Received: 22 July 2022 | Revised: 10 January 2023 | Accepted: 25 January 2023

DOI: 10.1002/mma.9115

SPECIAL ISSUE PAPER

WILEY

# Estimating and pricing commodity futures with time-delay stochastic processes

**Lourdes Gómez-Valle** | **Julia Martínez-Rodríguez**

Departamento de Economía Aplicada e IMUVA, Facultad de Ciencias Económicas y Empresariales, Universidad de Valladolid, Valladolid, Spain

**Correspondence**
Julia Martínez-Rodríguez, Departamento de Economía Aplicada e IMUVA, Facultad de Ciencias Económicas y Empresariales, Universidad de Valladolid, Valladolid, Spain.
Email: juliamr@uva.es

Communicated by: J. C. Cortés

**Funding information**
Spanish Agencia Estatal de Investigación, Grant/Award Number: PID2020-113554GB-I00/AEI/10.13039/501100011033; Junta de Castilla y León, Spain and European FEDER Funds, Grant/Award Number: VA193P20; University of Valladolid, Grant/Award Number: GIR Optimización Dinámica, Finanzas Matemáticas y Utilidad Recursiva

In commodity futures pricing models, the commodity present price is generally considered to reflect all information in the markets and past information is not regarded important. However, there is some empirical evidence that shows that this fact is unrealistic. In this paper, we consider some stochastic models with delay for pricing commodity futures. The functions of the commodity price stochastic process under the risk-neutral measure are necessary for pricing derivatives. However, the observations in the market have risk. Then, we use a technique that allows us to estimate the functions of the risk-neutral commodity spot price stochastic process, directly from futures prices traded in the market, and show how to price the commodity futures. Finally, we make an empirical application of this methodology with gold futures traded in the COMEX. Furthermore, we make clear the supremacy of the delay models in pricing gold futures.

**KEYWORDS**

commodity futures prices, delay stochastic processes, derivative securities, nonparametric estimation, PDEs with randomness

**MSC CLASSIFICATION**

35R60, 91G30, 91G60

## 1 | INTRODUCTION

In the derivative pricing literature, it is very common to assume that the efficient market hypothesis (EMH) is satisfied (see Fama[1] and Titan[2]). This hypothesis establishes that the present underlying asset price reflects all the available information and that no information on its history can improve the prediction of its future behavior; see Kemajou.[3] However, some empirical studies indicate that there exists some dependence on past returns; see, for example, Bernard and Thomas[4] and Scheinkman and LeBaron[5] for stock prices or Küchler and Platen[6] for commodities. Consequently, many scholars include a delay in these financial models, for example, it is considered for pricing options (see previous studies[7–11]), for modeling and pricing swaps (see Swishchuk and Xu[12]) and for bond valuation (see Flore and Nappo[13]). Furthermore, Chunxiang and Shao[14] model the dynamics of the wealth by a stochastic differential delay equation in a portfolio optimization problem, Benhabib[15] considers a model where the nominal interest rates are measured by a flexible distributed delay, and Kemajou-Brown et al.[16] apply it to pricing debt and loan guarantees.

In all the above-mentioned papers, the authors show that the past dependence of the underlying asset processes is an important feature in order to get the best prediction of the pricing process.

International markets for commodity derivatives have experienced a prominent growth in recent years. This makes the study of commodity futures pricing of especial interest in the present context. However, the behavior of the commodities is, in general, different from conventional financial assets. For example, prices may temporary be low or high but will tend toward an equilibrium level (see Kyriakou et al.[17]), and its volatility is both relatively higher and more variable over time (see Pyndick[18]).

There have been numerous studies looking at the pricing of commodity futures. In some cases, the geometric Brownian motion has been considered to model the commodity prices; see, for example, Brennan and Schwartz[19] and Gibson and Schwartz.[20] However, due to the interaction of demand and supply, a mean-reverting behavior usually arises in the commodity market, as in Schwartz[21] and Schwartz and Smith.[22] Some other authors consider a jump-diffusion process to take into account the usual abrupt changes that commodity prices suffer in the markets; see Gómez-Valle et al.[23] and Hilliard and Reis,[24] and Yan[25] also considers stochastic volatility. It is also widely showed that supply and demand of most commodities follow seasonal cycles, mainly agricultural and energetic products; see previous studies[26–28] among others.

The aim of this paper is to provide the necessary methodology to implement one-factor stochastic models with delay for pricing commodity futures (in both phases: estimation and valuation). This work shares some similarities with recent contributions in the financial literature focused on other markets such as, for example, stock options or bonds (see Arriojas et al.[8] and Flore and Nappo[13]).

In more detail, in this paper, we consider the effect of the past in the commodity futures prices. In particular, we assume that the commodity price satisfies a stochastic delay differential equation (SDDE) with a fixed delay on the volatility. Using a self-financed strategy and replication (following Arriojas et al.[8] and Kemajou-Brown et al.[16]), we establish that the commodity futures price follows a random partial differential equation (RPDE) whose coefficients are the drift and the volatility of the spot price under the well-known risk-neutral measure. As the RPDE solution is usually unknown and there are no observations of the spot price under the risk-neutral measure, its drift cannot be estimated. Hence, we prove a result, which allows us to estimate it using the prices of commodity futures traded in the markets. This result is model independent and versatile and can be applied to any one-factor commodity futures stochastic delay model. Moreover, we complement these results with an empirical illustration using gold futures traded in the Commodity Exchange Inc. (COMEX), CME-group.

The structure of this paper is as follows. Section 2 describes the futures pricing with delay setup and proposes different commodity futures pricing models. Section 3 proves a result that allows us to estimate the risk-neutral drift of the commodity spot price and an interesting approach to price commodity futures. Section 4 describes the available data and the obtained empirical results. Section 5 concludes.

## 2 | STOCHASTIC DELAY MODEL AND RPDE

In this section, we present a stochastic delay model for pricing commodity futures and the RPDE that any commodity futures price should satisfy. Then, we introduce some parametric and nonparametric models with delay.

The need for improving the understanding of the behavior of the financial variables has motivated the development of dynamic models that take into consideration the influence of past events on the current and future prices. In fact, some authors such as Kazmerchuk et al.[29] and Bernard and Thomas[4] establish that, in the market, a portion of the price response to new information is delayed.

It is also known that the assumption of a constant volatility in commodity prices is unrealistic, since empirical evidence shows that the volatility depends on time; see, for example, Pyndick.[18] Then, it seems natural to consider a commodity spot price process where volatility can be regarded as a function of past states, that is, has a high self-reinforcement; see Tambue et al.[30]

Suppose that $(\Omega, \mathcal{F}, \mathcal{P})$ is a given probability space equipped with a filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ satisfying the usual conditions; see Oksendal and Sulem[31] and Protter.[32] We consider that the commodity spot price $S$ follows a SDDE:

$$dS(t) = \mu(S(t))\,dt + \sigma(S(t), S(t-d))\,dW(t), \;\; t \in (0, T], \tag{1}$$

$$S(t) = \phi(t), \;\; t \in [-d, 0], \tag{2}$$

where $d$ is a positive constant, a fixed delay, which is incorporated in the volatility term in order to take into account the past events on the current and future states. Moreover, $\mu$ and $\sigma$ are deterministic continuous functions and are, respectively, the drift and volatility of the process, satisfying suitable regularity conditions; see Oksendal and Sulem.[31] These functions depend on the spot price in two instants of time. Moreover, $W$ is a Wiener process, and $\phi$, in (2), is a deterministic continuous function, $\phi \in C([-d, 0], \mathbb{R})$, which represents previous known price information.

Commodity futures are derivative financial contracts that obligate the buyer to purchase and the seller to sell a commodity at a predetermined future date (delivery date) and price (delivery price); see Kwok.[33] Then, the futures price is the value of the delivery price that makes the futures contract has no-arbitrage price zero at initial time, and therefore, no initial investment is required for the futures contracts. These contracts are very important in the markets not only for hedging risk but also for speculating.

A commodity futures price at time $t$ with maturity at time $T$, $t \leq T$, can be expressed under the above assumptions as $F(t, S; T)$, and at maturity, it verifies

$$F(T, S; T) = S. \tag{3}$$

That is, its value must be equal to the commodity spot price at that time.

Following, in order to price this financial derivative, we introduce its valuation equation by means of the non-arbitrage argument.

Consider a riskless strategy $\Pi$, which consists of holding $\alpha_1$ futures contracts expiring at $T_1$ and $\alpha_2$ futures contracts expiring at $T_2$; see Gabillon[34]:

$$\Pi = \alpha_1 F(t, S; T_1) + \alpha_2 F(t, S; T_2).$$

In the absence of arbitrage opportunities, as no initial investment is required for the futures contract (see Shreve[35]), its instantaneous return must be equal to zero:

$$d\Pi = \alpha_1 dF(t, S; T_1) + \alpha_2 dF(t, S; T_2) = 0.$$

As $S$ follows the SDDE (1), we apply Ito's formula (see Arriojas et al.[8] and Kemajou-Brown et al.[16]) and deduce

$$\begin{aligned}
&\alpha_1 \left( F_t(t, S; T_1) + \mu(S(t)) F_S(t, S; T_1) + \frac{1}{2}\sigma^2(S(t), S(t-d)) F_{SS}(t, S; T_1) \right) dt \\
&+ \alpha_1 \sigma(S(t), S(t-d)) F_S(t, S; T_1) \, dW(t) \\
&+ \alpha_2 \left( F_t(t, S; T_2) + \mu(S(t)) F_S(t, S; T_2) + \frac{1}{2}\sigma^2(S(t), S(t-d)) F_{SS}(t, S; T_2) \right) dt \\
&+ \alpha_2 \sigma(S(t), S(t-d)) F_S(t, S; T_2) \, dW(t) = 0,
\end{aligned}$$

where $F_t$ and $F_S$ represent the derivative of $F$ with respect to time and spot price, respectively, and $F_{SS}$ is the second derivative with respect to the spot price. As the deterministic terms must be equal

$$\begin{aligned}
&\alpha_1 \left( F_t(t, S; T_1) + \mu(S(t)) F_S(t, S; T_1) + \frac{1}{2}\sigma^2(S(t), S(t-d)) F_{SS}(t, S; T_1) \right) \\
&= \alpha_2 \left( F_t(t, S; T_2) + \mu(S(t)) F_S(t, S; T_2) + \frac{1}{2}\sigma^2(S(t), S(t-d)) F_{SS}(t, S; T_2) \right),
\end{aligned}$$

as well as the stochastic terms

$$\begin{aligned}
&\alpha_1 \sigma(S(t), S(t-d)) F_S(t, S; T_1) \\
&= \alpha_2 \sigma(S(t), S(t-d)) F_S(t, S; T_2),
\end{aligned}$$

hence, there must exist a function $\lambda(t, S(t), S(t-d))$ independent of maturity, commonly known as market price of risk, such that

$$\begin{aligned}
\lambda(t, S(t), S(t-d)) &= \frac{F_t(t, S; T_1) + \mu(S(t)) F_S(t, S; T_1) + \frac{1}{2}\sigma^2(S(t), S(t-d)) F_{SS}(t, S; T_1)}{\sigma(S(t), S(t-d)) F_S(t, S; T_1)} \\
&= \frac{F_t(t, S; T_2) + \mu(S(t)) F_S(t, S; T_2) + \frac{1}{2}\sigma^2(S(t), S(t-d)) F_{SS}(t, S; T_2)}{\sigma(S(t), S(t-d)) F_S(t, S; T_2)}.
\end{aligned}$$

That is, the market price of risk does not depend on the maturity of the futures.

Moreover, in this paper, we also assume it does not depend directly on the time $t$. As a result, we obtain that the futures price is the solution of this RPDE

$$F_t + (\mu(S) - \lambda(S, S(t-d))\sigma(S, S(t-d)))\, F_S + \frac{1}{2}\sigma^2(S, S(t-d))F_{SS} = 0, \tag{4}$$

with the final condition (3), where $S(t-d)$ follows a stochastic process. Note that the coefficient of $F_S$ is the drift of the stochastic process $S(t)$ under the risk-neutral measure $\mathcal{Q}$ (in fact, the price is a martingale under this measure), that is, $\mu^{\mathcal{Q}}(S(t), S(t-d)) = \mu(S(t)) - \lambda(S(t), S(t-d))\,\sigma(S(t), S(t-d))$ and the futures price is given by the conditional expectation

$$F(t, S; T) = E^{\mathcal{Q}}\left[S(T)\,|\,\mathcal{F}_t\right], \tag{5}$$

where

$$dS(t) = \mu^{\mathcal{Q}}(S(t), S(t-d))\,dt + \sigma(S(t), S(t-d))\,dW^{\mathcal{Q}}(t). \tag{6}$$

Therefore, for pricing commodity futures, we have to estimate the drift* in (6) under this measure $\mathcal{Q}$, but not under the physical measure $\mathcal{P}$ in (1).

In order to describe the dynamics of the commodity spot price, first of all, we assume some parametric delay stochastic processes which are modifications of others considered in the literature. In fact, we introduce the delay in the volatility of two very well-known processes. On the one hand, we assume that the functions $\mu$ and $\sigma$ in (1) are given by

$$\mu(S(t)) = \mu_1 S(t), \tag{7}$$

$$\sigma(S(t), S(t-d)) = \sigma_1 S(t-d)S(t), \tag{8}$$

where $\mu_1$ and $\sigma_1$ are constants. This is a modification of the geometric Brownian process which has already been used in some asset and commodity derivatives models; see Brennan and Schwartz[19] and Gibson and Schwartz.[20] This delay process has also been used to price several derivatives; see, for example, previous studies.[11,12,29] We denote it as delay geometric (**DG**).

On the other hand, we consider the mean reversion in the drift as in Vasiceck,[36] which is also an important feature in the commodity markets, and assume the delay in the volatility:

$$\mu(S(t)) = k(\mu_2 - S(t)), \tag{9}$$

$$\sigma(S(t), S(t-d)) = \sigma_2 S(t-d), \tag{10}$$

where $k$, $\mu_2$, and $\sigma_2$ are constant. This process has been previously employed to price zero-coupon bonds in Flore and Nappo.[13] We denote it as delay mean reversion (**DMR**) model.

In order to obtain the futures prices with both parametric models, we replace (7)–(8) or (9)–(10) into the valuation RPDE (4). In the **DG**, we have

$$F_t + (\mu_1 - \lambda_1\sigma_1 S(t-d))\, SF_S + \frac{1}{2}\sigma_1^2 S^2(t-d)S^2 F_{SS} = 0, \tag{11}$$

and in the **DMR**,

$$F_t + (k(\mu_2 - S) - \lambda_2\sigma_2 S(t-d))\, F_S + \frac{1}{2}\sigma_2^2 S^2(t-d)F_{SS} = 0, \tag{12}$$

where $\lambda_1$, $\lambda_2$ are assumed to be constant. In next sections, we will discuss its implementation to obtain the futures prices.

Finally, in this work, we also consider a nonparametric model (from now on, we will denote it as **Nonpar**). In this case, instead of imposing an arbitrary behavior to the commodity spot price, we use nonparametric estimation techniques to obtain the functions of the process. This type of models, but without delay, have been used to value different derivatives; see, for example, Gómez-Valle and Martínez-Rodríguez[37] for commodity futures, Stanton[38] for zero-coupon bonds, or Gómez-Valle et al.[39] for forward freight agreements. However, this estimation technique has not yet been applied to delay processes.

*Although the delay is not taken into account in the drift of the commodity spot price, in (1), the change of measure incorporates it in the risk-neutral drift.

## 3 | ESTIMATION AND VALUATION

In this section, we show how to estimate the risk-neutral functions of the different models and price commodity futures, considering the previous parametric and nonparametric models.

In order to price these derivatives with any of the delay models considered in Section 2, the risk-neutral drift and volatility must be previously obtained.

In the literature, the risk-neutral drift is usually estimated with the futures closed-form expression of the corresponding model jointly with spot and futures price data. However, as the solution of (4) is usually unknown and there are no observations under the risk-neutral measure, we cannot use (proceed with) this approach. Therefore, we provide a result that allows us to estimate the risk-neutral drift using prices of futures traded in the market.

**Theorem 1.** *Let $S$ be the spot price, which follows the delay stochastic process (6) and $F(t, S; T)$ the price of the future (5). Then,*

$$\frac{\partial F}{\partial T}(t, S(t); T) = E^{Q}\left[\mu^{Q}(S(T), S(T-d)) \,|\, \mathcal{F}_t\right]. \tag{13}$$

*Proof.* We consider the integral form of (6)

$$S(T+h) - S(T) = \int_{T}^{T+h} \mu^{Q}(S(z), S(z-d))dz + \int_{T}^{T+h} \sigma(S(z), S(z-d))dW(z). \tag{14}$$

Taking into account (5) and that the expected value of the Ito integral term is zero, if we take the expectation over the integral form (14) with respect to $Q$ measure, we obtain

$$F(t, S; T+h) - F(t, S; T) = \int_{T}^{T+h} E^{Q}\left[\mu^{Q}(S(z), S(z-d)) \,|\, \mathcal{F}_t\right] dz. \tag{15}$$

Then, dividing (15) by $h$ and taking limits, we obtain (13). □

This result avoids assuming that the market price of risk is zero, that is, the physical and the risk-neutral measure are equal, which is unrealistic and leads to misspecification. Similar results are shown in Gómez-Valle et al.[23], for pricing commodity futures with jump-diffusion processes, and Gómez-Valle et al.[39], for one-factor freight models.

In order to implement this result, we apply numerical differentiation. This fact allows us to consider futures prices with different maturities for its estimation. More precisely, we approximate terms in (13) in the following way: the expectation on the right-hand side by $\mu^{Q}(S(t), S(t-d))$ and the derivative on the left-hand side by a well-known fourth-order difference formula:

$$\mu^{Q}(S(t), S(t-d)) \simeq \frac{3F(t, S(t); T+2h) + 8F(t, S(t); T+h) - 8F(t, S(t); T-h) + F(t, S(t); T-2h)}{12h}, \tag{16}$$

where $T$, for example, can be chosen as 4 months and the increment $h$ as 1 month. Hence, four maturities would be involved in the estimation.

On the one hand, in **DG** and **DMR** models, we apply least squares to (16) for estimating the parameters of $\mu^{Q}$. On the other hand, in order to avoid arbitrary parametric restrictions on the functions of the stochastic process, we consider nonparametric estimation techniques. In this case, we approximate the function $\mu^{Q}(S, \eta)$ on (6) ($S$ and $\eta$ represent $S(t)$ and $S(t-d)$, respectively) by means of the Theorem 1 with (16) and the Nadaraya–Watson nonparametric estimator. This estimator works as follows. We consider a data set of $N$ observations, $(S_1, \eta_1, Z_1), \ldots, (S_N, \eta_N, Z_N)$, where $(S_i, \eta_i)$ are the explanatory variables and $Z_i$ is the response variable (in our case, the right-hand side of 16). We assume a model as $Z_i = f(S_i, \eta_i) + \epsilon_i$, where $f(S_i, \eta_i)$ is an unknown function and $\epsilon_i$ is a random error term in the observations, which are random independent variables and identically distributed with mean 0 and finite variance. The Nadaraya–Watson estimator has the form

$$\hat{f}(S, \eta) = \frac{\sum_{i=1}^{N} \omega_i(S, \eta) Z_i}{\sum_{i=1}^{N} \omega_i}, \tag{17}$$

where $\omega_i(S,\eta) = K\left(\frac{S-S_i}{h_S}\right)K\left(\frac{\eta-\eta_i}{h_\eta}\right)$ are weight functions, $K$ is the kernel (more specifically, we use the multivariate Gaussian kernel), and $h_S$ and $h_\eta$ the bandwidths; see Hardle[40] for more details about this estimation technique.

As far as the volatility term in (1) is concerned, as it is not affected by the change of measure, we approximate it by means of the Euler discretization, the infinitesimal generator and observations of the spot price in the market; see Oksendal and Sulem[31] and Gómez-Valle and Martínez-Rodríguez[37]:

$$\sigma^2(S(t), S(t-d)) = \frac{1}{\Delta}E\left[(S(t+\Delta) - S(t))^2 \,|\, \mathcal{F}_t\right] + \mathcal{O}(\Delta), \tag{18}$$

where we consider equidistant time points $\{t_i\}_{i=1}^n$ and increments $\Delta = 1/252$, that is, daily data.

In parametric models, **DG** and **DMR**, we estimate the parameters of the volatility with the method of least squares. However, for the **Nonpar** model, we use Nadaraya–Watson estimator (17) with $Z_i = (S_{i+1} - S_i)^2$.

To calculate the futures prices, we have to solve the RPDE (4) with the final condition (3), whose explicit solution is unknown, or obtain the expectation in (5).

In the parametric models, we proceed as follows. First, we calculate a high number of simulations of the stochastic delay equation (1) with the functions (7)–(8) or (9)–(10), depending on the model, by means of Euler discretization. Then, we substitute the values of these trajectories in the coefficients of Equation (4). As these coefficients are now deterministic, the obtained PDE can be solved. Therefore, the commodity futures prices are approximated with the means of the prices obtained by solving the PDE for each simulation. That is, if $F^i(t,S;T)$ is the price calculated from simulation $i$, $i = 1, \dots, M$, as solution to the PDE, then

$$F(t,S;T) = \frac{1}{M}\sum_{i=1}^{M} F^i(t,S;T).$$

This approach has already been used for pricing other derivatives; see Agrawal and Hu[7] and Tambue.[30]

Next, we describe the procedure we use to approximate the futures prices from the PDE for the parametric models proposed in this paper.

- **DG** model

  We assume that the expression of the price function in (11) is

$$F(t,S;T) = A(t)S + B(t), \tag{19}$$

where $A$ and $B$ are functions of time. Then, on the one hand, from the final condition (3), we get $A(T) = 1$ and $B(T) = 0$, and, on the other hand, replacing (19) in (11), we obtain the following system of ordinary differential equations:

$$\begin{aligned} &A' + (\mu_1 - \lambda_1\sigma_1 S(t-d))A = 0,\\ &B' = 0. \end{aligned}$$

Solving this system, we obtain $B(t) \equiv 0$ and $A(t) = e^{\mu_1(T-t) - \lambda_1\sigma_1\int_t^T S(z-d)\,dz}$. Therefore, the futures price is given by

$$F(t,S;T) = Se^{\mu_1(T-t) - \lambda_1\sigma_1\int_t^T S(z-d)\,dz}. \tag{20}$$

- **DMR** model

  We assume the price function (19), as in the **DG** model, and replace it in (12). The functions $A$ and $B$ verify the following system of ordinary differential equations:

$$\begin{aligned} &A' - kA = 0,\\ &B' + (k\mu_2 - \lambda_2\sigma_2 S(t-d)) = 0. \end{aligned}$$

Solving this system with the final conditions $A(T) = 1$ and $B(T) = 0$, we obtain

$$A(t) = \mathrm{e}^{-k(T-t)},$$
$$B(t) = \int_t^T (k\mu_2 - \lambda_2\sigma_2 S(z-d))\, dz.$$

Therefore, the futures price is given by

$$F(t,S;T) = S\mathrm{e}^{-k(T-t)} + \mu_2\left(1 - \mathrm{e}^{-k(T-t)}\right) - \lambda_2\sigma_2 \int_t^T S(z-d)\mathrm{e}^{-k(T-t)}\, dz. \tag{21}$$

In order to approximate the prices, in (20) and (21), we use the trapezoidal numerical integration formula for the integrals.

In nonparametric models, the RPDE (4) cannot be solved as before because its coefficients are not functions with explicit expressions. Then, the conditional expectation in (5) must be used.

- **Nonpar** model
  We approximate the futures prices by means of (5) and Monte Carlo method.

## 4 | EMPIRICAL APPLICATION

In this section, we estimate the risk-neutral drift and volatility of the commodity spot price, using the methodology in Section 3, and price gold futures traded in United States. All the implementation has been made using MATLAB.

As Brooks and Prokopczuk[41] establish, it is inappropriate to treat the commodities as a single-asset class because they have their own singularities. Therefore, we use different models and implement our methodology using gold futures market data. More precisely, in this paper, we employ daily gold futures data[†] traded in the COMEX.

Gold is a hard commodity, specifically a precious metal, which is relatively stable and also serves as a reserve asset for central banks. Available gold futures prices in our database cover more than 9 years, spanning October 2012 to June 2021 observations, with maturities up to 7 months.

As we do not know the gold spot price, we use the front-month futures price as a proxy, as usual in the literature; see previous studies.[42–44] We plot the gold spot price and its changes in Figure 1 and report the main descriptive statistics in Table 1.

So as to estimate all the functions of the models and show their robustness, we divide our database in three periods. We keep the first 3 months of the data to consider the delay (from October to December), retain the last year (from June 2020 to 2021) as the out-of-sample data, and the rest of the data will be the in-sample period, that is, from January 2013 to June 2020.

For comparison, we also price commodity futures using two very well-known Markovian models in the literature: a geometric Brownian motion (see Brennan and Schwartz[19] and Black and Scholes[45]) and a log-mean reversion model (see Schwartz[21]).

In the first model, we assume that the spot price follows a geometric Brownian process:

$$dS(t) = \mu_g S(t)\, dt + \sigma_g S(t)\, dW^{Q}(t),$$

where $\mu_g$ and $\sigma_g$ are constants. Then, assuming a constant market price of risk, we calculate the futures price (see Brennan and Schwartz[19]):

$$F(t,S;T) = S\mathrm{e}^{\mu_g(T-t)}. \tag{22}$$

We will refer to this futures model as **Geometric**.

[†]All data are obtained from Nasdaq Data. Link: https://data.nasdaq.com.

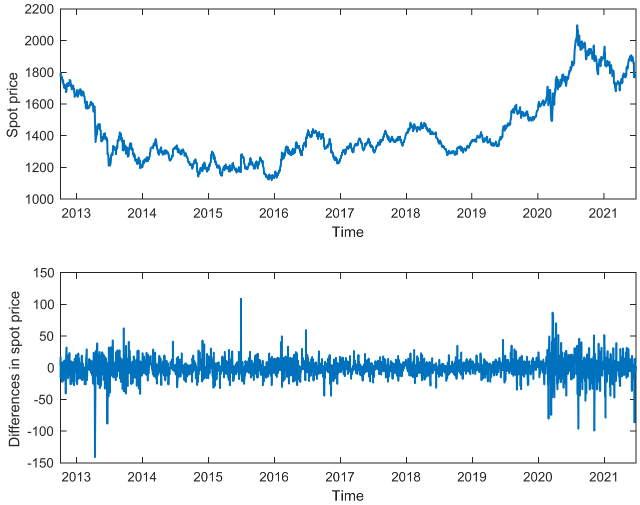


**FIGURE 1** Daily gold spot price and its differences: from October 2012 to June 2021. [Colour figure can be viewed at wileyonlinelibrary.com]

**TABLE 1** Summary of the statistics of the gold spot prices and their first differences, October 2012 to June 2021.

| | Variable | N | Mean | SD | Max | Min |
|---|---|---|---|---|---|---|
| Gold | | | | | | |
| | $S(t)$ | 2191 | 1429.0596 | 212.768 | 2097 | 1120.1 |
| | $S(t+1) - S(t)$ | 2190 | 0.0032 | 14.6021 | 108.7 | −140.4 |

**TABLE 2** Estimated parameters with the in-sample data and a time delay of 30 days.

| Parameters | |
|---|---|
| $\mu_g$ | $-2.190 \times 10^{-1}$ |
| $\sigma_g$ | $1.599 \times 10^{-1}$ |
| $k_s$ | $6.871 \times 10^{-5}$ |
| $\mu_s$ | $-311.89$ |
| $\sigma_s$ | $6.573 \times 10^{-3}$ |
| $\mu_1$ | $-1.802 \times 10^{-1}$ |
| $\sigma_1$ | $1.128 \times 10^{-4}$ |
| $\lambda_1$ | $-1.119$ |
| $k$ | $-1.418 \times 10^{-3}$ |
| $\mu_2$ | $1.849 \times 10^{5}$ |
| $\sigma_2$ | $1.616 \times 10^{-1}$ |
| $\lambda_2$ | $-1.156$ |

In the second model, the process is given by a log-mean reversion model

$$dS(t) = k_s(\mu_s - \ln S(t))S(t)\,dt + \sigma_s S(t)\,dW^Q(t),$$

where $k_s$, $\mu_s$, and $\sigma_s$ are constants and, assuming a constant market price of risk, the futures price verifies; see Schwartz[21]:

$$\ln F(t,S;T) = \mathrm{e}^{-k_s(T-t)} \ln S + \left(\mu_s - \frac{\sigma_s^2}{2k_s}\right)\left(1 - \mathrm{e}^{-k_s(T-t)}\right) + \frac{\sigma_s^2}{4k_s}\left(1 - \mathrm{e}^{-2k_s(T-t)}\right). \tag{23}$$

We will refer to this futures model as **Schwartz**.

In order to estimate all the parameters of both models, we use the expressions for the futures prices (22) and (23), respectively, the least squares method and spot and futures prices from the in-sample period. The first five rows of Table 2 show the values of the estimated parameters.

As far as the models with delay are concerned, using the in-sample and the delay period data, we estimate the risk-neutral drift and volatility of the spot price as established in Section 3. The volatility is not affected by the change of measure; then, we estimate it using (18) and spot prices. So as to estimate the risk-neutral drift, we apply (16) and use gold

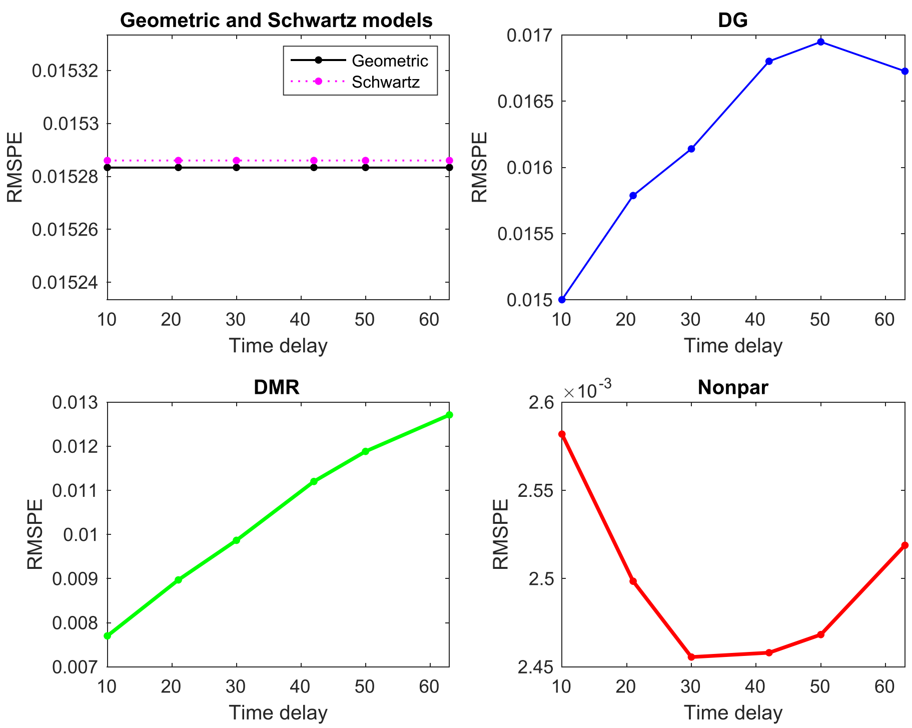


**FIGURE 2** Gold futures RMSPE obtained with different time delays and models. [Colour figure can be viewed at wileyonlinelibrary.com]

**TABLE 3** MAE, RMSE, MAPE, and RMSPE of the gold futures prices obtained with **Geometric**, **Schwartz**, **DG**, **DMR**, and **Nonpar** models, in the out-of-sample (July 2020–June 2021).

| | Geometric | Schwartz | DG | DMR | Nonpar |
|---|---|---|---|---|---|
| MAE | 25.8398 | 25.8445 | 28.7083 | 17.9814 | 3.6720 |
| RMSE | 28.9846 | 28.9898 | 30.7974 | 18.6636 | 4.7041 |
| MAPE | $1.3710 \times 10^{-2}$ | $1.3712 \times 10^{-2}$ | $1.5180 \times 10^{-2}$ | $9.5506 \times 10^{-3}$ | $1.997 \times 10^{-3}$ |
| RMSPE | $1.5283 \times 10^{-2}$ | $1.5286 \times 10^{-2}$ | $1.6140 \times 10^{-2}$ | $9.8682 \times 10^{-3}$ | $2.4556 \times 10^{-3}$ |

futures prices with all available equidistant maturities, that is, from 2 to 7 months (we don't use 1-month futures prices since they are used as a proxy of the spot price).

In order to analyze the effect of the delay, we compute the futures prices with the different models and calculate the errors in the out-of-sample period.

As measures of error, we use the mean absolute and percentage errors (MAE and MAPE, respectively):

$$MAE_\tau = \frac{1}{N}\sum_{i=1}^{N}\left|F^{\theta}_{i\tau} - F^{M}_{i\tau}\right|,$$

$$MAPE_\tau = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{F^{\theta}_{i\tau} - F^{M}_{i\tau}}{F^{M}_{i\tau}}\right|,$$

where $F^{M}_{i\tau}$ and $F^{\theta}_{i\tau}$ represent the observed and estimated futures prices, respectively, with maturity $\tau$. We also use the root mean square and percentage errors (RMSE and RMSPE, respectively):

$$RMSE_\tau = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|F^{\theta}_{i\tau} - F^{M}_{i\tau}\right|^2},$$

$$RMSPE_\tau = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|\frac{F^{\theta}_{i\tau} - F^{M}_{i\tau}}{F^{M}_{i\tau}}\right|^2}.$$

We denote as MAE, MAPE, RMSE, and RMSPE the corresponding errors with the whole maturities.

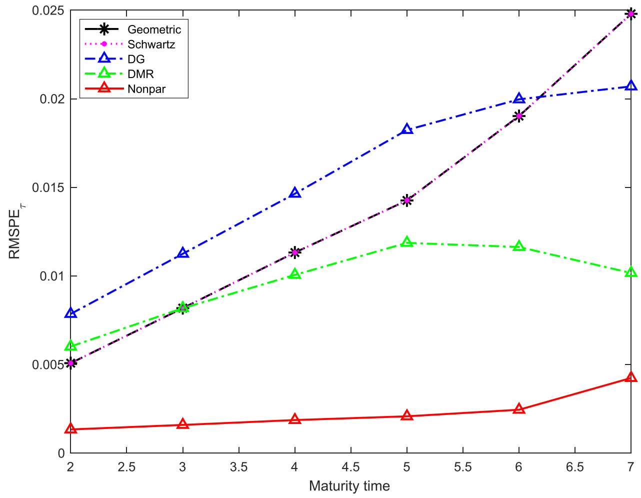


**FIGURE 3** Gold futures RMSPE$_\tau$ obtained with each model for the different maturities, with a time delay of 30 days. [Colour figure can be viewed at wileyonlinelibrary.com]

First of all, we analyze the effect of the delay by computing the futures pricing errors with several delays (up to 3 months) with the different models.

The results are shown in Figure 2. As the **Geometric** and **Schwartz** models don't have delay, their errors, represented in the first plot, are constant. In the second, third, and fourth plots, we represent the RMSPE obtained with the **DG**, **DMR**, and **Nonpar** models, respectively. We observe that the **DMR** and **Nonpar** models provide lower errors than the **DG**, **Geometric**, and **Schwartz** models. Therefore, it is interesting to remark that the delay is important, because even considering short delays the errors are, in general, lower than those without delay (**Geometric** and **Schwartz**). Moreover, the lowest error is obtained with the **Nonpar** model with a delay of 30 trading days. This fact means that the effect of the available information in the markets does not affect instantly the prices, but it suffers some delay. One of the possible reasons is that traders fail to assimilate the available information or because of certain costs, such as the cost of transacting or opportunity cost. Moreover, after a period of time its effect starts to decrease, probably because new information is continuously generated in the markets.

Finally, we also show the accuracy of the models during the out-of sample period. We consider a delay of 30 days, because as we have seen in Figure 2, it provides the lowest RMSPE. All the estimated parameters with the in-sample data are showed in Table 2.

In Table 3, we show the different errors for the whole out-of-sample period and different models. We observe that the **DG** model is not better than the models without delay. This is, possibly, because it does not include the mean reversion to the mean characteristic which is very common in commodities; for example, see Schwartz and Smith.[22] Moreover, the **Geometric** and **Schwartz** models are even better. However, this could be because these models have a closed-form solution for the price, but we just obtain an approximated solution for the models with delay. Nevertheless, the **DMR** reduces the errors of the models without delay (**Geometric** and **Schwartz**) about more than 36%.[‡] Finally, we see that the **Nonpar** model provides the most accurate prices. In fact, this model reduces the errors of the models without delay (**Geometric** and **Schwartz**) by more than 84%.

Figure 3 shows the RMSPE$_\tau$ of the different models for the different maturities. Note that the **DMR** and **Nonpar** models provide lower errors for nearly all the considered maturities, and the superiority of the **Nonpar** model is clearly showed.

In conclusion, we think that including memory in the spot price dynamics is very important for pricing futures. Hence, in this paper, we propose a new methodology to take into account both the delay in the commodity price process and the accurate estimation of its risk-neutral drift. Estimating the functions of the processes under the risk-neutral measure is essential in order to avoid misspecification[§]; for example, see Gómez-Valle et al.[39]

[‡]This percentage is computed as the difference between the errors of the compared (with and without delay) models divided by the error of the model without delay and multiplied by 100.
[§]We also priced the futures assuming that the risk-neutral drift of the spot price is equal to the drift under the physical measure, that is, considering that the market price of risk is equal to zero, and we saw that the errors increased.

Therefore, at least for pricing gold futures, market practitioners should take into account the memory in the dynamics of the spot price and nonparametric estimation techniques.

## 5 | CONCLUSIONS

In the derivatives pricing literature, Markov processes are usually considered to model the assets, that is, the memory is not taken into account. However, in finance, it is well recognized that time-delay models are more realistic than those without delay; see Kemajou[3] and Benth et al.[42] Although the solution of this kind of models is usually unknown, we consider that the memory can improve the commodity futures pricing. Moreover, assuming a constant volatility is not compatible with commodity futures prices observed in the market (see Bates[46]). Therefore, we introduce the dependence of the volatility term on the spot price at previous instants of time (giving as a result a stochastic model with delay).

For an accurate pricing of commodity futures, it is necessary to estimate all the functions of the stochastic state variable under the risk-neutral measure. However, this is not possible unless a closed-form solution for the futures price is known. Therefore, in this paper, we prove a result that allows us to estimate the functions of the the commodity spot price under the risk-neutral measure directly from futures prices in the market using any parametric or nonparametric estimation method. This new technique exhibits essential to obtain accurate futures prices because it avoids making any arbitrary assumption about the market price of risk, such as being equal to zero. We also propose a numerical technique to approximate futures prices accurately for the considered parametric delay models.

Finally, in order to analyze the effect of the delay in the commodity futures pricing, we have implemented the different proposed models and estimation technique with gold futures data traded in the COMEX. We observe that considering the delay, especially with a nonparametric model, gives as a result an improvement in the futures pricing with respect to traditional Markovian models.

### ACKNOWLEDGEMENTS

The authors are grateful to the anonymous referees for their constructive suggestions and helpful comments. This work was funded in part by the GIR Optimización Dinámica, Finanzas Matemáticas y Utilidad Recursiva of the University of Valladolid; Grant PID2020-113554GB-I00/AEI/10.13039/501100011033 of the Spanish Agencia Estatal de Investigación; Grant VA193P20 of the Junta de Castilla y León, Spain and European FEDER Funds.

### CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interests.

### ORCID

*Lourdes Gómez-Valle* https://orcid.org/0000-0002-1419-0752
*Julia Martínez-Rodríguez* https://orcid.org/0000-0001-8937-4217

### REFERENCES

1. Fama E. *Efficient capital markets: a review of theory and empirical work*. J Finance. 1970;**25**(2):383-417.
2. Titan AG. *The efficient market hypothesis: review of specialized literature and empirical research*. Procedia Econom Finance. 2015;**32**:442-449.
3. Kemajou E. *A stochastic delay model for pricing corporate liabilities*. *Dissertation*: OpenSIUC. Paper 547.
4. Bernard VL, Thomas JK. *Post-earnings-announcement drift: delayed price response or risk premium*. J Account Res. 1989;**27**:1-36.
5. Scheinkman JA, LeBaron B. *Nonlinear dynamics and stock returns*. J Bus. 1989;**62**(3):311-337.
6. Küchler U, Platen E. *Time delay and noise explaining cyclical fluctuations in prices of commodities*, Research Paper Series, vol. 195. Sydney: Quantitative Finance Research Centre, University of Technology; 2007.
7. Agrawal N, Hu Y. *Jump models with delay—option pricing logarithmic Euler-Maruyama scheme*. Mathematics. 2020;**8**:1932.
8. Arriojas M, Yaozhong H, Mohammed S, Pap G. *A delayed Black and Scholes formula*. Stoch An App. 2007;**25**:471-492.
9. Lee KA, Lee K, Park S. *A financial market of stochastic delay equation*. Bull Korean Math Soc. 2019;**56**(5):1129-1141.
10. Lee M, Kim J. *A delayed stochastic volatility correction to the constant elasticity of variance model*. Acta Math Appl Sin. 2016;**32**:611-622.
11. Mao X, Sabanis S. *Delay geometric Brownian motion in financial option valuation*. Inter J Prob Stoch Proc. 2013;**85**:295-320.
12. Swishchuk A, Xu L. *Pricing variance swaps for stochastic volatilities with delay and jumps*. Int J Stoch An. 2011;**2011**:435145.
13. Flore F, Nappo G. *A Feynman-Kac type formula for a fixed delay CIR model*. Stoch An Appl. 2019;**37**:550-573.

14. Chunxiang A, Shao Y. *Portfolio optimization problem with delay under Cox-Ingersoll-Ross model*. J Math Finance. 2013;**13**(4):527-542.
15. Benhabib J. *Interest rate policy in continuous time with discrete delays*. J Money Credit Bank. 1989;**27**:1-36.
16. Kemajou-Brown E, Mohammed S, Menoukeu-Pamen O, Tambue A. *A stochastic delay model for pricing debt and loan guarantees: theoretical results*. Int J Evol Equ. 2015;**10**:315-332.
17. Kyriakou I, Nomikos NK, Pouliasis P, Papapostolou C. *Affine-structure models the pricing of energy commodity derivatives*. Eur Financ Manag. 2016;**22**(5):853-881.
18. Pyndick RS. *Volatility and commodity price dynamics*. J Futures Mark. 2004;**24**:1029-1047.
19. Brennan M, Schwartz E. *Evaluating natural resource investments*. J Bus. 1985;**58**:135-157.
20. Gibson R, Schwartz ES. *Stochastic convenience yield and the pricing of oil contingent claims*. J Finance. 1990;**45**(3):959-976.
21. Schwartz ES. *The stochastic behavior of commodity prices: implications for valuation and hedging*. J Finance. 1997;**52**:923-973.
22. Schwartz ES, Smith JE. *Short-term variations and long-term dynamics in commodity prices*. Manag Sci. 2000;**46**(7):893-911.
23. Gómez-Valle L, Habibilashkary Z, Martínez-Rodríguez J. *A new technique to estimate the risk-neutral processes in jump-diffusion commodity futures models*. J Comput Appl Math. 2017;**309**:435-441.
24. Hilliard JE, Reis J. *Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot*. J Financ Quant Anal. 1998;**33**(1):61-86.
25. Yan X. *Valuation of commodity derivatives in a new multifactor model*. Rev Deriv Res. 1997;**5**:177-188.
26. Gómez-Valle L, Habibilashkary Z, Martínez-Rodríguez J. *A multiplicative seasonal component in commodity derivative pricing*. J Comput Appl Math. 2018;**330**:835-847.
27. Moreno M, Novales A, Platania F. *Long-term swings and seasonality in energy markets*. Eur J Oper Res. 2019;**279**:1011-1023.
28. Sorensen C. *Modeling seasonality in agricultural commodity products*. J Futures Mark. 2002;**5**:393-426.
29. Kazmerchuk Y, Swishchuk A, Wu J. *A continuous-time GARCH model for stochastic volatility with delay*. Can App Math Q. 2005;**13**:123-149.
30. Tambue A, Kemajou E, Mohammed S. *A stochastic delay model for pricing debt and equity: numerical techniques and applications*. Commun Nonlinear Sci Numer Simul. 2015;**20**:281-297.
31. Oksendal B, Sulem A. *Stochastic Differential Equations: An Introduction with Applications*. Heidelberg: Springer-Verlag; 2000.
32. Protter C. *Stochastic integration and differential equations*. Berlin, Heidelberg: Springer-Verlag; 2007.
33. Kwok Y. *Mathematical models of financial derivatives*. Berlin, Heidelberg: Springer; 2008.
34. Gabillon J. *The term structures of oil futures prices*. Oxford, United Kingdom: Oxford Institute for Energy Studies; 1991.
35. Shreve SE. *Stochastic calculus for finance II. Continuous-time models*. New York: Springer; 2008.
36. Vasiceck O. *An equilibrium characterization of the term structure*. J Financ Eco. 1997;**5**:177-188.
37. Gómez-Valle L, Martínez-Rodríguez J. *Advances in pricing commodity futures: multifactor models*. Math Comput Model. 2013;**57**:1722-1731.
38. Stanton R. *A nonparametric model of term structure dynamics and the market price of interest rate risk*. J Finance. 1997;**52**:1973-2002.
39. Gómez-Valle L, Kyriakou I, Nomikos NK. *Estimating risk-neutral freight rate dynamics: a nonparametric approach*. J Futures Mark. 2021;**41**:1824-1842.
40. Hardle W. *Applied nonparametric regression*, Econometric Society Monographs, vol. 19. New York: Cambridge University Press; 1999.
41. Brooks C, Prokopczuk M. *The dynamics of commodity prices*. Quant Finance. 2013;**13**(4):527-542.
42. Benth F, Khedher A, Vanmaele M. *Pricing of commodity derivatives on processes with memory*. Risks. 2020;**8**:1-32.
43. Geman H, Smith WO. *Theory of storage, inventory and volatility in the LME base metals*. Resour Policy. 2013;**38**:18-28.
44. Henderson B, Pearson N, Wang L. *New evidence on the financialization of commodity markets*. Rev Financ Stud. 2015;**28**(5):1285-1311.
45. Black F, Scholes M. *The pricing of options and corporate liabilities*. J Pol Econ. 1973;**81**:637-654.
46. Bates D. Testing option pricing models, statistical models in finance. In: Maddala GS, Rao GR, eds. *Handbook of statistics*, Statistical Methods in Finance, vol. 14. Amsterdam: North-Holland; 1996:567-611.

**How to cite this article:** Gómez-Valle L, Martínez-Rodríguez J. Estimating and pricing commodity futures with time-delay stochastic processes. Math Meth Appl Sci. 2023;1-12. doi:10.1002/mma.9115