SIMULATING THE DYNAMICS OF THE GOLD MARKET USING COMPUTATIONAL MARKET DYNAMICS

Examiners: docent Ph.D. Tuomo Kauranne
D.Sc. Matylda Jabłońska
ABSTRACT

Lappeenranta University of Technology
Faculty of Technology
Degree Programme in Technomathematics and Technical Physics

Denise Uwamariya

Simulating the Dynamics of the Gold Market Using Computational Market Dynamics

Master’s thesis

2012

45 pages, 30 figures, 5 tables

Examiners: docent Ph.D. Tuomo Kauranne
D.Sc. Matylda Jabłońska

Keywords: Gold prices, Animal spirits, Population dynamics, Ensemble simulation.

Ability to predict behavior of prices in various financial and commodity markets has been of big concern for researchers and analysis for a long time. However, this need increased over the past half a decade when the world faced a dramatic economy meltdown. These days econometricians still lack modelling tools that would be able to simulate, not to mention predict, extreme events in financial markets. The aim of this study is to simulate the dynamics of the gold market using models from a novel approach called computational market dynamics. The models used are Jabłońska-Capasso-Morale (JCM) and Kalman Dynamics (KD) models. Maximum likelihood has been used to fit model to the data in both cases. The models have been applied to daily London PM fix gold price, this price is quoted in Euro, and they include observations over the period from December 1969 up to April 2012. After carrying out the simulation analysis it was observed that simulated prices closely follow original price and two models have the same ability as they produced similar results.
Acknowledgements

I am deeply grateful to the Almighty God who has guided me throughout the whole period of my studies.

May this be a good opportunity to thank all those without their assistance and moral support; this study would have not achieved its goals. Thus, I would like to express my sincere gratitude firstly to the LUT for giving me a study opportunity and the Department of Mathematics and Physics for providing me financial support.

Special thanks to my supervisor, prof. Tuomo Kauranne for his technical and wise advice, suggestion and corrections that made this study successful. Also many thanks to my second examiner D.Sc. Matylda Jabłońska for her guidance, comments, advice and insightful review which made this study a success. Thanks to Anna Shcherbacheva for her technical support, and to all who generously gave their time and helped me to sharpen my knowledge during this year I have been training in LUT and contributed significantly to the realization of my studies.

And finally but not least, I am very thankful to my lovely husband; Niyinzi Billy Second for his patience all the time when I was away from home. He has showed me enough support accompanied with encouraging words, care and love and to my brothers and sisters and my family in law for their prayers and support throughout this year of my studies. Many thanks also to all my friends in Finland.

Thank you all

Lappeenranta, December 3, 2012.

Denise Uwamariya
Contents

1 Introduction 7

2 The gold market 8
   2.1 Gold market description and history 8
   2.2 Market functioning 9
   2.3 Market past modelling – literature review 9

3 Gold prices 11

4 Modelling background 17
   4.1 Introduction to stochastic differential equations (SDEs) and stochastic
       modelling 17
   4.2 Jabłońska-Capasso-Morale (JCM) model 19
   4.3 Kalman Dynamics (KD) model 22
       4.3.1 KD model parameters 24

5 Gold price modelling results 25
   5.1 JCM results 25
   5.2 KD results 29
   5.3 Comparison of model performance 33

6 Results summary and discussion 39

7 Conclusion 40

REFERENCES 41
CONTENTS

List of Tables 43

List of Figures 44
List of Symbols and Abbreviations

ACF           Autocorrelation Function
APARCH        Asymmetric Power ARCH
APGARCH       Asymmetric Power GARCH model
AR            Autoregressive
ARCH          Autoregressive Conditionally Heteroscedastic
GARCH         Generalized Autoregressive Conditionally Heteroscedastic
GED           Generalized Error Distribution
JCM           Jabłońska-Capasso-Morale
KD            Kalman Dynamics
Kt            karat
ODE           Ordinary Differential Equation
OTC           Over-the-Counter
PACF          Partial Autocorrelation Function
SDE           Stochastic Differential Equation
RMSE          Root Mean Square Error
RMSD          Root Mean Square Deviation
TARCH         Threshold ARCH model
VEnKF         Variational Ensemble Kalman Filtering
1 Introduction

Gold is a precious metal which has been used as a monetary asset and is classified as a commodity. Gold has played an important role as a precious metal and many people buy gold as a way of preserving their wealth throughout ages. Its value and the area of usage are increasing until today. Gold was the way of money system in the past in that it acts as a storage of wealth, a medium of exchange and a unit value. After 1973 some European Countries stopped their fixed exchange rate with dollar and convertibility of dollar against gold at the same time, and under these issues gold lost its property as mean of exchange and started being a personal saving tool and a part of Central Bank Reserves.

In the recent years, demand of gold has expanded with the use of gold both in industrial goods and in the jewelery sector. However, the quantity of gold required is determined by the quantity demanded for industrial and in the jewelery sector. Therefore, the price of gold increases as the quantity demanded by the industry and jewelery sector increases and price changes can also be the result of a change in the Central Bank's holding gold as a metal [17].

There are many studies that provide different models and modification of suggested models to explain dynamics and forecast future prices for the financial and commodity markets, including the gold market, but none of them had the power to capture their spiky behavior due to high volatility found in the financial and commodity markets, in which high volatility means the larger variation in the demand and supply in this markets.

Recently, some researchers have found that price spikes may be caused by price dynamics resulting from market momentum and the psychology of traders. This analogy statement is related to animal spirits, suggested first by Keynes in 1936 [10]. This study aims to simulate the dynamics of gold prices by two mathematical models, namely, Jabłońska-Capasso-Morale (JCM) model introduced by Jabłońska [8] and Kalman Dynamics (KD) model to captures dynamics of gold markets. For KD model part of JCM model has been analyzed by Kalman Filttering Techniques.

The format for this work is as follows; section 2 presents the gold market, that includes market description and history, market functioning and market past modeling-literature review. In section 3 the data set used in this study is
described. In section 4, modelling background is presented together with some introduction to stochastic differential equations. Section 5 presents simulation results of the two models separately, comparison of model performance and summary and discussion of the results. Section 6 provides the conclusions of the study.

2 The gold market

This section aims to provide a background to the gold market. This section describes the metal itself and its history, gold market functioning and past market modelling efforts.

2.1 Gold market description and history

Gold is a brilliant yellow precious metal. It is very resistant to air and water corrosion, which is why it is also used as a commodity and a monetary asset. It is a very soft and pure metal (24 kt). A karat (kt) is the unit that measures the purity of gold jewelery it is hallmarked with a three-digit number that indicates the parts per thousand of gold.

Gold is the most malleable, ductile and popular metal found on earth. Often, alloys of gold are created through mixing it with other metals, usually copper and silver, to make it less expensive and harder. The alloyed gold comes in many colors and may not always be bright yellow [4, 7].

The natural gold was first found in Spanish caves used by the Paleolithic Man around 40,000 B.C. Gold has been used throughout history as a form of payment and has been a standard for currency equivalents to many economic regions or countries. The first use of gold as money in 700 B.C. "is claimed by the citizens of the Kingdom of Lydia (Western Turkey)" [7].

The gold market is highly liquid. Gold held by central banks, other major institutions and distribution of jewelery is reinvested in the market. The world gold markets are located in London, New York, Zurich, Istanbul, Dubai, Singapore, Tokyo and Mumbai. The world’s largest gold producing country is South Africa, whereas India is the world’s largest consumer of gold in jewelery
as investment. World’s gold demand is constantly increasing and it is nearing
record levels at 4000 tons per year while the mine production is constant at
2250 tons per year [7, 13].

2.2 Market functioning

Gold is traded as a commodity, but primarily it is a monetary asset [4]. The
trade in gold market exists in two different forms, that is, the Over-the-Counter
(OTC) market and the exchange-based market. As with other equities, there
are also the spot and the derivative segments. The spot markets are essen-
tially over-the-counter markets. Their participants are those who are directly
involved in production and distribution of the commodity, such as, the pro-
ducers, refiners, central bank, etc. Derivative trading takes place in the form
of exchange-based markets with standardized contracts, settlements, etc.

The trading that takes place in these OTC markets is delivery-based. The
buyer as well as the seller have their set of brokers who negotiate the prices for
them. In addition, in the case of a deal taking place, the gold and the money
would be exchanged directly between the buyer and the seller.

Furthermore, the exchange trade market’s characteristics are very specific and
are similar in their functioning to equity derivatives. Here, a person can pur-
chase a contract by paying only a small percentage of the contract value. Many
people who participate in the exchanges are those who are not involved with
the physical trading of the commodity [12].

The gold trade is heavily influenced by what happens in equity markets. Gold
performs well when major stock markets are in turmoil, and if equity markets
underperform, trading activity in the gold market is likely to rise. This suggests
that volatility dynamics of gold returns are likely to be different from other
commodity markets [2].

2.3 Market past modelling – literature review

There are a number of different models on the price of gold in the literature.
This section presents some of them.
Abken (1980) investigated factors that influence gold price movements [1]. For this purpose, gold prices were considered as endogenous variables and lagged values of gold prices and interest rates were considered as exogenous variables in a regression analysis. Monthly data between 1973-January and 1979-December have shown that the explanatory ratio of the regression equation is low. When similar relationship was sought between future prices and future spot prices, the explanatory ratio improved significantly. Also, the study found that the spot prices of all storable commodities, including gold, are particularly influenced by anticipations of future spot prices [1].

Another study was made by Dooley et al. (1992) [5]. Monthly data between 1976 – 1990 was used to apply multivariate vector regression and cointegration modelling techniques to test short- and long-run influence of gold prices on exchange rates, conditional on other monetary and real macroeconomic variables. Also, the study provided tests that exchange rates may be largely coterminous with changes in preferences for holding claims on different countries. It was found that exchange rates explain the changes in gold prices and changes in country preferences will be systematically reflected in the price of gold. It can be interpreted that exchange rates are largely coterminous with a change in preferences for holding claims on different countries [5].

Another study was made by Ghosh et al. (2002) using monthly data between 1976 and 1999, investigating apparent inconsistency between significant short-run price volatility and long-run movements in the price of gold with a cointegration regression [6]. It was concluded that gold price can be regarded as a long-run inflation hedge. The analysis also confirms that movements in the nominal price of gold are dominated by short-run influences [6].

Tully and Lucey (2005) used APGARCH model in order to investigate main influences on gold prices and whether the model captures the dynamics of gold market [17]. In the 1984 – 2003 period, a relationship was found between daily and future prices of gold and US Dollar, and APGARCH model provided the best fit for the data [17].

Canarella and Pollard (2008) used daily London gold fixing observation over the period from January 4, 1993 to December 30, 2004 to the London Gold Market Fixing for the applicability of the asymmetric power ARCH (APARCH) study found that the APARCH models estimated using student $t$ or Generalized Error Distribution (GED) leptokurtic conditional error distribution ex-
hibit evidence of unequal volatility responses to market shocks. Study also found that volatility in the London Gold Market Fixing is impacted more by positive shocks (‘good news’) rather than negative market shocks (‘bad news’) [2].

Another study was made by Stefan and Kevin (2012) who used daily PM fixings price from 4th January 1999 to 30th December 2008 to investigate the modelling of volatility dynamics of gold market returns in London. The study found that AR as well as TARCH models provided the best results [16].

3 Gold prices

The main focus of this work is the analysis of gold price behavior in spot markets by mathematical models. In this work, the theory discussed is illustrated by secondary data downloaded from world gold council website. The main dataset suitable to models discussed in this work is daily London PM fix gold price, this price is quoted in Euro, and includes observations over the period from December 1969 up to April 2012 (the weekend prices were not included) [18].

In this section, the data used in this study are presented. It is done through graphical presentations, such as time line plots, autocorrelation functions, partial autocorrelation functions and histograms. Also, some basic statistics such as mean, standard deviation, skewness and kurtosis have been used for exploratory analysis.

Most time series patterns can be described in terms of most basic classes of components: trend, seasonality and cyclic. Thus, a time line plot has been used to plot series and examine main features of the graph to check whether there are any patterns in the series behavior. Should any pattern be observed, it must be removed in the original series before applying classical time series analysis. However, models used in this study do not strictly require stationarity as they use the concept of moving mean reversion.

Figure 1 shows the time line plot of gold price data set, as it can be observed, there are enough observations for further analysis. But the series is non-stationary, that is, its mean value does not remain constant over time. The
data seems to have a trend to be removed before analysis. The series is also non-stationary with respect to variance.

The next step is to examine Autocorrelation function (ACF) and Partial Autocorrelation function (PACFs) plots. These are presented in Figure 2. Autocorrelation plot can be used to determine whether the series is random or whether there are any periodicities in the series, and whether the data is stationary or not. ACF and PACF plots are commonly used tools for Autoregressive Moving Average model identification in Box-Jenkins approach.

From ACF plot we can see a strong decaying autocorrelation within the series.
which is an evidence of the fact that the series is not stationary. In the PACF
plot, the correlation values lie approximately within the confidence limits but
there is a very high significant coefficient at lag 1.

We also need to verify the distribution of the series. In many classical econo-
metric models the data are required to have normal distribution. The data
histogram is present in Figure 3. Apparently, the series is not normally dis-
tributed. There is a very high peak which results in higher than normal dis-
tribution kurtosis.

Figure 3: (a) Histogram of gold price. (b) Normal histogram of gold price.

From those graphical representations the important features have been ob-
erved, that is, the series is not stationary and it has a trend. Thus, the next
step is transformation of the data. This has been done by difference and log
difference transformations to remove the trend and to make it stationary with
respect to mean value. Hence, Figure 4 presents the time line plot of gold price
transformations for both difference and log difference.

As can be seen in Figure 4, the difference series oscillate around their zero
mean, but their variances are not constant and vary over time.

The transformed data are less autocorrelated than the original data and simi-
lar features can be observed for difference and log difference transformations.
Figures 5 and 6 present the ACF and PACF plots of gold price difference and
log difference transformations.

By difference and log difference transformations the series becomes less auto-
correlated as Figures 5 and 6 have shown. There is no significant autocorrelation for neither ACF nor PACF of both transformations. From the histograms of gold price transformations presented in Figures 7 and 8 we observe that for both transformations the series are not normally distributed, though the distributions are a lot closer to bell-shape than the original prices. The series still have distribution with very high peak and tails that are significantly fatter than a normal distribution.

Figure 4: (a) Time line plot of gold price difference transformation. (b) Time line plot of gold price log difference transformation.

Figure 5: (a) ACF of gold price difference transformation. (b) PACF of gold price difference transformation.
Figure 6: (a) ACF of gold price log difference transformation. (b) PACF of gold price log difference transformation.

Figure 7: (a) Histogram of gold price difference transformation. (b) Normal histogram of gold price difference transformation.

The basic statistics of the data have been summarized in Table 1, which includes attributes such as mean, standard deviation, skewness and kurtosis.
Figure 8: (a) Histogram of gold price log transformation. (b) Normal histogram of gold price log transformation.

From above table, positive and negative skewness and high kurtosis can be observed for original and transformed data, which indicates a distribution with an asymmetric fat tail extending towards more positive values for original and log difference transformed data, for gold price difference transformed data an asymmetric fat tail extending towards negative values and relatively peaked distribution. This is also supported by the results observed from histogram figures. Also we observe that mean and standard deviation are positive, which indicates that gold prices were increasing during considered time period. Hence, the series are not normally distributed because the model values for skewness is different from 0 and kurtosis is greater than 3 for original and transformed data.

Table 1: Basic Statistics of gold price

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St.Dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold price</td>
<td>399.9554</td>
<td>202.8468</td>
<td>2.5360</td>
<td>9.6485</td>
</tr>
<tr>
<td>Gold price difference</td>
<td>0.1335</td>
<td>5.7777</td>
<td>-0.2601</td>
<td>23.1783</td>
</tr>
<tr>
<td>Gold price log</td>
<td>1.1259e-004</td>
<td>0.0053</td>
<td>0.1157</td>
<td>14.8148</td>
</tr>
</tbody>
</table>
4 Modelling background

The main focus of this work is the analysis of the gold price behavior using mathematical models. This section reviews briefly an introduction to stochastic differential equations and stochastic modelling. The stochastic models discussed are Jabłońska-Capasso-Morale (JCM) and Kalman Dynamics (KD) models.

4.1 Introduction to stochastic differential equations (SDEs) and stochastic modelling

Suppose a probability model defined on a probability space \((\Omega, F, P)\) where \(\Omega\) is a set known as sample space, \(F\) is subset of \(\Omega\) and \(P\) is a probability defined on \(F\), probability model can be used in the presence of uncertainties while stochastic differential equation can be defined in the absence of uncertainties [11]. A stochastic differential equation is a differential equation with one or more stochastic terms.

Consider the following ordinary differential equation:

\[
\frac{dx}{dt} = f(x, t), \quad x(t_0) = x(0), \tag{1}
\]

The differential equation 1 can be:

\[
dx(t) = f(x, t)dt \tag{2}
\]

and the integral equation can be written as:

\[
x(t) = x(0) + \int_{t_0}^{t} f(x(s), s)ds, \tag{3}
\]

where \(x(t)\) is the state of the system at time \(t\) and it is a solution with initial condition \(x(t_0) = x(0)\).

Furthermore, let us consider a stochastic differential equation (SDE) as an ordinary differential equation (ODE) with stochastic process \(X(t)\). Then consider stochastic differential equation as follows

\[
\frac{dX(t)}{dt} = f(X(t), t) + g(X(t), t)\xi(t), \quad X(t_0) = x_0. \tag{4}
\]
in which the white noise $\xi(t)$ is a stochastic process and equation 4 is known as stochastic differential equation. Due to the presence of $\xi(t)$, $X(t)$ becomes a stochastic process. In the integral form equation 4 become:

$$X(t) = X(t_0) + \int_{t_0}^{t} f(X(s), s) ds + \int_{t_0}^{t} g(X(s), s) \xi(s) ds$$  \hspace{1cm} (5)$$

Let us consider the white noise process $\xi(t)$ as a derivative of Brownian motion $W(t)$.

$$dW(t) = \xi(t) dt.$$  \hspace{1cm} (6)$$

Equation 5 can be written in the form

$$dX(t) = f(X(t), t) dt + g(X(t), t) dW(t)$$  \hspace{1cm} (7)$$

or in the integral form equation 7

$$X(t) = X(t_0) + \int_{t_0}^{t} f(X(s), s) ds + \int_{t_0}^{t} g(X(s), s) dW(s)$$  \hspace{1cm} (8)$$

is satisfied and is called the stochastic differential equation.

The Gaussian stochastic process $W(t)$ is called Brownian motion or Wiener process if

- $E(W(t_2) - W(t_1)) = 0$;
- $E((W(t_4) - W(t_3))(W(t_2) - W(t_1))) = 0, t_4 \geq t_3 \geq t_2 \geq t_1$;
- $E(W(t_2) - W(t_1))^2 = (t_2 - t_1), t_2 \geq t_1$.

We can replace $dW(t)$ by $\xi(t) dt$ where we assume that the stochastic process $\xi(t)$ is Gaussian. Furthermore, we assume that $\xi(t)$ is white noise process which means that $\xi(t)$ satisfies the following conditions

- $E(\xi(t)) = 0$
- $E(\xi(t_1)(\xi(t_2)) = \delta(t_2 - t_1), t_2 \geq t_1$

where $\delta(t)$ is the Dirac delta function.
4.2 Jabłońska-Capasso-Morale (JCM) model

This model is a combination of the models discussed first by Capasso and Morale in 2005, to which Jabłońska added a new term in order to improve model performance in a very specific financial application. The aim was to complement the existing econometric and stochastic models with a novel approach [8]. In financial markets, it has been claimed by many researchers that many human actions are related to animal behavior. The main idea of this model is related to the animal spirit behavior first discussed by Keynes in 1936 [10] that a critical fraction of the total population can drive the whole population in a specific direction. For instance, "a flock of birds cruising the sky is driven by a single leader in front" [9]. And this situation can be considered as behavior between traders in financial markets. Traders’ bids are related to their neighbourhood and if a sufficiently big subgroup of the whole population of traders is bidding far enough from the others, the others would follow that direction and the price will change dramatically at times [8].

In general form, the corresponding model reads as

$$dX^k_N(t) = [f(X^k_t) + h(k, X_t)]dt + \sigma dW^k(t), \quad (9)$$

where $f(X^k_t)$ represents the forces acting on the whole population, which in this particular application is a group of traders in the spot market and the measure of their distance is the price, and $h(k, X_t)$ is the interaction within the closest neighbors of each individual in the population of $N$ members.

Equation 9 is the Capasso-Bianchi system of stochastic differential equations used in mathematical biology for modelling animal population dynamics or price herding. The idea of using animal dynamics for modelling the trader’s prices came first as a combination of the mean-reverting jump-diffusion model with the system term in equation 9.

Hence, the model became:

$$dX^k_t = \gamma_t[(X^*_t - X^k_t) + (f(k, X_t) - X^k_t)]dt + \sigma_t dW^k_t + J^k_t dN_t + - J^k_t dN_t \quad (10)$$

for $k = 1, \ldots, N$, where

- $X^k_t$ is the price of trader $k$ at time $t$,
- $X^*_t$ is the global price reversion level at time $t$. 
\[ dX^k_t = \left[ \gamma_t (X^*_t - X^k_t) + \theta_t (h(k, X_t) - X^k_t) + \xi_t (g(k, X_t) - X^k_t) \right] dt + \sigma_t dW^k_t, \] (11)

Where

- \( h(k, X_t) \) is a global interaction,
- \( g(k, X_t) \) is a local spread.

The remaining notations are the same as in the previous model (equation 10).

Here, the new component \( g(k, X_t) \) represents the maximally distanced member of \( k \)-th trader’s neighborhood, formed by the closest \( p\% \) of the population. That is,

\[ g(k, X_t) = \max_{k \in I} \left\{ X^k_t - X_t \right\}, \] (12)

where \( I = \{k \mid X^k_t \in N^k_{p\%} \} \) and \( N^k_{p\%} \) means the neighborhood of the \( k \)-th individual formed by the closest \( p\% \) of the population.

The term \( h(k, X_t) \) presents a price level which is obtained by multiplying the difference between mean and mode of the ensemble of traders’ prices with
the mode of the ensemble traders’ prices in which both mean and mode are calculated in the time interval \( t - 1 \), hence the spikes are produced as follow, \( M(X_t)[E(X_t) - M(X_t)] \), where \( M(X_t) \) is the ensemble mode and \( E(X_t) \) is the ensemble mean.

The idea is derived from the Burgers’ equation which is a one-dimensional form of the Navier-stokes momentum equations without the pressure term and volume forces. It is widely used in different areas of applied mathematics like in modeling of fluid dynamics and traffic flow [3]. And it is presented by equation 13

\[
    u_t + \alpha uu_x + \beta u_{xx} = f(x, t) \tag{13}
\]

where

- \( u \) stands for the price and could present one dimensional measurement of fluid pressure along a periodic domain.
- \( u_x \) is the spread between any given day’s average and most common bids. In other words, the mean and the mode of the trades’ prices on a given day.
- \( uu_x \) is the momentum term that expresses traders’ movement toward the most common price, this effect is magnified at higher prices. And this term is similar with the term \( h(k, X_t) \) presents in our model see equation 11.
- \( \beta u_{xx} \) is the diffusion term that is related to the fact that the sport markets tend to reach an equilibrium price.
- \( f(x, t) \) describes the fundamentals (possibly of a periodic character).

An important note here is that the above stochastic model originated from the Ornstein-Uhlenbeck model. It is widely used for modelling mean-reverting processes that are naturally applicable for modelling commodity prices.

The process \( s \) is modelled as

\[
    ds = \lambda(\mu - s)dt + \sigma dW_t \tag{14}
\]

Where

- \( W_t \) is a Brownian-motion, so \( dW_t \sim N(0, \sqrt{dt}) \),
• \( \lambda \) measures the speed of mean reversion,
• \( \mu \) is 'long run mean', to which the process tends to revert towards the mean,
• \( \sigma \), is a measure of the process volatility.

The Ornstein-Uhlenbeck process is the solution of the above stochastic differential equation.

4.3 Kalman Dynamics (KD) model

The present approach is based on Jabłońska-Capasso-Morale (JCM) model and Variational Ensemble Kalman filter algorithm, which incorporates information of the second order (in other words, covariance) via propagating a set of particles through a nonlinear evolutionary operator, see [15]. The motivation of such an attempt originates from the nature of financial market, where psychological reaction of market participants has a certain reflection in price formation, since decision making of traders is influenced by behaviour, expectations and opinions of their partners and other participants accessible for observations. The mutual influence is not necessarily dependent on physical distance between interacting members, due to extensive availability of virtual resources and various internet services.

The objective of synthetic model developed is to simulate the real price, combining information from model output with information about correlations between decisions of the market participants accumulated in the covariance matrix of the state vector. In the present work KD model was fitted to preliminary detrended and deseasonalized original daily data, also referred to as pure trading series. The original gold prices are taken to estimate the \( H \) days moving average and the standard deviation of the KD simulated price.

To initialize the model we replicate the first value of the real price \( N \) times, hence, we obtain the vector \( u \in \mathbb{R}^{[N,1]} \), where \( N \) is the number of market participants. Initial particles \( s = (s_1, \ldots, s_{N_p}) \) constitute the set of \( N_p \) replicates of the main trajectory \( u \).
Kalman dynamics based algorithm

1. Estimate a mean reversion level from original data via Least Squares:

\[
\begin{pmatrix}
1 & x_{n-1} \\
1 & x_{n-2} \\
\vdots & \vdots \\
1 & x_{n-H}
\end{pmatrix}
\begin{pmatrix}
p_1 \\
p_2 \\
\vdots \\
p_H
\end{pmatrix}
= 
\begin{pmatrix}
dx_{n-1} \\
dx_{n-2} \\
\vdots \\
dx_{n-H}
\end{pmatrix}
\]  \hspace{1cm} (15)

where \( x_k, k = n-H, n-1 \) is the original pure trading price, \( dx_n = x_{n+1} - x_n \) is the price return. Equivalently, in vector-matrix representation the system (15) can be rewritten in the following way:

\[
X_p = dx.
\]  \hspace{1cm} (16)

We also compute the standard deviation for \( H \) days assimilation window. The latter two values (\( H \) days moving average and standard deviation) are parameters required for JCM evolutionary operator.

2. define the prior of the price increment \( \delta u^p = (\delta u^p_1, \ldots, \delta u^p_N) \), where \( N \) is the number of participants, from JCM model, which is employed as an evolutionary operator;

3. propagate the particles \( \delta s = (\delta s_1, \ldots, \delta s_{N_p}) \) via JCM operator, where each of the particles \( ds_i, i = 1, N_p \) represents price increments \( \delta u^p \) supplied with a small perturbation;

4. perform lbfgs-minimization of the quadratic cost function with respect to the vector of price increments \( \delta u^* \):

\[
l(\delta u^* | \delta u_{obs}) = \frac{1}{2}(\delta u^* - \delta u^o)^T C_p^{-1}(\delta u^* - \delta u^o) \\
+ \frac{1}{2}(\delta u_{obs} - K\delta u^o)^T C_e^{-1}(\delta u_{obs} - K\delta u^o),
\]  \hspace{1cm} (17)  \hspace{1cm} (18)

where components of artificial observational vector \( \delta u_{obs} \) are composed by the particles \( \delta s \) in the following way:

\[
(\delta u_{obs})_i = \sum_{j=1}^{N_p} \omega_{ij} \delta s_{ij}, \hspace{1cm} i = 1, N.
\]  \hspace{1cm} (19)

Here \( \delta s_{ij} \) denotes \( j-th \) component of the \( i-th \) vector-particle \( \delta s_i \), coefficients \( \omega_{ij} \) are specified as Gaussian weights \( \omega_{ij} = f(s_{ij}|u_i, \sigma) \), standard
deviation $\sigma$ is model parameter, $f$ stands for normal distribution probability density function; observational operator $K$ is an identity matrix, the inverse of observational vector error covariance $C^{-1}_{\varepsilon_o} = c_0 I_N$ is represented by diagonal matrix with scalar model parameter $c_0$ as a diagonal entity.

The inverse of the prior state vector $\delta u^*$ covariance is obtained from Sherman-Morrison-Woodbury (SMW) formula:

$$C^{-1}_p = \left(XX^T + C^{-1}_{\varepsilon_m}\right)^{-1} = C^{-1}_{\varepsilon_m} - C^{-1}_{\varepsilon_m}X \left(1 + XC^{-1}_{\varepsilon_m}X^T\right)^{-1}X^T C^{-1}_{\varepsilon_m}$$

where model error covariance $C^{-1}_{\varepsilon_m} = c_m I_N$ is diagonal, and $c_m$ is the KD model parameter. We calculate sample covariance $\bar{X}$ in the following way (see [15]):

$$\bar{X} = (\delta s_1 - \delta u^p, \delta s_2 - \delta u^p, \ldots, \delta s_N - \delta u^p) / \sqrt{N_p - 1}. \quad (21)$$

5. Sample new ensemble $\delta s^* \sim N(\delta u^*, C^*_{est})$ using low-storage representation of the covariance estimate $C^*_{est}$;

6. Increment the main trajectory $u^* = u + \delta u^*$ and the particles $s^*_i = s_i + \delta s^*_i, i = 1, N_p$;

7. $s \to s^*$, $u \to u^*$ and go to step 1.

4.3.1 KD model parameters

- $(\gamma, \theta, \xi)$ - JCM evolutionary operator parameters;
- $N_p$ - number of particles employed for filtration;
- $H$ - the length of assimilation window for computing moving average and standard deviation;
- $c_0, c_m$ - the inverse values of observation and model error variances, correspondingly. These coefficients define how strongly we trust observations and model. When model error variances are small in comparison to observational error variances, decisions of the traders are affected by change expectations of surrounding agents greater than by rational reasons based on previous price dynamics.
- $\sigma$ - observational operator parameter; the bigger $\sigma$ is, the more deviant are expectations, taken into account by market agents.
5 Gold price modelling results

This section provides the model fitting results of this work. The whole downloaded data discussed in section 3 has been used for each model separately and for the comparison of the two models to apply simulation analysis. Original versus simulated data have been plotted in the same figure such as time plot and histograms, ACFs, PACFs of each and comparison of basic statistic values for original versus simulated data, to capture the performance of those models.

5.1 JCM results

Firstly, Jabłońska-Capasso-Morale (JCM) results are discussed. Figure 9 presents time line plot for original price in blue color together with simulated price in red color. Clearly, there is a resemblance between the simulated and original gold price. However, in the overall JCM model fails to follow precisely the original price in some periods which are observed where the price looks very high before and after 8000th observation.

![Time line plot for original and simulated gold price (Jabłońska-Capasso-Morale model).](image)

Figure 9: Time line plot for original and simulated gold price (Jabłońska-Capasso-Morale model).

Autocorrelation figures for original and simulated gold price are drawn in figure 10 and figure 11. There is no difference in ACFs where both price series have positive and negative ACF coefficients.
5 GOLD PRICE MODELLING RESULTS

The PACFs figures are presented in Figures 12 and 13. There is very high significant coefficient at lag one and the remaining PACF values lie approximately within the confidence limits. Even though there are some coefficient values standing out of significance limits especially for simulated price, that level of autocorrelation could be considered as noise.

The price was also presented in the form of histogram and is shown in figure 14 to see the distribution of the original and simulated price. The same features can be observed for original and simulated prices, in both cases there is a

![Figure 10: ACF for original gold price.](image)

![Figure 11: ACF for simulated gold price (Jabłońska-Capasso-Morale model).](image)
positive skewed and very high peak but the original price has a very high peak and is more skewed than simulated price. This result can be supported by the values of the basic statistics summarized in table 2.

Mean, standard deviation, and five consecutive central moments that are kurtosis, skewness, and $5^{th}$, $6^{th}$ and $7^{th}$ moment have been used for basic statistic comparison of original and simulated prices. Values of these moments in the simulated price histogram are close to the corresponding values of moments of the real price histogram. Only $5^{th}$ and $7^{th}$ moments are much higher for

Figure 12: PACF for original gold price.

Figure 13: PACF for simulated gold price (Jabłońska-Capasso-Morale model).
original price than for simulated price by JCM model.

Therefore, after the analysis and interpretation of the figures, we can say that the Jabłońska-Capasso-Morale model fairly captures the dynamics of gold price since simulated results and basic statistical comparison shows reasonable resemblance.

![Figure 14](image)

Figure 14: Histograms for original and simulated gold price (Jabłońska-Capasso-Morale model).

Table 2: Statictical properties of real price and simulated price (JCM model).

<table>
<thead>
<tr>
<th></th>
<th>real</th>
<th>JCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>399.95</td>
<td>384.37</td>
</tr>
<tr>
<td>St.Dev</td>
<td>202.84</td>
<td>174.01</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.54</td>
<td>2.22</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>9.65</td>
<td>8.66</td>
</tr>
<tr>
<td>5th moment</td>
<td>36.29</td>
<td>30.61</td>
</tr>
<tr>
<td>6th moment</td>
<td>143.89</td>
<td>120.42</td>
</tr>
<tr>
<td>7th moment</td>
<td>583.94</td>
<td>473.60</td>
</tr>
</tbody>
</table>
5.2 KD results

Secondly, Kalman Dynamics (KD) results are presented. In Figure 15, the original price and simulated price are plotted together. It can be seen that the simulation nicely follows the original data especially in terms of spikes.

Figure 15: Time line plot for original and simulated gold price (Kalman Dynamics model).

Figure 16 shows ACFs for original price and for simulated price, one can say that they are similar. Also, in Figure 17 the PACFs are similar each having one significant coefficient at lag one.

The original and simulated histograms are similar in Figure 18. Hence, it is necessary to find the difference by comparing the basic statistic values, that is mean, standard deviation, skewness and kurtosis values in Table 3.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St.Dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>real</td>
<td>0.0001</td>
<td>57.7299</td>
<td>0.2759</td>
<td>3.3436</td>
</tr>
<tr>
<td>KD</td>
<td>-0.3814</td>
<td>62.4685</td>
<td>0.2375</td>
<td>3.3524</td>
</tr>
</tbody>
</table>

From Table 3 we observe that especially skewness and kurtosis for simulated price are having values very close to the real price. The standard deviation is
Figure 16: ACFs for original and simulated gold price (Kalman Dynamics model).

Figure 17: PACFs for original and simulated gold price (Kalman Dynamics model).

still comparable. But the mean is negative for the simulated price distribution whereas for the real price it is zero.

The series oscillate around mean zero and their variances vary over time in Figure 19 for real gold price return and simulated gold price return. As shown in Figure 20 and Figure 21, the autocorrelation and partial autocorrelation functions are similar for original and simulated data.
Figure 18: Histograms for original and simulated gold price (Kalman Dynamics model).

Figure 19: Time line plot for real and simulated gold price return (Kalman Dynamics model).

Simulation results show remarkable resemblance. These are observed by comparing figures and numerical values of basic statistics for real price and simulated ones. Hence, simulation using Kalman Dynamics model reconstructs the features of original gold price.
Figure 20: ACFs for real and simulated gold price return (Kalman Dynamics model).

Figure 21: PACFs for real and simulated gold price return (Kalman Dynamics model).
5.3 Comparison of model performance

In this subsection, the study of gold price dynamics using two models (KD and JCM model) is presented in order to compare the performance of these models. It is done through a statistical comparison of the simulated series.

In Figure 22, time line plot results for both original price and simulated price by KD and JCM model are presented. It can observed that both models follow very well the original price and among these two model it is difficult to distinguish which model is better.

![Figure 22: Time line plot for original and simulated price (KD and JCM model).](image)

Figure 23 shows autocorrelation figures. The first figure is for original price, the second figure is the simulated price by the JCM model and the last one is the simulated price by the KD model. Similar features in autocorrelation disposition can be seen when comparing the JCM model and the KD model to the original price.

![Figure 23: Autocorrelation figures.](image)

The partial autocorrelation figures are presented by Figure 24, arranged as follows: plot one is for original price, plot two is for simulated price by the KD model and the last one is for simulated price by the JCM model. This situation is similar when one considers the position of first lag. For the simulated price there are a number of coefficients lying out of significance level but simulated price by JCM model is closer to the original price since the coefficient values approximately lie within significance level than in KD model when JCM model...
and KD model is compare to the original price.

Simulated price histograms by KD and JCM models are similar to the original price and are shown in Figure 25, arranged in the same order as Figure 24. The histogram is right skewed for the original price and for the simulated price. When the simulated price of the two models is compared to the original price, it is difficult to tell which model is better since both models have similar features.

Figure 23: ACFs for original and simulated price (KD and JCM model).

Figure 24: PACFs for original and simulated price (KD and JCM model).
The basic statistics for this comparison analysis are in table 4 and close numerical values between original and simulated prices by JCM and KD model can be seen especially for kurtosis and skewness. Hence, for comparison of the original and the simulated price of the two models, one can not say that a particular model produces values closer to the original price in each moment of distribution. This result is supported by histogram figures.

Table 4: Statistical properties of real and simulated price (JCM and KD model).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St.Dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>real</td>
<td>0.0001</td>
<td>57.7299</td>
<td>0.2759</td>
<td>3.3436</td>
</tr>
<tr>
<td>JCM</td>
<td>-0.4136</td>
<td>52.1309</td>
<td>0.2819</td>
<td>3.6976</td>
</tr>
<tr>
<td>KD</td>
<td>-0.3814</td>
<td>62.4685</td>
<td>0.2375</td>
<td>3.3524</td>
</tr>
</tbody>
</table>

An analysis of real price returns and simulated price returns by KD and JCM has been conducted. Figure 26 is the result in time line plots for real price returns in green, simulated price returns by the JCM model in blue and simulated price returns by KD in red. Both simulated price return produced similar features and it is difficult to tell which model is better when real price return is compare to the simulated price return for JCM model and KD model.

Figure 27 are autocorrelation plots for real price returns, simulated price return.
returns by JCM model and simulated price returns by KD model. For both simulated price returns, there are number of coefficients out of significance level but they are not well identified and autocorrelation for simulated price return by JCM is better when real price returns are compared to the simulated price returns of the two models, since the autocorrelation coefficients lie approximately within confidence limit than for simulated price return by KD model.

These price return are presented also in the form of partial autocorrelation function. In Figure 28 there are simulated price return PACF by KD model, simulated price return PACF by JCM model and real price return PACF. There is no big difference when real price return is compared to simulated price return PACF for both model.

The Root Mean Square Error (RMSE), also called the root mean square deviation (RMSD), plots are also presented and RMSE values have been used to distinguish the performance of the two models. In this case, the RMSE is the absolute value between the real price and simulated price and was calculated as follow:

\[ [\text{rmse}]_k = |x_{k}^{\text{mo}} - x_{k}^{\text{true}}| \]  

(22)

Therefore, the comparison of RMSE plots for simulated price by JCM model

![Figure 26: Time line plot for real and simulated return price (KD and JCM model).](image)
in blue and KD model in red are presented by Figure 29, one can say that JCM model produced small RMSE value since the highest value is around 120 where RMSE values for KD model is almost 150.

Figure 30 shown RMSE histogram and it can observed that KD model is more left skewed than JCM. Those statements are supported by RMSE properties for KD and JCM model simulated price given in Table 5 those are mean, minimum, maximum and standard deviation.

From Table 5 mean and maximum value for JCM model is smaller than for KD model. But minimum and standard deviation is small for KD model.

Table 5: RMSE properties for KD and CJM simulated prices.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>min</th>
<th>max</th>
<th>st.dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>KD</td>
<td>8.0799</td>
<td>0</td>
<td>164.8958</td>
<td>10.2184</td>
</tr>
<tr>
<td>JCM</td>
<td>7.3249</td>
<td>0.5632·1.0e-003</td>
<td>148.2047</td>
<td>10.7690</td>
</tr>
</tbody>
</table>

Figure 27: ACFs for real and simulated return price (KD and JCM model).
After the study of these figures, we can observe that there is a strong relationship between the JCM model and the KD model, since simulated prices by both models capture the movement dynamics of real prices. In most cases

Figure 28: PACFs for real and simulated return price (KD and JCM model).

Figure 29: RMSE time line plots for simulated price (KD and JCM model).
it was very difficult to say whether a particular model performed better when
the JCM model was compared to the KD model, and basic statistics for these
prices can also support this statements due to the closeness of the statistical
values of real prices and simulated prices of the two models. There is no big
difference between those three prices.

6 Results summary and discussion

The results of this study is in three parts, one is Jabłońska-Capasso-Morale
(JCM) results, second is Kalman Dynamics (KD) results and third one is the
comparison of them. The results are related to the simulation analysis of gold
prices using two models and the analysis done in terms of figures and tables
of basic statistics summary in which simulated prices and original prices were
compared. Obtained results were done by MATLAB software and maximum
likelihood method has been used to calibrate model parameters to the data.

Figures 9-14 were for original and simulated prices by JCM model and Table
2 is statistical properties for the two prices. Using KD model the simulated
and original price results were presented by Figure 15-21 and their statistical
properties were summarized in Table 3. For comparison of model performance
the similar analysis was done. Figure 22-30 were the results for original and
simulated price by JCM and KD model, 4 and 5 is statistical and RMSE

Figure 30: RMSE histograms for simulated price (KD and JCM model).
properties of two models.

Generally, the results have shown that both models were able to capture the
dynamics of gold price since simulated prices follow original price and this was
supported by the closeness of statistical distributions of original and simulated
prices. The same ability of the two models was found thus there is insignificant
difference between three prices those are original and simulated price by JCM
and KD model.

7 Conclusion

Jabłońska-Capasso-Morale (JCM) and Kalman Dynamics (KD) models have
been used to simulate the dynamics of gold price in spot market and the
comparison of them for model performance.

The study started by giving the reader an introduction to the gold market and
the presentation of data used in this study and was done through graphical
presentations and basic statistic values to capture the important features and
it was observed that the series are not stationary and not normally distributed.
However, the models used in this study did not require stationarity as well as
normality of the series, since they use the concept of moving mean reversion.

Furthermore, by using those two mathematical models it has been observed
that both models capture the dynamics of gold market, since simulated results
by JCM and KD model are very close to the original prices. There was no
significant difference when JCM model was compared to KD model thus both
models produced similar results and one can say that two models have the
same capacity to follow the movement in real data.

From the obtained results it implies that the assumption about human psy-
chology and market momentum may be true in gold market which is related
to the JCM model idea and the fact that KD use JCM model.
References


REFERENCES


List of Tables

1. Basic Statistics of gold price ........................................... 16
2. Statistical properties of real price and simulated price (JCM model). ........................................... 28
3. Statistical properties of real and simulated price (KD model). ........................................... 29
4. Statistical properties of real and simulated price (JCM and KD model). ........................................... 35
5. RMSE properties for KD and CJM simulated prices. ........................................... 37
List of Figures

1  Time line plot of gold price. ........................................ 12

2  (a) ACF of gold price. (b) PACF of gold price. .................. 12

3  (a) Histogram of gold price. (b) Normal histogram of gold price. 13

4  (a) Time line plot of gold price difference transformation. (b) Time line plot of gold price log difference transformation. .... 14

5  (a) ACF of gold price difference transformation. (b) PACF of gold price difference transformation. .......................... 14

6  (a) ACF of gold price log difference transformation. (b) PACF of gold price log difference transformation. ..................... 15

7  (a) Histogram of gold price difference transformation. (b) Normal histogram of gold price difference transformation. ....... 15

8  (a) Histogram of gold price log transformation. (b) Normal histogram of gold price log transformation. ...................... 16

9  Time line plot for original and simulated gold price (Jabłońska-Capasso-Morale model). ................................. 25

10 ACF for original gold price. ........................................... 26

11 ACF for simulated gold price (Jabłońska-Capasso-Morale model). .................................................. 26

12 PACF for original gold price. ........................................... 27

13 PACF for simulated gold price (Jabłońska-Capasso-Morale model). .................................................. 27

14 Histograms for original and simulated gold price (Jabłońska-Capasso-Morale model). ................................. 28

15 Time line plot for original and simulated gold price (Kalman Dynamics model). .......................................... 29

16 ACFs for original and simulated gold price (Kalman Dynamics model). .................................................. 30
17 PACFs for original and simulated gold price (Kalman Dynamics model). .................................................. 30
18 Histograms for original and simulated gold price (Kalman Dynamics model). ........................................ 31
19 Time line plot for real and simulated gold price return (Kalman Dynamics model). ................................... 31
20 ACFs for real and simulated gold price return (Kalman Dynamics model). ................................................ 32
21 PACFs for real and simulated gold price return (Kalman Dynamics model). ........................................... 32
22 Time line plot for original and simulated price (KD and JCM model). ...................................................... 33
23 ACFs for original and simulated price (KD and JCM model). ... 34
24 PACFs for original and simulated price (KD and JCM model). ... 34
25 Histograms for original and simulated price (KD and JCM model). 35
26 Time line plot for real and simulated return price (KD and JCM model). ................................................. 36
27 ACFs for real and simulated return price (KD and JCM model). 37
28 PACFs for real and simulated return price (KD and JCM model). 38
29 RMSE time line plots for simulated price (KD and JCM model). 38
30 RMSE histograms for simulated price (KD and JCM model). ... 39