Energy Economics 44 (2014) 212–221

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

What do market-calibrated stochastic processes indicate about the long-term price of crude oil?

Warren J. Hahn [a,*], James A. DiLellio [b], James S. Dyer [a]

[a] The University of Texas at Austin, 1 University Station, B6000, Austin, TX 78712, USA
[b] Pepperdine University, 18111 Von Karman Ave, Irvine, CA 92612, USA

ARTICLE INFO

Article history:
Received 28 May 2013
Received in revised form 3 April 2014
Accepted 13 April 2014
Available online 25 April 2014

JEL classification:
C52
C53
Q47

Keywords:
Oil prices
Futures markets
Stochastic processes
Kalman filter
Forecasting

ABSTRACT

Stochastic process models of commodity prices are important inputs in energy investment evaluation and planning problems. In this paper, we focus on modeling and forecasting the long-term price level, since it is the dominant factor in many such applications. To provide a foundation for our modeling approach we first evaluate the empirical characteristics of crude oil price data from 1990 to 2013 using unit root and variance ratio tests. Statistical evidence from these tests shows only weak support for the applicability of stationary mean-reverting type processes up through 2004, with non-stationary Brownian motion type processes becoming more plausible when the data from 2005 to 2013 is added. We then apply a Kalman filtering method with maximum likelihood approach to estimate the model parameters for both a single-factor Geometric Brownian motion (GBM) process as well as the two-factor Schwartz and Smith (2000) process. The latter process decomposes the spot price into unobservable factors for the long-term equilibrium level and short-term deviation, and it accommodates aspects of both a GBM process and a mean-reverting process. Both empirical and simulated data are analyzed with these models, and we quantify the increases in both the drift rate and volatility of these processes that result from developments in the crude oil markets since the middle of the last decade. We conclude by comparing and contrasting both historical accuracy and forecasts from the parameterized models, and show that when the emphasis is on the long-term expectations, a single factor GBM forecast may be sufficient.

1. Introduction

Given that crude oil is generally expected to play a significant role in meeting the world's energy needs for the foreseeable future, forecasting the price of this commodity will be important for planning continuing oil exploration and production investments. Additionally, oil price forecasting can provide insights for determining the potential impact of energy costs on the broader economy and developing appropriate policies for the eventual transition to alternative energy sources. These issues are more likely to depend on the long-term expectations for crude oil price rather than short-term fluctuations; therefore, a primary objective of this paper is to evaluate different approaches for developing long-term forecasts based on the most recent developments in crude oil prices.

There are several different approaches to developing longer-term forecasts for commodity prices, including many types of econometric models, equilibrium models, and expert survey forecasts. In this paper, we use an approach that is based upon calibrating some of the commonly-used stochastic process models with data from the commodities markets. Schwartz (1997), Schwartz and Smith (2000), Manoliu and Tompaidis (2002), and others describe how the parameters for these types of process models can be obtained with the Kalman filter and maximum likelihood estimation, and evaluate the performance of these models for capturing the dynamics of futures prices. We extend this previous work by first expanding the scope of the parameterization study, including an update through the last decade of data that includes several significant developments in the crude oil market. We also conduct back-testing of forecasts at different points in time, and validate the parameterization process through a simulation study. We then use the capability of the Kalman filter to isolate the unobservable long-term component of the crude oil price process, and use this information to develop long-term forecasts.

It is important to distinguish this work from other related empirical research, and to note what we believe are important characteristics of the approach we utilize. First, there are several studies which evaluate the historical performance of futures prices as direct predictors of future spot prices (e.g., Alquist and Kilian, 2010). We instead use futures prices to parameterize different stochastic price models, and then use those models to generate forecasts of spot prices, so that the relationship between futures and spot prices is established within the context of a risk-neutral valuation framework. In this approach, futures prices are equal to the expected future spot price under a risk-neutral stochastic process (Duffie, 1992).

* Corresponding author. Tel.: +1 512 232 6703; fax: +1 512 471 4131.
E-mail address: joe.hahn@mccombs.utexas.edu (W.J. Hahn).

http://dx.doi.org/10.1016/j.eneco.2014.04.007
0140-9883/

There are also other studies which use econometric methods to model oil market structure or construct equilibrium models to capture the relationships between oil price and fundamental factors such as demand, supply, storage and other variables (e.g., Alquist and Killian, 2010; Chevillon and Rifflart, 2009; Dees, et. al., 2007; Kaufmann et. al., 2008; Kaufmann and Ullman, 2009; Pindyck, 2001). In addition to testing these fundamental relationships, still other research has the added objective of identifying breaks or changes in the market structure over time (e.g., Fan and Xu, 2011; Kapetanios and Tzavalis, 2010; Krichene, 2002; Miller and Ratti, 2009).

Fundamental supply and demand relationships, as well as new features of the increasingly complex global oil market are present in the approach that we use as well, but they are specified by the consensus of market participants as they set futures prices through their transactions. The aggregation of these factors into futures prices makes their individual effects unobservable, and would seem to limit the usefulness of the price information set. However, the application of Kalman filtering allows us to recover those effects through optimal estimation of the stochastic process parameters, so that we can decompose prices into, for example, short and long term components, where both components are affected by supply, demand and other market factors. The Kalman filter also allows us to deal with changes in the market structure over time as well, since it is a recursive procedure for estimating the state variables at a given time, based on the information available at that time, thereby enabling continuous estimation as new information becomes available (Schwartz, 1997).

Using this approach, we seek to address two basic research objectives. First, based on the current data, we wish to investigate which of the most commonly-used stochastic process forms is most appropriate, and what are the most likely model parameters. In particular, we want to determine whether the typical assumption of stationary oil prices is still valid, in light of the market developments over the past decade. To our knowledge, this is the first work to re-parameterize the forms of stochastic processes that we consider with the Kalman filter using data from this period. Given the results for this first objective, the secondary purpose of this paper is to outline the implications for forecasting oil prices over the longer term. We believe that this work is the only application of these parameterized models for this purpose.

In the next section, we begin with a description and comparison of some of the popular forms of commodity price process models. Section 3 presents an empirical analysis of the historical crude oil price data to evaluate the potential fit with the modeling frameworks discussed in Section 2. In Section 4, we address our first research objective as we discuss in detail the stochastic process parameter specification and obtain parameter estimates for what appear to be the most appropriate processes for modeling the long-term price of crude oil using both actual futures data and synthetic data obtained through simulation. In Section 5 we address our second research objective by presenting our forecasts using the parameter estimates obtained in Section 4 and compare and contrast the results. We conclude in Section 6 with a summary of findings and implications.

2. Commodity price process models

The most basic models used for commodity prices are simple one-factor stochastic processes. Perhaps the most common of these models is a Geometric Brownian motion (GBM) process, in which commodity prices P evolve according to the stochastic differential equation

$$DX = \mu dt + \sigma dz,$$

where $X = \ln(P)$, μ is the expected rate of change or drift rate of the process over an increment of time dt, σ is the process volatility, and $dz = \varepsilon\sqrt{dt}$ is a random increment of a standard Brownian motion process with $\varepsilon \sim \text{Normal}(0,1)$. The GBM model implies that $\ln(P)$ is normally distributed with mean $\ln(P_0) + \mu t$ and variance $\sigma^2 t$, where P_0 is the price at $t = 0$.

The economic assumptions behind the GBM model are that commodity prices are expected to increase over time at a continuous rate, due to inflation and other growth factors, with the variance of prices also increasing in relation to time. A GBM model of prices also implies the assumption that markets are efficient, so that all relevant past price information is impounded in current prices and that future price movements are conditionally independent of past price movements. Under these assumptions, rational investors drive a non-stationary process with the expectation of normally-distributed returns and a lognormal distribution of prices. The GBM process is the most commonly assumed model of prices in the markets for equities and other financial assets that are traded by investors that are generally assumed to have such expectations. A GBM model of prices is simple to implement, flexible to use, and depends on a limited number of parameters. However, some research (e.g., Pindyck and Rubinfeld, 1991) has shown cases when prices were not modeled well with a GBM, and were instead shown to exhibit mean reverting behavior over time.

Mean reverting processes are an alternative type of Markov process where the sign and degree of the drift are dependent on the current level of the variable being modeled, which reverts to a long-term equilibrium level that we typically assume is the long-term mean. The economic assumptions behind this model are that, unlike the GBM model of constant expected growth of commodity prices, prices will tend to increase or decrease depending on the relationship between the price at a given time and the long-term equilibrium price level. In terms of a market hypothesis, the assumption implicit in mean reverting processes is that the price discovery process is one of so-called rational expectations. Under the rational expectations hypothesis, an inverse relationship between spot prices and the slope of the futures price curve indicates that investors expect mean reversion in spot prices, since it implies a lower expected future spot when prices increase and vice-versa. To illustrate the practical reasons behind these expectations, if we suppose that the current price rises above the long-term equilibrium value, investors might expect additional production capacity to be brought on-line and/or use of substitutes to be increased. These activities would result in downward pressure on commodity prices, forcing them back toward the long-term equilibrium level. Conversely, when prices are below the long-term equilibrium price, investors might expect capacity to be reduced and/or commodity use to be shifted away from substitutes causing the price of the commodity to rise.

The simplest form of mean reverting process is the one factor Ornstein–Uhlenbeck process, also called an arithmetic mean reverting process, which has the form

$$dX_t = \kappa(\overline{X} - X_t)dt + \sigma dz_t.$$

For commodity price modeling, X_t is the log of price, κ is the mean reversion coefficient, $\overline{X}$ is the log of the long term mean price, σ is the process volatility and dz_t is an increment of a standard Brownian motion process. The log of price is commonly used since it is generally assumed that commodity prices are lognormally distributed. This is convenient, because the price cannot be negative and it also allows future price movements to be modeled based on the stochastic behavior of returns.

The expected value and variance of the Ornstein–Uhlenbeck process are given by Eqs. (1) and (2):

$$E[X_t] = \overline{X} + (X_0 - \overline{X})e^{-\kappa T} \tag{1}$$

$$Var[X_t] = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa T}\right). \tag{2}$$

Eqs. (1) and (2) show that when $T \to \infty$, then $E[X_t] \to \overline{X}$ and $VAR[X_t] \to \frac{\sigma^2}{2\kappa}$, as opposed to a GBM where the variance approaches ∞