold\_price_t = \beta_0 + \beta_1 \ln Crude\_oil\_price_t + \beta_2 \ln S\&P500_t + \beta_3 \ln US\$/EUR_t + +\beta_4 \ln CPI_t + \beta_5 \ln Gold\_price_{t-1} + \varepsilon_t \tag{14}$$

The results of the estimation for the structural parameters of model (14) are presented in Table 4.

Table 4

OLS estimation results for model (14), observations used: 2004:03–2023:11 (N = 237)

Dependent variable (Y): l_Gold_price				
	Coefficient	**Standard error**	**t-Statistic**	**p-value**
Const	0,313256	0,110129	2,844	0,0048
l_CPI	0,262414	0,0895438	2,931	0,0037
l_Crude_oil_price	−0,0293083	0,00980851	−2,988	0,0031
l_USEUR	−0,146114	0,0435838	−3,352	0,0009
l_SP500	−0,0234551	0,0146622	−1,600	0,1110
l_Gold_price_1	0,963081	0,0130942	73,55	<0,0001

Source: Gretl software.

Table 5

Selected fit measures and statistics for model (14)

Mean of dependent variable	7,041621	**Standard deviation of dependent variable**	0,440925
Sum of squared residuals	0,301687	**Standard error of residuals**	0,036139
R-squared	0,993425	**Adjusted R-squared**	0,993282
Residual autocorrelation –rho1	0,146390	**Durbin's h-statistic**	2,300881

Source: Gretl software.

To examine the stochastic properties of the error term, the following tests were conducted: White's test for heteroskedasticity, the LM test for autocorrelation of the error term, and the Doornik-Hansen test for normality of the error term distribution (see Table 6).

Table 6

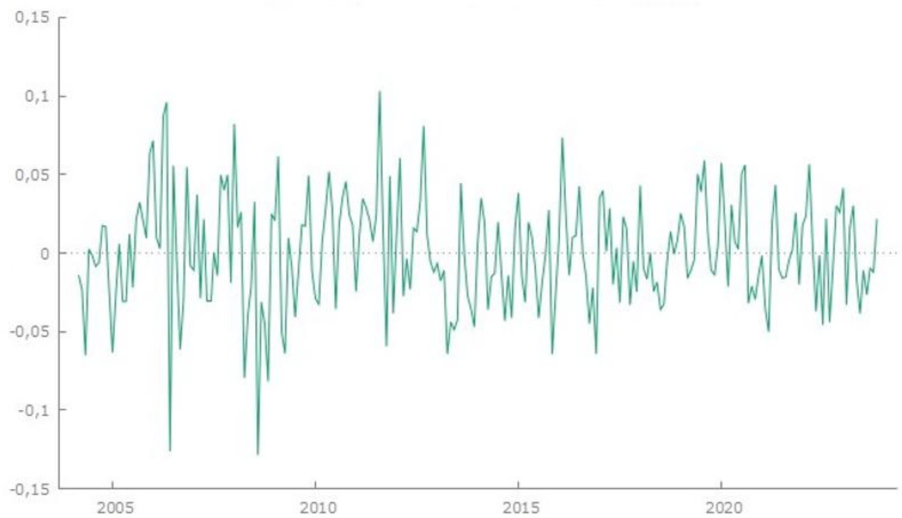
Testing the stochastic properties of the model

Test	Null hypothesis	Test statistic	p-value
White's test	No heteroskedasticity in residuals	35,2317	0,0189045
Doornik- Hansen test	Error term is normally distributed	6,96452	0,0307378
LM test for 12-th order autocorrelation	No autocorrelation in error term	1,28534	0,228483

Source: Gretl software.

Figure 10

Residual plot of the error term for model (14)



Source: Gretl software.

The S&P 500 index does not have a statistically significant impact on changes in the gold price ($|t - Statistic| \leq 2$), which contradicts the research hypothesis (see Table 9). The CPI index has a positive impact on the gold price, while the oil price and the US$/EUR exchange rate have a negative impact on the gold price (see Table 9). According to the results in Table 11, the error term is heteroskedastic ($p - value \leq 0{,}05,$ so there is a reason to reject H_0). The distribution of the error term is not normal ($p - value \leq 0{,}05,$ so there is a reason to reject H_0). There is no autocorrelation in the error term ($p - value > 0{,}05,$ so there is no reason to reject H_0). The estimated autoregressive parameter value close to 1 indicates instability of the model in the long term.

To verify the stationarity of the time series, the Augmented Dickey–Fuller (ADF) test was applied to all variables. The test's hypotheses are as follows:

H_0: The series contains a unit root (i.e., is non-stationary)

H_1: The series is stationary

Table 7

Results of the Augmented Dickey–Fuller (ADF) test for logarithmic variables

Variable	p-value
l_Gold_price	0,2575
l_CPI	0,997
l_SP500	0,5494
l_USEUR	0,2124
l_Crude_oil_price	0,009

Source: Gretl software.

The test results (see Table 7) indicate that the original variables are non-stationary (p-value > 0.05 in most cases). Therefore, in the next step, the analysis was conducted using logarithmic returns (logarithmic increments) of the variables, which are commonly used to stabilize variance and achieve stationarity:

$$\Delta \ln Gold\_price_t = \beta_0 + \beta_1 \Delta \ln Crude\_oil\_price_t + \beta_2 \Delta \ln S\&P500_t + \\ + \beta_3 \Delta \ln US\$/EUR_t + \beta_4 \Delta \ln CPI_t + \beta_5 \Delta \ln Gold\_price_{t-1} + \varepsilon_t \quad (15)$$

The model also includes dummy variables related to events such as the COVID-19 pandemic and the financial crisis. The results of the estimation for the structural parameters of model (15) are presented in Table 8.

Table 8

OLS estimation results for model (15), observations used: 2004:04-2023:11 (N = 236)

Dependent variable (Y): ld_Gold_price				
	Coefficient	**Standard error**	**t-Statistic**	**p-value**
const	0,00259679	0,00250659	1,036	0,3013
ld_USEUR	−0,481506	0,0864761	−5,568	<0,0001
ld_CPI	1,93050	0,649616	2,972	0,0033
ld_SP500	−0,188553	0,0536759	−3,513	0,0005

u0606	−0,151765	0,0318166	−4,770	<0,0001
u0808	−0,0926070	0,0320126	−2,893	0,0042
u0811	0,0851222	0,0315872	2,695	0,0076
u0420	0,0946007	0,0327479	2,889	0,0042
u0216	0,0813866	0,0314905	2,584	0,0104
ld_Gold_price_1	0,203871	0,0563506	3,618	0,0004

Source: Gretl software.

Table 9

Selected fit measures and statistics for model (15)

Mean of dependent variable	0,006718	**Standard deviation of dependent variable**	0,037338
Sum of squared residuals	0,221372	**Standard error of residuals**	0,031367
R-squared	0,324301	**Adjusted R-squared**	0,294270
Residual autocorrelation – rho1	0,024330	**Durbin's h-statistic**	0,761204

Source: Gretl software.

To examine the stochastic properties of the error term, the following tests were conducted: White's test for heteroskedasticity, the LM test for autocorrelation of the error term, and the Doornik-Hansen test for normality of the error term distribution (see Table 10).

Table 10

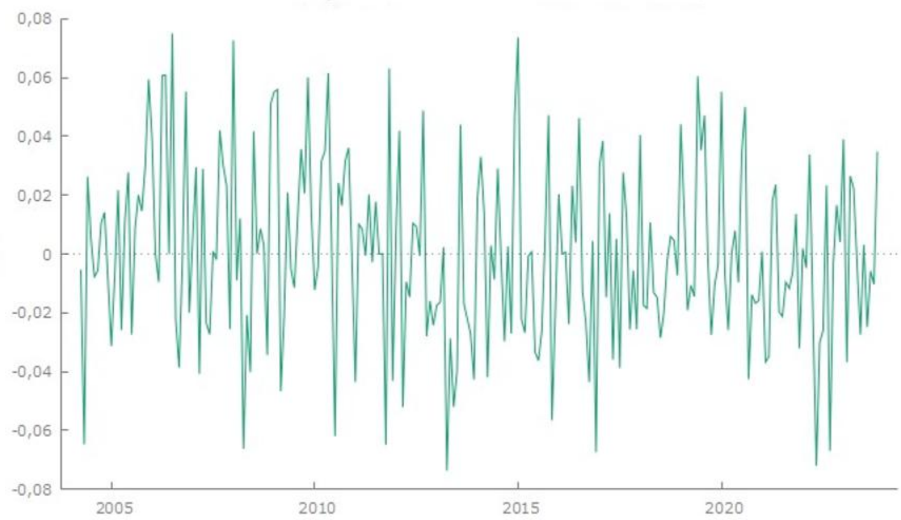
Testing the stochastic properties of the model

Test	Null hypothesis	Test statistic	p-value
White's test	No heteroskedasticity in residuals	26,9449	0,358642
Doornika-Hansen test	Error term is normally distributed	1,17641	0,555323
LM test for 12th-order autocorrelation	No autocorrelation in error term	1,59451	0,094803

Source: Gretl software.

Figure 11

Residual plot of the error term for model (15)



Source: Gretl software.

Due to the statistical insignificance of the oil price variable ($|t - Statistic| \leq 2$), it is not included in the Table 7. The other explanatory variables have a statistically significant effect on the dependent variable. As hypothesized, the CPI index has a positive impact on the gold price, while the S&P 500 index and the US$/EUR exchange rate have a negative impact on the gold price. Based on Table 14, the error term is homoskedastic ($p - value > 0{,}05,$ so there is no reason to reject H_0). The error term is normally distributed ($p - value > 0{,}05,$ so there is no reason to reject H_0). There is no autocorrelation in the error term ($p - value > 0{,}05,$ so there is no reason to reject H_0).

Due to potential instability in the long term of the model for logarithmic variables, the model for the logarithmic increments (see model (15)) appears to be the best choice.

To conclude, the CPI index has a positive impact on changes in the gold price, which is consistent with the research hypothesis. The S&P 500 index and the US$/EUR exchange rate negatively affect changes in the gold price, which also aligns with the proposed hypotheses. After excluding the insignificant variable, the stochastic properties of the model remained unchanged. The error term is homoskedastic, normally distributed, and no autocorrelation is present.

Utilization of Microsoft Copilot for Gold Price Analysis

Microsoft Copilot is a chat assistant powered by OpenAI models. During the research work, Copilot Pro- a paid version was utilized, which differs from the free version in terms of performance and priority access to GPT-4 and GPT-4 Turbo during peak hours. Additionally, AI features can be enabled in Office applications. The use of this tool involves asking questions, and the responses generated by Microsoft Copilot included information such as:

1) Correlation between gold prices and the examined independent variables.

According to the analysis provided by Copilot, there is a strong and consistent correlation between gold prices and crude oil prices. The relationship between gold and oil prices is positively correlated. Copilot also identified a weak negative correlation between gold prices and the S&P 500 index. However, it noted the influence of varying economic conditions, which can affect the relationships between these variables, leading to periods of positive correlation. The analysis found a negative correlation between gold prices and the US$/EUR exchange rate. This is related to investor behavior during periods of U.S. dollar depreciation. Lastly, Copilot highlighted that gold prices and the CPI (Consumer Price Index) tend to exhibit a general positive correlation. However, this correlation is not constant and may fluctuate depending on various factors, such as economic conditions.

2) Other determinants of gold prices.

Microsoft Copilot also identified additional factors that influence gold prices, including:

- central bank reserves,
- demand for jewelry and industrial use,
- gold production,
- investment demand.

3) Methodologies for forecasting gold prices.

Copilot suggested several methods for estimating model parameters, including:

- regression models,
- time series models,
- deep learning models.

4) Literature on gold price modelling.

Here are some examples of scholarly articles recommended by Microsoft Copilot that pertain to gold price modeling:

- Amini, A., Kalantari, R. (2024). Gold price prediction by a CNN-Bi-LSTM model along with automatic parameter tuning,
- Jabeur, S.B., Mefteh-Wali, S., Viviani, J.L. (2021). Forecasting gold price with the XGBoost algorithm and SHAP interaction values,
- Livieris, I.E., Pintelas, E., Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting,
- Ismail, Z., Yahya, A., Shabri, A. (2009). Forecasting gold prices using multiple linear regression method.

5) R code.

In the course of this study, Microsoft Copilot was employed to generate R code for analyzing the time series of gold prices. A natural language prompt requesting code to model gold prices using the ARIMA approach was provided, and Copilot produced a complete script including data importation, transformation of prices into logarithmic returns, graphical visualization, and application of the auto.arima() function from the forecast package, which automatically selects the most appropriate ARIMA model based on statistical information criteria such as AIC and BIC. The code generated by Copilot is not retrieved from a fixed template (or pre-written library) but is dynamically composed based on patterns learned during the training of the underlying language model, GPT-4, which was trained on large volumes of publicly available code, technical documentation, and statistical programming examples, allowing it to predict plausible sequences of code given a specific instruction. This generation process is guided by statistical associations learned from data rather than logical reasoning or domain-specific rules. As a result, while the produced code is often syntactically and technically accurate, it lacks an explicit connection to economic theory or justification grounded in domain expertise. The auto.arima() function in R, recommended by Microsoft Copilot, selected the ARIMA(1,0,0) model as the best tool for forecasting gold prices (the relevant R code is provided in the appendix).

Justification for the use of the ARIMA Model in Gold Price Forecasting

As part of this academic analysis, a reasoned justification for the selection of the ARIMA model in forecasting gold prices was requested from Microsoft Copilot. The model recommendation provided by the AI highlights both the empirical characteristics of the time series data and the theoretical strengths of ARIMA as a univariate time series framework.

According to the response, the choice of the AutoRegressive Integrated Moving Average (ARIMA) model is well-founded due to its robustness in handling non-stationary financial time series, such as gold prices, which are influenced by a wide range of macroeconomic and geopolitical factors and are known to exhibit volatility and trends. The integrated component of ARIMA allows the transformation of non-stationary data into a stationary series through differencing, enabling meaningful statistical modeling and forecasting.

The AI also emphasized that, unlike seasonal models, ARIMA does not require the presence of cyclical patterns, which is advantageous when dealing with time series affected by irregular or external shocks rather than regular seasonality. Furthermore, ARIMA is widely used and validated for short- to medium-term forecasting and provides not only point estimates but also confidence intervals, which are essential for quantifying uncertainty in financial contexts.

To enhance objectivity and reproducibility, the AI recommended the use of the auto.arima() function in R. This function automates the selection of optimal model parameters (p, d, q) based on established information criteria such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

A key point raised by the AI was the deliberate decision to exclude exogenous variables – such as interest rates, inflation expectations, or exchange rates – from the model. Although theoretically relevant, their exclusion is justified due to practical concerns related to data availability, forecastability, and the risk of multicollinearity. Including such variables could also complicate model interpretation. The AI argued that focusing on a univariate approach allows for a clearer understanding of the internal dynamics of the gold price series and ensures greater methodological transparency.

In summary, Microsoft Copilot recommended the ARIMA model due to its statistical flexibility, suitability for non-stationary financial data, ease of implementation, and interpretability in a univariate setting. These characteristics make ARIMA an appropriate and defensible modeling strategy for forecasting the behavior of gold prices within the scope of this academic study.

Comparison of Results Between the AR(1) Model and the Causal Model

To identify a better model, the ARIMA(1,0,0) model was compared with model (15). The results for the ARIMA(1,0,0) model, as suggested by Copilot, are presented in Table 11 (for the logarithm of the gold price).

Table 11

OLS estimation results, observations used: 2004:03–2023:11 (N = 237)

Dependent variable (Y): l_Gold_price				
	Coefficient	**Standard error**	**t-Statistic**	**p-value**
Const	0,0888882	0,0381607	2,329	0,0207
l_Gold_price_1	0,988318	0,00541372	182,6	<0,0001

Source: Gretl software.

Table 12

Selected fit measures and statistics for the ARIMA(1,0,0) model

Mean of dependent variable	7,041621	**Standard deviation of dependent variable**	0,440925
Residual sum of squares	0,321259	**Standard error of residuals**	0,036974
R-squared	0,992998	**Adjusted R-squared**	0,992968
Residual autocorrelation – rho1	0,141855	**Durbin's h-statistic**	2,191457

Source: Gretl software.

In the next step, the parameters of an autoregressive model for the log return of the gold price were estimated. The results for the ARIMA(1,0,0) model for the log return of the gold price are shown in Table 13.

Table 13

OLS estimation results, observations used 2004:04–2023:11 (N = 236)

Dependent variable (Y): ld_goldprice				
	Coefficient	**Standard error.**	**t-Statistic**	**p-value**
Const	0,00573454	0,00244576	2,345	0,0199
ld_goldprice_1	0,149472	0,0647309	2,309	0,0218

Source: Gretl software.

Table 14

Selected fit measures and statistics for the ARIMA(1,0,0) model for the log return of gold price

Mean of dependent variable	0,006718	**Standard deviation of dependent variable**	0,037338
Residual sum of squares	0,320320	**Standard error of residuals**	0,036998
R-squared	0,022279	**Adjusted R-squared**	0,018101
Residual autocorrelation – rho1	0,006913	**Durbin's h-statistic**	1,006232

Source: Gretl software.

The model based on literature for the logarithmic levels of variables (see model (14)) exhibits a higher adjusted R-squared compared to the autoregressive model suggested by the auto.arima() procedure (see Table 12). This indicates that the model, which also accounts for other determinants and outliers, better explains the dependent variable. The same conclusion can be drawn when comparing the results for models with log returns (see Table 9 and Table 14). Microsoft Copilot suggested an automated ARIMA model selection procedure, which does not account for other factors influencing gold price changes. Adding explanatory variables (i.e., the S&P500 index, US$/EUR exchange rate, CPI index) improved the explanatory power of the model.

Artificial Intelligence as a Suport Tool in Econometrics

Artificial intelligence represents a disruptive technology that is fundamentally altering approaches to knowledge acquisition and task execution. Its application is increasingly observed within the field of econometrics. AI systems, particularly large language models (LLMs), possess the capability to process and analyze both numerical and unstructured textual data (e.g., news articles, social media content), thereby enabling a more nuanced interpretation of economic phenomena. Moreover, AI significantly enhances the efficiency of literature discovery and data retrieval by rapidly identifying thematically relevant sources. Despite these advantages, notable limitations remain. In econometric analysis, the interpretability of model coefficients is essential for drawing valid inferences – an aspect where AI-generated outputs frequently fall short. Such models often lack transparency in how specific variables influence the dependent outcome. Furthermore, AI approaches are predominantly data-driven and not grounded in established economic theory. Consequently, while AI may generate empirically adequate results, it often fails to provide theoretical justification or causal explanation.

Conclusion

All the objectives set out in this study were achieved. Based on the literature, factors influencing the change in gold prices were identified. Crude oil prices were found to be statistically insignificant. The best results were obtained for a model explaining the log return of gold prices, in which the explanatory variables included the S&P500 index, the US$/EUR exchange rate, and the CPI index. Lagged gold prices, representing the autoregressive factor, also proved to be significant. The statistical significance of these variables' impact on the dependent variable was confirmed. Additionally, the model demonstrated good stochastic properties, indicating the absence of autocorrelation, homoscedasticity, and normality of the residuals. The positive impact of the CPI index and the negative impact of the US$/EUR exchange rate and the S&P500 index were confirmed. COVID-19, the financial crisis, and other outlier observations were accounted for in the model by introducing appropriate dummy variables.

In addition to the econometric model, the Microsoft Copilot tool was used to study the phenomenon under investigation. This program proved helpful in finding some of the literature and identifying determinants of gold prices. However, the R code suggested by Copilot only applied an automatic procedure for the ARIMA model, which indicated a simple AR(1) model. The model that includes both the autoregressive factor and other determinants of gold prices has a higher coefficient of determination and better explains the variable under study. In summary, the advantage of causal models is that they consider factors affecting the phenomenon under study, allowing for more effective modeling of economic processes. The Microsoft Copilot tool is useful in finding information that aids in constructing the model. However, human input remains essential in the modeling process, particularly in evaluating the model's substantive content and making decisions regarding the choice of its appropriate form.

Acknowledgements

The article is based on the findings of the research project “ChatGPT czy model ekonometryczny?”, conducted under the academic supervision of dr Emilia Gosińska. This work was developed within the framework of the 7th edition (2023/2024) of the “Talented at School, Brilliant at the University” programme at the University of Lodz. The author gratefully acknowledges the support and guidance received throughout the project.

References

Amini, A., Kalantari R. (2024). Gold price prediction by a CNN-Bi-LSTM model along with automatic parameter tuning. *PLoS ONE* 19(3). https://doi.org/10.1371/journal.pone.0298426

Bankier.pl – https://www.bankier.pl/

Borkowski, B., Dudek H., Szczęsny W. (2007). Ekonometria. wybrane zagadnienia. *Wydawnictwo Naukowe PWN*.

Bukowski, S.I. (2016). The main determinants of gold price in the international market. *International Business and Global Economy,* Tom 35/1, 402–413. https://doi.org/10.4467/23539496IB.16.029.5610

Choong, P., Kwoo, P., Piong, C., Wong, W. (2012). Determinants of Gold Price: Using Simple and Multiple Linear Regression. *Universiti Tunku Abdul Rahman*, 1–99.

Chudy-Hyski, D. (2006). Ocena wybranych uwarunkowań rozwoju funkcji turystycznej obszaru. *Infrastruktura i Ekologia Terenów Wiejskich,* 2/1.

Greene W.H. (2002). Econometric analysis, 5th edition, *Pearson.*

Gretl software – https://gretl.sourceforge.net/

Guha, B., Bandyopadhyay, G. (2016). Gold price forecastin using ARIMA model. *Journal of Advanced Management Science,* vol. 4, no. 2., 117–121. http://dx.doi.org/10.12720/joams.4.2.117-121

Ismail, Z., Yahya, A., Shabri, A. (2009). Forecasting gold prices using multiple linear regression method. *American Journal of Applied Sciences,* vol. 6 (8), 1509–1514. https://doi.org/10.3844/ajassp.2009.1509.1514

Jabeur, S.B., Mefteh-Wali, S., Viviani, J.L. (2021). Forecasting gold price with the XGBoost algorithm and SHAP interaction values. *Annals of Operations Research,* vol. 334, 679–699. https://link.springer.com/article/10.1007/s10479-021-04187-w

Levin, E.J., Wright, R.E. (2006). Short-run and long-run determinants of the price of gold. *World Gold Council*, Research Study no. 32.

Livieris, I.E., Pintelas, E., Pintelas, P. (2020). A CNN–LSTM model for gold price time-series forecasting. *Neural Comput & Applic,* vol. 32, 17351–17360. https://link.springer.com/article/10.1007/s00521-020-04867-x

Maddala, G.S. (2001). Introduction to econometrics, 3rd edition. *Wiley India*, ISBN: 9788126510955.

Makalala, D., Li, Z. (2021). Prediction of gold price with ARIMA and SVM. *Journal of Physics: Conference Series,* vol. 1767. http://dx.doi.org/10.1088/1742-6596/1767/1/012022

Polyus https://sustainability.polyus.com/en/esg_data_and_reports/

Puci, J., Demi, A., Pjeshka, A. (2022). The effect of the S&P500 on gold prices. *Journal of Financial and Monetary Economics,* vol. 10, 308–313.

Setyowibowo, S., As'ad, M., Sujito, S., Farida, E. (2021). Forecasting of daily gold price using ARIMA-GARCH hybrid model. *Jurnal Ekonomi Pembangunan,* vol. 19, 257–270. https://www.researchgate.net/publication/358593313_Forecasting_of_Daily_Gold_Price_using_ARIMA-GARCH_Hybrid_Model

Verbeek, M. (2012). A guide to modern econometrics, 4th Edition, *John Wiley & Sons, Ltd.*

Yang, X. (2019). The prediction of gold price using ARIMA model. *2nd International Conference on Social Science, Public Health and Education.* http://dx.doi.org/10.2991/ssphe-18.2019.66

Zhang, P., Ci, B. (2020). Deep belief network for gold price forecasting. *Resources Policy,* vol. 69. http://dx.doi.org/10.1016/j.resourpol.2020.101806

Appendix

R code suggested by Microsoft Copilot

```
# Load necessary libraries
library(forecast)
library(tseries)
# Load the data
gold_data <- read.csv("gold_price_data.csv", sep=";", header=TRUE)
# Convert the 'obs' column to date format
gold_data$obs <- as.yearmon(gold_data$obs, format = "%Y.%m")
# Create a time series object
gold_ts <- ts(gold_data$gold_price, start=start(gold_data$obs), frequency=12)
# Fit an ARIMA model
fit <- auto.arima(gold_ts)
# Display the model summary
summary(fit)
# Forecast future gold prices
forecasted_values <- forecast(fit, h=12)
# Print the forecasted values
print(forecasted_values)
```

Descriptive statistics of the variables

Below are the descriptive statistics of the analyzed variables (see Table 15, Table 16, Table 17, Table 18, Table 19). All descriptive statistics were generated using Gretl.

Table 15

Descriptive statistics for the $Gold\_price$ variable

Mean	Median	Minimum	Maximum
1239,1	1268,8	383,80	2000,4
Standard deviation	**Coefficient of variation**	**Skewness**	**Kurtosis**
449,73	0,36295	–0,28255	–0,83719
5th Percentile	**95th Percentile**	**Interquartile range (Q3-Q1)**	
423,97	1913,2	702,60	

Source: Gretl software.

Table 16

Descriptive statistics for the $Crude\_oil\_price$ variable

Mean	Median	Minimum	Maximum
72,052	68,735	21,040	132,83
Standard deviation	**Coefficient of variation**	**Skewness**	**Kurtosis**
23,983	0,33285	0,28893	–0,86590
5th Percentile	**95th Percentile**	**Interquartile range (Q3-Q1)**	
37,549	108,83	38,765	

Source: Gretl software.

Table 17

Descriptive statistics for the $S\&P500$ variable

Mean	Median	Minimum	Maximum
2137,8	1827,1	735,09	4766,2
Standard deviation	**Coefficient of variation**	**Skewness**	**Kurtosis**
1080,0	0,50519	0,90364	–0,38139
5th Percentile	**95th Percentile**	**Interquartile range (Q3-Q1)**	
1030,2	4375,0	1517,4	

Source: Gretl software.

Table 18

Descriptive statistics for the $US\$/EUR$ variable

Mean	Median	Minimum	Maximum
0,81582	0,81810	0,63410	1,0201
Standard deviation	**Coefficient of variation**	**Skewness**	**Kurtosis**
0,083472	0,10232	–0,013089	–0,81796
5th Percentile	**95th Percentile**	**Interquartile range (Q3-Q1)**	
0,68312	0,94521	0,14080	

Source: Gretl software.

Table 19

Descriptive statistics for the CPI variable

Mean	Median	Minimum	Maximum
2,3596	2,3473	1,8670	3,0802
Standard deviation	**Coefficient of variation**	**Skewness**	**Kurtosis**
0,29446	0,12479	0,59079	–0,079229
5th Percentile	**95th Percentile**	**Interquartile range (Q3-Q1)**	
1,9170	2,9889	0,37994	

Source: Gretl software.

Received: 18.08.2025
Accepted: 26.09.2026
Available online: 30.09.2025