The Convenience Yield and the Informational Content of the Oil Futures Price

Jean-Thomas Bernard, Lynda Khalaf**, Maral Kichian***, and Sebastien McMahon****

ABSTRACT

Recent studies have shown that futures prices do not generally outperform naive no-change forecasts of spot prices, calling into question the usefulness of futures prices for forecasting purposes. However, such usefulness is predicated on the question of whether certain modeling strategies are able to yield more of the information found in futures prices. Applying a forecast-based approach, we study the extent to which alternative ways of modeling futures prices can reveal the extent of the information present in futures prices. Using weekly and monthly data, and futures of maturities of one to four months, we notably examine the out-of-sample predictability of futures prices over various forecast horizons, and in real-time, whereby parameters are updated prior to each sequential forecast.

Our results with weekly data are particularly interesting. We find that models allowing for a time-varying convenience yield often produce considerably more precise forecasts over the three forecast horizons considered. Thus, more of the informational content of futures prices is attainable when both the price level and the distance of the latter from spot price are jointly considered, rather than when only the price level is considered. We also document that forecast performances improve with longer date-to-maturity futures, suggesting that the role of the convenience yield is greater when physical oil inventories are held for longer durations. Finally, we show that forecast accuracy is highest at the one year horizon, though the time-varying convenience models have a much higher accuracy than unit-root-based models even over the three and five-year horizons.

Keywords: Oil Price, Convenience Yield, Futures markets, Oil price forecasting

http://dx.doi.org/10.5547/01956574.36.2.2

1. INTRODUCTION

Oil prices have exhibited much variability since the first oil well was drilled in Pennsylvania in 1859. Moreover, price fluctuations have been particularly marked since 1973; the level of the series has averaged more than twice its 20$/barrel (in today's dollars) mean over the 1880–1972 period, and as its volatility has increased by almost 50 per cent relative to the pre-1973 era

* Economics Department, University of Ottawa, 120 University, Ottawa, Canada, K1N 6N5. E-mail: jbernar3@uOttawa.ca. Tel: (613) 562 5800 x1374.

** Economics Department, Carleton University, Centre interuniversitaire de recherche en économie quantitative (CIREQ), and GREEN. E-mail: Lynda_Khalaf@carleton.ca.

*** Economics Department and Graduate School of Public and International Affairs, University of Ottawa. E-mail: mkichian@uottawa.ca.

*** Industrielle Alliance, Québec, Canada. Email: sebastien.mcmahon@inalco.com.

(see the graphs in BP (2013) and the Review of World Energy (2012)). These fluctuations are due to various supply and demand shocks, and are also related to a multitude of geopolitical considerations and the uncertainties that the latter entail.[1]

With the continuing high volatility in the recent decades, and the expanding use of oil in the global economy, individuals with a vested interest increasingly looked to methods for hedging against future price changes. As a result, the market for oil futures steadily developed, bringing with it an increased trading volume and more liquidity, notably for contracts of maturities of a year or less. In turn, this led policymakers and firms to become more and more reliant on futures prices for their prediction and planning needs. Given such a widespread dependence on the futures market, oil futures would be expected to contain useful information, notably on the future behaviour of spot prices. Yet, recent research has shown that naive no-change forecasts can often outperform forecasts based on futures prices (see, for instance, Alquist-Kilian (2010) and Chinn-Coibion (2013)), which suggests that futures prices do not hold information that improves on random-walk forecasting.

In this paper we contribute to the above debate by focusing on the extent of information that can be obtained from futures prices when alternative modeling strategies are used to specify the behaviour of oil futures prices. Our analysis is conducted using a forecast-based perspective. More specifically, we study the extent to which futures prices are predictable out-of-sample, based on different available models and in real-time, that is, based on updated parameter estimates for these models ahead of each individual forecast. We show that models which, at each time period, also take into account the difference between the futures price and the spot price (thereby allowing for a time-varying convenience yield), often produce considerably more precise forecasts. In other words, futures prices are found to contain a certain amount of economic and financial information that is useful for forecasting, but more importantly, that more of this information can be reached when both the price level and the distance of the latter from spot price are jointly considered. By extension, we would expect models that rely only on the futures price level (or on the first difference in the level) not to be particularly succesful at outperforming random-walk-based forecasts of spot prices. In this regard, our analysis appears to reconcile, to a certain extent, some of the seemingly contradictory positions and findings in the literature.

One branch of the literature has relied mainly on structural settings to explain the behaviour of oil prices. Well-known examples of such equilibrium models include dynamic Hotelling-type models for non-renewable resource markets, storage and inventory models for commodity markets, and financial-theory based risk and hedging models for options and futures markets.[2] At the centre of most of these models is the concept of a time-varying convenience yield which can be defined as the flow of goods and services that accrues to the owner of a commodity but not to the owner of a futures contract.[3] It is conjectured that (random) changes in the quantities of held inventories will influence spot prices more than they will affect futures prices, so that the convenience yield will vary over time. Examples of studies that propose models of time-varying convenience yields

1. The research on oil price modeling is challenging and also extensive. Examples of studies on the topic include Abosedra-Laopodis (1996), Berck-Roberts (1996), Wilson-Aggarwal-Inclan (1996), Ahrens-Sharma (1997), Schwartz (1997), Pindyck (1999), Schwartz-Smith (2000), Routledge-Spratt (2000), Pindyck (2001), Cortazar-Schwartz (2003), Cortazar-Naranjo (2006), Lee-List-Strazicich (2006), Moshiri-Foroutan (2006), Postali-Picchetti (2006), Sadorsky (2006), Regnier (2007), Trolle-Schwartz (2009), Tabak-Cajueiro (2007), CrespoCuaresma-Jumah-Karbuz (2009), Smith (2009), Hamilton (2009), Kilian (2009), Alquist-Kilian (2010), Hamilton-Wu (2012), and Chinn-Coibion (2013). A summary of some of the used approaches can also be found in Wirl (2008).

2. For a recent exposition, see Hamilton (2009) and Smith (2009).

3. For a cogent exposition in the context of oil price, see Pindyck (2001).

include, for example, Schwartz (1997), Schwartz-Smith (2000) and Pindyck (1999). In addition, Wu-McCallum (2005) and, to a lesser extent Alquist-Kilian (2010), report that the Hotelling model does relatively well in forecasting the future price of oil.

Another branch of the literature has used time-series methods to specify oil price dynamics, often relying on statistical testing to determine the model features. In general, futures prices are assumed to follow a random walk, with or without a time-varying variance. To examine whether futures prices are predictors of spot oil prices or whether they perform better than no-change forecasts, statistical tests are applied to the (first-differenced) series within the context of simple linear regression models. Recent findings from this research show that futures prices generally do not outperform no-change forecasts, though results also appear to somewhat depend on the maturity considered for the futures, on the sample period, as well as on the postulated linear relationship (see, for example, Chinn-Coibion (2013) and Alquist-Kilian (2010), as well as the references cited therein). As will be discussed below, implicit in these testing strategies are underlying models where the convenience yield is assumed to be constant.

It thus appears that the answer to the question of whether futures prices contain useful information on the behaviour of future oil prices is predicated first on the assumed underlying model. In this regard, the empirical literature on oil price modeling reveals that while almost all of the studies acknowledge volatility clusters, heavy tails, and structural breaks, statistical support has been claimed for otherwise fundamentally different types of models. The reasons for such disparity has of course to do with underlying theoretical setups, but also with the various ways data challenges are handled econometrically. The latter include discerning deterministic trends from stochastic ones, disentangling structural change in the fundamentals from inherent fluctuations, using different ways to allow for unexpected discontinuities in price levels or price changes, and accounting differently for non-constancy in the variance of prices and of price changes. Commonly used econometric tools to pin down sample data idiosyncracies include a plethora of unit root and break tests, a battery of generalized autoregressive conditional heteroskedasticity [GARCH] based inference methods, Kalman-filter based time-varying-parameter or neural-network based estimations, or jump-diffusion based procedures. No consensus emerges on what constitutes the best model.

Importantly, conclusions also depend on whether the convenience yield in the underlying model is assumed to be constant or time-varying. In constant convenience yield models, oil prices are considered to be completely unpredictable, so that a unit root is found to be the best way to represent the behaviour of the mean of the series (the high persistence of oil price has been reported, among others, by Diebold-Kilian (2000)). In studies proposing models with time-varying convenience yield (such as Schwartz (1997), Schwartz-Smith (2000), Pindyck (2001) and Pindyck (1999)) various theoretical and practical reasons are given to suspect that simple unit-root or Geometric Brownian Motion specifications are not appropriate to model natural resources or commodity prices. Instead, these models suggest that prices revert to a mean that is a long-run equilibrium, but that nevertheless this mean may itself change over time due to resource depletion, technological change or product innovations. Reversion to the mean occurs because when prices are higher (or lower) than some equilibrium level, high-cost producers will enter (or exit) the market, which pushes prices downward (or upward). They also conclude for mean reversion by analyzing the relationship between futures prices at different maturities and the spot price, in other words, the time-varying convenience yield. Thus, the convenience yield reflects the information contained in the oil futures prices regarding future values of the spot price and can be exploited.[4]

4. For details please refer to Pindyck (2001).

In light of the above discussion on the lack of concensus on the best-fitting empirical model, in this paper we follow an agnostic approach to assess whether oil futures contain useful information. Thus, rather than applying various types of tests that, depending on the data sample considered and the idiosyncraticities therein, could have little or no power for some models, or worse, could be oversized and reject spuriously, we prefer to rely on a forecast-based strategy where the exact same criterion is applied to all the models. Importantly, our analysis is also real-time-based; before a new forecast is made, model parameters are first re-estimated and thus updated. We also argue that, in these types of analyses, it is vitally important to consider data of different frequencies, as some models can make better use of the information present in one or the other frequency data.

Our results show that, with monthly data, forecast performances derived from the different models are largely comparable. However, with weekly frequency data, forecasts based on models that assume time-varying convenience yield are much more accurate compared to forecasts from models that do not make this assumption. This (often substantial) gain in accuracy comes from considering the joint behaviour of spot and futures prices, which shows that deviations at every time period of futures prices from spot prices (reflecting the time-varying convenience yield) are informationally extremely relevant.[5]

For the same model case, we also document that forecast performances increase with longer date-to-maturity futures. This result suggests that the role of the convenience yield becomes more important with increasing maturity durations, and, by the same token, that longer-maturity oil futures prices have a higher information-to-noise ratio. Finally, we show that forecast accuracy improves as the forecasting period is shortened to one year.

In the next section we explain the particular classes of models that we work with and discuss the rationale for focusing on these. Section 3 discusses the forecasting applications and the obtained results. Section 4 offers some conclusions.

2. MODEL CONSIDERATIONS

2.1 Models Motivated by Equilibrium Arguments

To account for a possibly time-varying convenience yield, we focus on a class of multivariate models that has been fairly popular in the finance literature, notably since the mid-nineties. Such models take into account the equilibrium relationship between futures prices at different maturities and the spot price. When applied to commodity markets, demand and supply pressures, as well as non-constant convenience yields, play a role in these models. The models suggest that prices should revert to long-run equilibrium prices that themselves may change over time due to random technological and resource innovations. Mean reversion and the positive correlation between spot price and convenience yield changes are thus consistent with the theory of storage with random shocks: when inventories decrease (or increase), the spot price will increase (or decrease) and the convenience yield will also increase (or decrease), however futures prices will not increase (or decrease) as much as the spot prices. Studies such as Schwartz (1997), Schwartz-Smith (2000) or Pindyck (1999), Pindyck (2001) refute the hypothesis of a constant

5. The relevance of the information in the term structure of futures prices was also documented in financial studies such as Fuertes-Miffre-Rallis (2010) who consider the design of profitable trading strategies in commodity futures markets and focus on the combined roles of momentum and term structure signals in these markets.

convenience yield and their results suggest mean reversion to a long-run equilibrium that itself can change randomly over time.[6]

Mean reversion thus stems from the crucial role that inventories play in the case of commodities.[7] Numerous intuitive arguments may be envisaged to explain holding inventories. These include hedging in view of demand shifts against adjustment, marketing, scheduling or delivery costs, with obvious implications on price volatility. Indeed, in response to shifts in demand for commodities, producers make joint decisions on production and inventory levels, accounting for: (i) a spot price for the commodity, and (ii) a storage price determined from the so-called "marginal convenience yield." The latter is the flow of benefits that accrue to inventory holders from holding a marginal unit of inventory. In equilibrium it is equal to the cost of storage plus the interest loss on the purchase of the commodity (spot price) and minus the difference between the spot price and futures price. Two markets thus interact in the case of commodities, so equilibria in both markets are relevant, and interactions between these markets may be captured by analyzing the relationship between spot and futures markets. The following scenario may illustrate this dual market characteristic: in response to an unexpected temporary demand shock, spot prices are expected to increase and inventory holding to decrease while the futures prices stay constant. The price of inventory holding *i.e.* the convenience yield goes up; however it falls back as the situation returns to normal. Changes in spot price volatility may also trigger similar effects. Inventory adjustments are accompanied with adjustments in both spot and convenience yields.

Tractable time series models are available to capture such effects, and are typically characterized by mean reversion, but a time-varying convenience yield implies that the mean to which price reverts is, itself, time-varying. Such models are also related to the affine factor models recently proposed by Hamilton-Wu (2012) who, focusing on the arbitrageurs who take the other side of the contract with respect to commercial and financial hedgers, allow for time-varying risk premia in oil futures.

This paper specifically considers the mean-reverting class of models from Schwartz-Smith (2000) and Schwartz (1997) and conducts a forecast-based analysis.[8] The considered mean-reverting structural form not only presumes a stochastic convenience yield that comes about from the joint behaviour of spot and different future prices, but also allows one to disentangle the persistent or long-run element from the transitory or short-run component of oil price. That is, the two-factor model formally allows for a time-varying long-run mean and integrates both short- and long-run movements by construction. More precisely, the long-run equilibrium component follows a Brownian motion, whereas the short-run deviations follow an Ornstein-Uhlenbeck process that reverts towards zero. The model can be written in continuous time for the log level of the spot price as:

$$
\begin{aligned}
\ln(Y_t) &= \chi_t + \xi_t, \qquad (1)\\
d\chi_t &= -\kappa\chi_t dt + \sigma_\chi dz_\chi,\\
d\xi_t &= \mu_\xi dt + \sigma_\xi dz_\xi,\\
dz_\chi dz_\xi &= \rho_{\chi\xi} dt,
\end{aligned}
$$

6. Note that Pindyck (1999) models long-run prices, whereas Schwartz (1997) or Schwartz-Smith (2000) propose models that incorporate both short- and long-run considerations.

7. The summary of the theoretical underpinning of such models provided here borrows heavily from Pindyck (2001).

8. CrespoCuaresma-Jumah-Karbuz (2009) use also forecasting performance as the model selection criterion in their analysis of quarterly oil prices but they assume that the oil price trend has a unit root.

where ξ_t is the log equilibrium price of oil at time t, χ_t is the deviation of the log price at time t with respect to the equilibrium price, and dz_χ and dz_ξ are correlated increments of Brownian motions. The mean-reversion coefficient κ represents the rate of speed at which the price reverts to its equilibrium *i.e.*, the rate at which short-run deviations disappear, μ_ξ is the growth rate of the log of equilibrium price, and σ_χ and σ_ξ are the short-run and equilibrium volatilities of the process, respectively. For estimation purposes, (1) can be discretized as follows:

$$\ln(Y_t) = \chi_t + \xi_t, \tag{2}$$

$$\chi_t = e^{-\kappa}\chi_{t-1} + \epsilon_t^\chi, \tag{3}$$

$$\xi_t = \mu_\xi + \xi_{t-1} + \epsilon_t^\xi. \tag{4}$$

Such a decomposition allows one to disentangle the persistent from the transitory components of oil price.

To explain how a stochastic convenience yield intervenes in this model, Schwartz-Smith (2000) relate (1) to the following model from Schwartz (1997):

$$dX_t = \left(\mu - \delta_t - \frac{1}{2}\sigma_1^2\right)dt + \sigma_1 dz_1, \tag{5}$$

$$d\delta_t = \kappa(\alpha - \delta_t)dt + \sigma_2 dz_2,$$

$$dz_1 dz_2 = \rho dt,$$

where $X_t = \ln(Y_t)$ (in our notation, X_t gives the log of the current spot price), dz_1 and dz_2 are correlated increments of Brownian motions, and δ_t is the convenience yield, which intervenes as a reduction in the drift term of (5). Formally, Schwartz-Smith (2000) show that processes (1) and (5) are equivalent, in the sense that factors of (1) can be written as a linear combination of the factors in (5). In particular, $\chi_t = \frac{1}{\kappa}(\delta_t - \alpha)$; note that κ gives the short-term mean-reversion rate in both versions of the model, which justifies the overlap in notation.

Given the above specification, closed form solutions for the prices of futures can be obtained using standard valuation methods. For this purpose, Schwartz-Smith (2000) derive the corresponding risk-neutral process as:

$$d\chi_t = (-\kappa\chi_t - \lambda_\chi)dt + \sigma_\chi dz_\chi^*,$$

$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi dz_\xi^*,$$

$$dz_\chi^* dz_\xi^* = \rho_{\chi\xi} dt$$

where λ_χ and λ_ξ are constant reductions in the drifts of each process, and again, dz_χ^* and dz_ξ^* are correlated increments of Brownian motions. Risk-adjustment now implies mean-reversion to $-\kappa/\lambda_\chi$ (rather than zero) for the short-run process. In the latter valuation framework, assuming that future prices are given by the expected future spot price leads to the following specification for future prices:

$$\ln(F_{n,t}) = e^{-kn}\chi_t + \xi_t + A(n), \tag{6}$$

where $F_{n,t}$ represents the market price, at time t, for a futures contract with time n until maturity and

$$\xi_t = \mu_\xi + \xi_{t-1} + \epsilon_t^\xi,$$

$$\chi_t = e^{-\kappa}\chi_{t-1} + \epsilon_t^\chi,$$

$$A(n) = (\mu_\xi - \lambda_\xi)n - (1-e^{-\kappa n})\frac{\lambda_\chi}{\kappa} + \frac{1}{2}\left((1-e^{-2\kappa n})\frac{\sigma_\chi^2}{2\kappa} + \sigma_\xi^2 n + 2(1-e^{-\kappa n})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{\kappa}\right).$$

The system can then be written in a state-space form, and the latter is amenable to estimation via the Kalman filter.[9]

2.2 Time-Series-Based Models

We also consider various popular time-series models which can allow for non-constancies in the level and in the volatility of the futures price series. We first consider the class of GARCH specifications which have been widely used to capture one of the major features documented in the literature and that we feel is relevant for our purpose, namely that the volatility of price returns changes over time. We also consider GARCH processes with random jumps, given the importance of accounting for structural discontinuities.[10] Such jumps can be seen as an integral part of the empirical price process leading to relatively rare adjustments that can be distinguished from frequent and relatively "small" ordinary price fluctuations. These adjustments can result from an accident that shuts down the production of a large oil field for some time, or the unexpected decision of a producer to boost or cut down its production, or more generally from the arrival of unexpected information.[11] On breaks versus ARCH in oil prices, the reader may refer to Wilson-Aggarwal-Inclan (1996).[12] In this paper, we account for non-constancies in the mean and conditional variance via jump-GARCH processes as an alternative to dating and fitting breaks in addressing the question of mean reversion.

The above models impose a unit root in the futures price series, but we do not carry out unit root testing to confirm or refute this. We are partly motivated in this by Pindyck (1999) who expresses the minimum sample size for which the Dickey-Fuller test is significant when data are generated by a stationary autoregressive process in terms of the autocorrelation coefficient. Given the persistence characterizing oil prices, the author concludes that a very long series indeed is

9. For a detailed treatment of this method, see Kim-Nelson (1999).

10. The relevant literature is vast; see for instance Merton (1976), Ball-Torous (1985), Jorion (1988), Bates (2000), Bakshi-Cao-Chen (2000), Das (2002), and Chernov-Gallant-Ghysels-Tauchen (2003).

11. For a technical discussion on the relationship between jumps, breaks and GARCH processes, see Drost-Nijman-Werker (1998) and Ait-Sahalia (2004).

12. These authors study whether ARCH effects are wrongly included because the oil series present breaks-in-volatility. They consider daily 1-month futures from January 1984 to December 1992. Using a pre-test approach to identify multiple breaks, they find significant but less persistent ARCH effects in the presence of breaks.

required to perform reliable tests, something that is virtually unavailable to researchers. Moreover, his reasoning is made ignoring all the added complications that would arise from the arrival of structural shifts, which are much more likely to occur as the data sample lenght increases. Recent work applying random-walk tests that correct for breaks provides evidence against the unit-root hypothesis, yet debate is still on-going [see *e.g.* Lee-List-Strazicich (2006), Postali-Picchetti (2006), and Hamilton (2009)].

While we do not aim to take a stand on the worth of such tests for the problem at hand, we do believe that available evidence on the importance of parameter non-constancies, whether for refuting or for justifying the unit-root hypothesis, should be taken seriously. Formally, one of the main conclusions to be drawn from this literature, at least as we interpret it, is that structural discontinuities should be accounted for in examining stochastic models for oil prices whether one adopts a unit-root or a mean-reverting model.

We classify the above models into two categories, together covering a large proportion of recent popular models applied to commodity prices. They are: (i) models without discontinuities in the mean, that is random-walk models with GARCH effects, and with normal or Student-*t* innovations, and (ii) models with discontinuities in the mean, or more specifically with Poisson jumps and GARCH effects, and with normal or Student-*t* innovations.

In what follows, specifications (i) and (ii) will be referred to as the random-walk-based class of models. For further reference, let

$$y_t = \ln(Y_t) - \ln(Y_{t-1}) \tag{7}$$

where Y_t is the nominal price level at time t, $t = 1, \ldots T$.

GARCH models are among the most popular volatility models in practice, because they are capable of describing the well known volatility clustering feature as well as observed heavy tails, and are computationally and analytically tractable. The models that we consider in this class include the GARCH-in-mean model that is motivated by theory on storage,[13] the exponential GARCH model that allows for asymmetric effects, and a standard GARCH model with Student-*t* shocks that account for conditional heavy-tailed fundamentals.

Formally, we first consider the GARCH-in-mean model denoted GARCH-M(1,1):

$$y_t = v_t + \beta h_t, \tag{8}$$

$$v_t = \mu + \sqrt{h_t} z_t, \tag{9}$$

$$h_t = \alpha_0 + \alpha_1 (y_{t-1} - \mu)^2 + \phi h_{t-1} \tag{10}$$

where the fundamental shocks z_t, $t = 1, \ldots T$ are independently and identically distributed (i.i.d.) as standard normal or Gaussian variable

$$z_t \overset{i.i.d.}{\sim} N(0,1). \tag{11}$$

This model directly captures the relationship between level of price changes and volatility, via the ϕ coefficient in (10) and it is designed to allow for volatility clustering.

13. This was studied in particular by Beck (2001) in the context of commodity prices.

By setting $\beta = 0$, the model defined by (8)–(11) nests the standard GARCH(1,1) specification which we also consider. In this case, we also adopt an alternative specification for the fundamental shocks z_t, $t = 1, \ldots T$, namely the i.i.d. Student-t distribution with τ degrees of freedom denoted as $t(\tau)$

$$z_t \overset{i.i.d.}{\sim} Student - t(\tau) \tag{12}$$

where τ is unknown. It is worth recalling that even under (11), the implied unconditional distribution of v_t is non-normal, and in particular, the unconditional kurtosis exceeds 3 (that is the Gaussian value). Hence even with normal shocks, GARCH processes would capture fat-tails. However, in applications of such models to financial high frequency data, it was observed that the unconditional kurtosis compatible with Gaussian-based GARCH models still understates the observed kurtosis in the data. Processes of the (12) form have thus been proposed to possibly address this problem. For completion, we also consider the unit root (with drift) case, which corresponds to the above models setting the GARCH parameters to zero.

We next analyze the asymmetric specification denoted EGARCH(1,1) consisting of (8)–(9) with $\beta = 0$ and with

$$\ln(h_t) = \alpha_0 + \phi \ln(h_{t-1}) + \gamma \left(\frac{y_{t-1} - \mu}{\sqrt{h_{t-1}}} \right) + \eta \left[\frac{|y_{t-1} - \mu|}{\sqrt{h_{t-1}}} - \frac{\sqrt{2}}{\sqrt{\pi}} \right]. \tag{13}$$

With this configuration, the sign of shocks is relevant. This property is particularly interesting for oil price, since negative shocks or news may conceivably affect volatility quite differently than positive ones.

Basically, the above models capture time-varying volatility as a function of the magnitude and/or the sign of lagged fundamental shocks. Except for the GARCH-M(1,1) model, these specifications restrict focus to non-constancies in the volatility, and not in the mean. There are alternative GARCH frameworks to model non-constancies in both mean and variance, for example by allowing for random jumps. From this class of models, we consider the following specification with Poisson jumps:

$$y_t = v_t + \sum_{i=1}^{n_t} \ln P_{it}, \tag{14}$$

where v_t is as defined in (9)–(10), n_t is the number of jumps that occur between $t-1$ and t, and P_{it} $(i = 1, \ldots, n_t)$ is the size of the ith jump over this time interval. Jumps follow a Poisson process with arrival rate λ, in other words, a jump occurs on average, every $1/\lambda$ periods, and the jump sizes P_{it} are i.i.d. according to a lognormal distribution with mean θ and variance δ^2. This definition implies that n_t is an integer random variable and (14) also nests (if $n_t = 0$, $t = 1, \ldots T$, or if $\lambda = 0$) the Gaussian and/or Student-t GARCH(1,1) specification. We also consider a conditionally normal ARCH(1) with jumps, given its preponderance in the literature. This specification corresponds to (14) where v_t is as defined in (9)–(10) and where the ϕ parameter in (10) is set to zero.

To conclude, two issues are also worth noting. First, the GARCH models that we consider are formulated on returns and explicitly assume that price is non-stationary. In contrast, Schwartz and Smith's model treats price as the sum of two processes, one of which is a Brownian motion and non-stationary by construction. This model thus differs from standard random walks mainly via the Ornstein-Uhlenbeck process for the short term deviation. Secondly, even though volatility clustering is not explicitly modelled in the Schwartz and Smith specification, the time-varying coefficients of the models can nevertheless adapt, fitting fat tails sufficiently adequately.

3. EMPIRICAL ANALYSIS

Our empirical analysis focuses on forecasts of oil futures prices. Using weekly and monthly frequencies, and for three forecast horizons, we examine the forecasts from the various models based on the mean square prediction error (MSPE) criterion.[14] We notably use one-step-ahead out-of-sample forecasts, where parameter estimates are updated at every step of the procedure. Forecasting in real time has various practical advantages given our focus on time-varying parameter models; in particular, drift parameter estimates in the various models are allowed to adjust with additional observations, which allows the time-series models to produce better forecasts. In addition, in models with jumps, where analytical formulae are not readily available for obtaining conditional expected forecast errors, we devise a simple simulation-based procedure to approximate these errors.[15]

More specifically, we use crude oil price futures at daily frequency obtained from the *U.S. Department of Energy, Energy Information Administration*, and for 1, 2, 3 and 4 month futures. Our data extends from January 2, 1986 to February 8, 2012. From these, we construct weekly and monthly prices, the former using Wednesday values and the latter using the price on the Wednesday that is closest to the 15th day of that month. For the few cases where the Wednesday value is not available, the Tuesday value closest to the 15th day of that month is used. The three forecast horizons that we consider are for one, three, and five year durations.

We proceed as follows. Given a sample of size $T+K$, we first set apart K observations at the end of the sample, which correspond to the forecast horizon considered. The model is then estimated on the remaining sample (i.e., until T); the dependent variable's value is forecast for period $T+1$ and denoted $\ln(\hat{Y}_{T+1|T})$. The $T+1$ forecast error resulting from the comparison of $\ln(\hat{Y}_{T+1|T})$ and $\ln(Y_{T+1})$ is computed. Next, the $T+1$ observed value of the dependent variable is added to our sample, and the model is re-estimated. The $T+2$ observation is then forecast and denoted $\ln(\hat{Y}_{T+2|T+1})$, the $T+2$ forecast error is computed, and so on, until all K observations are covered. The MSPE criterion is then defined as:

$$MSPE = \frac{1}{K}\sum_{k=1}^{K}[\ln(\hat{Y}_{T+k|T+k-1}) - \ln(Y_{T+k})]^2.$$

For the models that include jump features, a closed-form analytical solution for the forecast error is unavailable. We propose, as in Bernard-Khalaf-Kichian-McMahon (2008) and Khalaf-Saphores-Bilodeau (2003), a simulation-based approximation described as follows.

For a given model with jumps, say (14), we first estimate the model parameters over the sample of size T; denote the latter estimates $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\mu}$, $\hat{\lambda}$, $\hat{\theta}$, $\hat{\delta}^2$ and $\hat{\tau}$ when relevant. Then, drawing from a normal or $t(\hat{\tau})$-distribution for the residuals, a Poisson distribution with estimated mean $\hat{\lambda}$

14. We also calculated the mean absolute forecast error of the different models, but as they yielded almost identical model ranking, the results are not reported here.

15. See Khalaf-Saphores-Bilodeau (2003) and Bernard-Khalaf-Kichian-McMahon (2008).

for the arrivals of the jumps, and a normal distribution with mean $\hat{\theta}$ and variance $\hat{\delta}^2$ for the amplitude of each jump, we generate 1,000 simulated values of the dependent variable $\tilde{Y}_{T+1}$. The forecast value of Y_{T+1} is then taken to be the average value of these 1,000 $\tilde{Y}_{T+1}$, and the $T+1$ forecast error is computed. At this point, the observed value of the dependent variable, Y_{T+1}, is added to the sample, the model is re-estimated, and the entire simulation process is repeated. Thus, $\tilde{Y}_{T+2}$ is obtained, as well as the forecast error for $T+2$. The above steps are repeated until $T+K$ forecast errors are obtained, which are then used to construct the MSPE.

The parameters of all the models are estimated by numerical maximization of the likelihood function. In the case of the mean-reverting model, and for each futures maturity, the state-space form associated with model (6) is estimated *jointly* using the Kalman filter. Model parameters are thus obtained based not only on the individual behaviour of spot and futures prices, but also on their simultaneous joint evolution.[16]

For each frequency, out-of-sample one-step-ahead dynamic forecasts are conducted for the three forecast horizons. As mentioned before, forecasts in real time are particularly useful here, in order to give the GARCH models a fair chance relative to the time-varying mean case. Indeed, it is well known that, except for the GARCH-in-mean and the jump-GARCH case, typical GARCH forecasts will not adapt unless adaptive estimation and forecasting is considered, since the associated optimal forecast is equal to the conditional mean. We thus view the rolling estimation burden as a worthy exercise from our perspective. The resulting MSPE are reported in Tables 1 and 3, while relative forecasting performances are found in Tables 2 and 4, showing the ratio of the MSPE of a given model to that of the corresponding GARCH(1,1) model. For illustrative purposes, Table 5 also provides parameter estimates, averaged over 1-year and 5-year horizons, for selected models.

Taken collectively, the tables reveal two general outcomes. First, the forecast performances from the various models are fairly similar when monthly frequency data is considered, but they are substantially different when the data used is at the weekly frequency. Second, the Schwartz and Smith model with time-varying convenience yield outperforms the other models in many of the examined cases, notably with weekly frequency data.

Looking more closely at the results from the weekly frequency data in Table 1, we can see that the Schwartz and Smith model always yields the smallest mean square prediction errors, and often by a substantial margin. Amongst the remaining models, the ARCH-with-jumps specification is the only model that is able to match this performance, but only over the one-year forecast horizon. In contrast, the Schwartz and Smith model performs consistently and overwhelmingly better than all of the other models over the 3- and 5-year horizons. For example, over our middle forecasting horizon, and compared to the second-best-performing specification, it does 83 per cent better when the one-month futures is used, 86 per cent better with two-month futures, 87 per cent better with three-month futures, and 88 per cent better with four-month futures case. Similarly, over our longest forecast horizon, and compared to the second-best-performing specification, it does 71 per cent better with the one-month futures case, 76 per cent better with two-month futures, 74 per cent better with three-month futures case, and 77 per cent better with four-month futures case. In addition, both the Schwartz and Smith and ARCH-with-jumps models perform 95, 96, 96 and 84 per cent better for the one-, two-, three- and four-month futures cases, respectively, compared to the second-best-performing specification and over our shortest forecasting horizon.

16. For parsimony purposes and following Schwartz-Smith (2000), a diagonal matrix is assumed for the measurement errors in prices.

The relatively exceptional performance of the mean-reverting model can also be seen from the results in Table 2 which reports the mean square prediction error of a given model relative to the MSPE of the GARCH(1,1). From the same table, and looking at the autoregressive conditionally heteroskedastic category of models, we find that the EGARCH, the GARCH-with-jumps specifications, and, even to some extent, the GARCH-in-mean models have forecast performances comparable to that of the GARCH(1,1), while the ARCH-with-jumps exhibits the best accuracy in the short-term and the GARCH-*t* specification consistently performs the worst.

Two additional conclusions can be drawn from Tables 1 and 2 regarding the performance of the Schwartz and Smith model. First, forecast accuracy increases when, in conjunction with the spot rate, a longer date-to-maturity futures is considered. Thus, using four-month futures instead one-month futures improves forecast performance by 21 per cent over the 1-year horizon, 27 per cent for the 3-year horizon, and by 22 per cent for the 5-year horizon. Second, regardless of the futures maturity used, forecast accuracy decreases as the forecast horizon increases. For example, with one-month futures, and compared to the 5-year horizon, the three-year MSPE is 15 per cent smaller and the one-year MSPE is 33 per cent more accurate. Similarly, with the four-month futures, the three-year MSPE is 20 per cent better than the five-year, MSPE, while the one-year MSPE is 32 per cent better.

Turning our attention to the results based on using monthly data (Tables 3 and 4), we see that the relative forecast superiority hitherto exhibited by the Schwartz and Smith model is no longer a given. We find that it outperforms the other models in only three out of twelve cases, and only over the five-year forecast horizon. Moreover, MSPEs from all of the models are much more similar in these tables than with the results using weekly frequency, so that forecast performances are likely not to be significantly different from one another. In the best case, the performance of the best model over the worst is about 13 per cent (for the three-year horizon and using the two-month futures series), while it is often within 5 to 6 per cent for the remaining cases. In addition, no one model consistently yields the smallest MSPE, though the GARCH(1,1) has the smallest MSPE in five cases.

Our results appear to be partly in line with the conclusions in Hamilton-Wu (2012) and Chinn-Coibion (2013) in stating that time-variation in the relationship between future and spot prices is important and needs to be adequately accounted for. The former study documents that the the risk premium to holding the other side of hedging contracts has exhibited changes over time, notably becoming more volatile since 2005, while the latter study points to changes in the predictive content of futures over time, and notably decreases in predictive content since 2000, based primarily on unbiasedness and linear prediction tests. Interestingly, both these studies are based on the use of monthly data, which may shed some light on why the forecasting performance of our mean-reverting model, when we also use monthly data, is only sometimes superior relative to the forecasts from the various GARCH alternatives. This requires an extension of the present work to investigate comparative forecast performance over various sub-periods, and is left for future work.

In the meantime, the contrast between the results obtained from the different data frequencies highlights an important informational comparative advantage that, with the higher frequency data, the mean-reverting model appears to have over the other considered models. Given that the Schwartz and Smith model describes the evolution of the time-varying convenience yield in the oil market, the explanation appears to lie in the way oil inventories are actually adjusted by commercial counterparties. If these adjustments are made (and published) weekly, then changes in the marginal convenience yield are much more likely to be immediately and accurately reflected in the weekly divergences between the spot and futures prices. This information is then efficiently revealed by the mean-reverting model with the weekly frequency data and is therefore found to be useful for forecasting. In comparison, monthly data likely dilutes this higher frequency information, and diminishes the relative efficacy of the mean-reverting model.

4. CONCLUSION

The increased dependence of investors in recent decades on oil futures markets has brought firms and policymakers to rely overwhelmingly on the price of oil futures for their forecasting needs. However, recent studies have shown that futures prices do not generally outperform naive no-change forecasts of spot prices, calling into question the usefulness of futures prices for forecasting purposes. At the same time, it could be that specific modeling strategies are able to yield more of the information that might be available in futures prices. In this paper, we examined this issue from a forecast-based perspective. Using maturities of one to four months and three forecast horizons, and applying both weekly and monthly data, we notably studied the extent to which futures prices are predictable out-of-sample, based on different available models and in real-time.

We showed that forecast performances are generally comparable when monthly data is used, but that, with weekly data, models that allow for a time-varying convenience yield often produce considerably more precise forecasts. In other words, compared with models that rely only on the futures price level or on its first difference, more of the informational content of futures prices is attainable when both the price level and the distance of the latter from spot price are jointly considered. We also documented, again based on weekly data, that forecast performances improve with longer date-to-maturity futures, suggesting a bigger role for the convenience yield in those cases. Finally, we showed that forecast accuracy is highest at the one year horizon, though the time-varying convenience models were found to have a much higher accuracy than unit-root-based models even over the three and five-year horizons.

Our results appear to reconcile, to a certain extent, some of the seemingly contradictory positions and findings in the literature, particularly that when futures prices are found not to consistently outperfom no-change forecasts, it may not be so much because there is no useful forecasting information in futures prices, but because not all of the information available in these prices was exploited by the underlying model.

ACKNOWLEDGMENTS

The authors would like to thank the editor, the associate editor and the referees for helpful comments and suggestions. This work was supported in part by the Institut de Finance Mathématique de Montréal, the Canadian Network of Centres of Excellence [program on Mathematics of Information Technology and Complex Systems], the Social Sciences and Humanities Research Council of Canada, the Fonds FQRSC (Government of Quebec), the Chair on the Economics of Electrical Energy, and the Canada Research Chair in Environment. All remaining errors are our own.

REFERENCES

Abosedra, S. S. and N. T. Laopodis (1996), "Stochastic behavior of crude oil prices: A GARCH investigation," *Journal of Energy and Development* 21(2): 283–91.

Ahrens, W. A. and V. R. Sharma (1997), "Trends in natural resource commodity prices: Deterministic or stochastic?," *Journal of Environmental Economics and Management* 33: 59–74. http://dx.doi.org/10.1006/jeem.1996.0980.

Aït-Sahalia, Y. (2004), "Disentangling diffusion from jumps," *Journal of Financial Economics* 74: 487–528. http://dx.doi.org/10.1016/j.jfineco.2003.09.005.

Alquist, R. and L. Kilian (2010), "What do we learn from the price of crude oil futures?," *Journal of Applied Econometrics* 25: 539–573. http://dx.doi.org/10.1002/jae.1159.

Bakshi, G., C. Cao, and Z. Chen (2000), "Pricing and hedging long-term options models," *Journal of Econometrics* 94: 277–318. http://dx.doi.org/10.1016/S0304-4076(99)00023-8.

Ball, C. and W. Torous (1985), "On jumps in common stock prices and their impact on call option pricing," *Journal of Finance* 40: 155–73. http://dx.doi.org/10.1111/j.1540-6261.1985.tb04942.x.

Bates, D. S. (2000), "Post-87 crash fears in S&P 500 futures options," *Journal of Econometrics* 94: 181–238. http://dx.doi.org/10.1016/S0304-4076(99)00021-4.

Beck, S. (2001), "Autoregressive conditional heteroscedasticity in commodity spot prices," *Journal of Applied Econometrics* 16: 115–32. http://dx.doi.org/10.1002/jae.591.

Berck, P. and M. Roberts (1996), "Natural resource prices: Will they ever turn up," *Journal of Environmental Economics and Management* 9: 122–137.

Bernard, J.-T., L. Khalaf, M. Kichian, and S. McMahon (2008), "Forecasting commodity prices: GARCH, jumps, and mean-reversion," *Journal of Forecasting* 27: 279–291. http://dx.doi.org/10.1002/for.1061.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2003), "Alternative models for stock price dynamics," *Journal of Econometrics* 116: 225–57. http://dx.doi.org/10.1016/S0304-4076(03)00108-8.

Chinn, M. and O. Coibion (2013), "The predictive content of commodity futures," *Journal of Futures Markets* forthcoming. http://dx.doi.org/10.1002/fut.21615.

Cortazar, G. and L. Naranjo (2006), "An N-factor gaussian model for oil futures prices," *Journal of Futures Markets* 26: 243–268. http://dx.doi.org/10.1002/fut.20198.

Cortazar, G. and E. S. Schwartz (2003), "Implementing a stochastic model for oil future prices," *Energy Economics* 25: 215–238. http://dx.doi.org/10.1016/S0140-9883(02)00096-8

Crespo-Cuaresma, J., A. Jumah, and S. Karbuz (2009), "Modelling and forecasting oil prices: The role of asymmetric cycles," *The Energy Journal* 30(3): 81–90. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol30-No3-4.

Das, S. R. (2002), "The surprise element: Jumps in interest rates," *Journal of Econometrics* 106: 27–65. http://dx.doi.org/10.1016/S0304-4076(01)00085-9.

Diebold, F. and L. Kilian (2000), "Unit root tests are useful for selecting forecast models," *Journal of Business and Economic Statistics* 18: 265–73.

Drost, F. C., T. E. Nijman, and B. J. M. Werker (1998), "Estimation and testing in models containing both jump and conditional heteroskedasticity," *Journal of Business and Economic Statistics* 16: 237–43.

Fuertes, A.-M., J. Miffre, and G. Rallis (2010), "Tactical allocation in commodity futures markets: Combining momentum and term structure signals," *Journal of Banking and Finance* 34: 2530–48. http://dx.doi.org/10.1016/j.jbankfin.2010.04.009.

Hamilton, J. D. (2009), "Understanding crude oil prices," *The Energy Journal* 30(2): 179–206. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol30-No2-9.

Hamilton, J. and C. Wu (2012), "Risk premia in crude oil futures prices," *Journal of Intenational Money and Finance* forthcoming.

Jorion, P. (1988), "On jump processes in the foreign exchange and stock markets," *Review of Financial Studies* 1: 427–45. http://dx.doi.org/10.1093/rfs/1.4.427.

Khalaf, L., J.-D. Saphores, and J.-F. Bilodeau (2003), "Simulation-based exact jump tests in models with conditional heteroskedasticity," *Journal of Economic Dynamic and Control* 28: 531–553. http://dx.doi.org/10.1016/S0165-1889(03)00034-4.

Kilian, L. (2009), "Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market," *American Economic Review* 99: 1053–1069. http://dx.doi.org/10.1257/aer.99.3.1053.

Kim, C.-J. and C. R. Nelson (1999), *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, The MIT Press, Cambridge, Massachusetts.

Lee, J., J. A. List, and M. C. Strazicich (2006), "Non-renewable resource prices: Deterministic or stochastic trends?," *Journal of Environmental Economics and Management* 51: 354–370. http://dx.doi.org/10.1016/j.jeem.2005.09.005.

Merton, R. C. (1976), "Option pricing when underlying stock returns are discontinuous," *Journal of Financial Economics* 3: 125–44. http://dx.doi.org/10.1016/0304-405X(76)90022-2.

Moshiri, S. and F. Foroutan (2006), "Forecasting nonlinear crude oil future prices," *The Energy Journal* 27(4): 81–85. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol27-No4-4.

Pindyck, R. S. (1999), "The long run evolution of energy prices," *The Energy Journal* 20(2): 1–27. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol20-No2-1.

Pindyck, R. S. (2001), "The dynamics of commodity spot and futures markets: a primer," *The Energy Journal* 22(3): 1–29. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol22-No3-1.

Postali, F. A. S. and P. Picchetti (2006), "Geometric brownian motion and structural breaks in oil prices: A quantitative analysis," *Energy Economics* 28: 506–522. http://dx.doi.org/10.1016/j.eneco.2006.02.011.

Regnier, E. (2007), "Oil and energy price volatility," *Energy Economics* 29: 405–427. http://dx.doi.org/10.1016/j.eneco.2005.11.003.

Routledge, B., S. D. and C. Spratt (2000), "Equilibrium forward curves for commodities," *Journal of Finance* 55: 1297–1338. http://dx.doi.org/10.1111/0022-1082.00248.

Sadorsky, P. (2006), "Modeling and forecasting petroleum futures volatility," *Energy Economics* 28: 467–488. http://dx.doi.org/10.1016/j.eneco.2006.04.005.

Schwartz, E. S. and J. E. Smith (2000), "Short-term variations and long-term dynamics in commodity prices," *Management Science* 46: 893–911. http://dx.doi.org/10.1287/mnsc.46.7.893.12034.

Smith, J. L. (2009), "World oil: Market or mayhem," *Journal of Economic Perspectives* 23: 145–164. http://dx.doi.org/10.1257/jep.23.3.145.

Tabak, B. M. and D. O. Cajueiro (2007), "Are the crude oil markets becoming weakly efficient over time? a test for time-varying long-range dependence in prices and volatility," *Energy Economics* 29: 28–36. http://dx.doi.org/10.1016/j.eneco.2006.06.007.

Trolle, A. and E. Schwartz (2009), "Unspanned stochastic volatility and the pricing of commodity derivatives," *Review of Financial Studies* 22: 4423–4461. http://dx.doi.org/10.1093/rfs/hhp036.

Wilson, B., R. Aggarwal, and C. Inclan (1996), "Detecting volatility changes across the oil sector," *Journal of Futures Markets* 16(3): 313–330. http://dx.doi.org/10.1002/(SICI)1096-9934(199605)16:3＜313::AID-FUT4＞3.0.CO;2-M.

Wirl, F. (2008), "Why do oil prices jump (or fall)?," *Energy Policy* 36(3): 1029–1043. http://dx.doi.org/10.1016/j.enpol.2007.11.024.

Wu, T. and A. McCallum (2005), "Do oil futures prices help predict future oil prices?," Federal Resesrve Bank of San Francisco Economic Letter 2005.

APPENDIX

Tables 1 and 3 show the oil price forecasts for weekly and monthly data, while Tables 2 and 4 present the corresponding forecasts relative to the forecasts obtained from the GARCH(1,1) model. Table 5 shows, for a few illustrative model cases, the average values of the obtained parameter estimates over the considered forecast horizon. In particular, for both the weekly and monthly data, we report the average of the estimates over the 1-year and 5-year horizons for the Schwartz and Smith (2000) and the GARCH (1,1) models, and for the case where the 4-month futures is used.

Table 1: Oil Price Forecast Errors, Weekly Data

Weekly Frequency	1 month Futures			3 month Futures		
Forecast Horizon	1 year	3 years	5 years	1 year	3 years	5 year
Schwartz & Smith	**0.0022**	**0.0028**	**0.0033**	**0.0018**	**0.0021**	**0.0026**
GARCH(1,1)	0.0439	0.0160	0.0115	0.0461	0.0164	0.0109
GARCH-M(1,1)	0.0512	0.0194	0.0145	0.0542	0.0202	0.0146
EGARCH(1,1)	0.0441	0.0162	0.0116	0.0461	0.0164	0.0109
GARCH-*t*(1,1)	0.2622	0.1222	0.1171	0.2237	0.1015	0.0880
ARCH(1) with jumps	**0.0022**	0.0163	0.0117	0.0022	0.0177	0.0118
GARCH(1,1) with jumps	0.0438	0.0160	0.0115	0.0461	0.0164	0.0109
GARCH-*t*(1,1) with jumps	0.0438	0.0160	0.0114	0.0461	0.0164	0.0109
Weekly Frequency	2 months Futures			4 months Futures		
Forecast Horizon	1 year	3 years	5 years	1 year	3 years	5 years
Schwartz & Smith	**0.0020**	**0.0023**	**0.0028**	**0.0017**	**0.0020**	**0.0025**
GARCH(1,1)	0.0450	0.0161	0.0109	0.0109	0.0166	0.0109
GARCH-M(1,1)	0.0532	0.0198	0.0143	0.0160	0.0195	0.0138
EGARCH(1,1)	0.0449	0.0161	0.0109	0.0109	0.0166	0.0109
GARCH-*t*(1,1)	0.2758	0.1255	0.1107	0.0880	0.0843	0.0719
ARCH(1) with jumps	**0.0020**	0.0164	0.0112	**0.0017**	0.0165	0.0108
GARCH(1,1) with jumps	0.0449	0.0161	0.0115	0.0109	0.0166	0.0166
GARCH-*t*(1,1) with jumps	0.0449	0.0161	0.0109	0.0109	0.0166	0.0166

Note: Numbers shown are MSPE; numbers in bold refer to the model which minimizes the forecast error measure for a given futures series and a given forecast horizon.

Table 2: Relative Oil Price Forecast Errors, Weekly Data

Weekly Frequency	1 month Futures			3 month Futures		
Forecast Horizon	1 year	3 years	5 years	1 year	3 years	5 year
Schwartz & Smith	0.05	0.17	0.28	0.04	0.13	0.24
GARCH-M(1,1)	1.17	1.21	1.27	1.18	1.23	1.34
EGARCH(1,1)	1.01	1.01	1.01	1.00	1.00	1.00
GARCH-*t*(1,1)	5.98	7.63	10.23	4.85	6.19	8.06
ARCH(1) with jumps	0.05	1.02	1.02	0.48	1.08	1.08
GARCH(1,1) with jumps	1.00	1.00	1.00	1.00	1.00	1.00
GARCH-*t*(1,1) with jumps	1.00	1.00	1.00	1.00	1.00	1.00
Weekly Frequency	2 months Futures			4 months Futures		
Forecast Horizon	1 year	3 years	5 years	1 year	3 years	5 years
Schwartz & Smith	0.04	0.14	0.26	0.16	0.12	0.23
GARCH-M(1,1)	1.18	1.23	1.31	1.47	1.17	1.26
EGARCH(1,1)	1.00	1.00	1.00	1.00	1.00	1.00
GARCH-*t*(1,1)	6.13	7.78	10.13	8.06	5.07	6.57
ARCH(1) with jumps	0.04	1.02	1.02	0.16	0.99	0.99
GARCH(1,1) with jumps	1.00	1.00	1.05	1.00	1.00	1.52
GARCH-*t*(1,1) with jumps	1.00	1.00	1.00	1.00	1.00	1.00

Note: Numbers shown are the MSPE of a given model relative the MSPE of the corresponding GARCH(1,1).

Table 3: Oil Price Forecast Errors, Monthly Data

Monthly Frequency	1 month Futures			3 month Futures		
Forecast Horizon	1 year	3 years	5 years	1 year	3 years	5 year
Schwartz & Smith	0.0086	0.0117	**0.0135**	0.0067	0.0080	**0.0100**
GARCH(1,1)	0.0079	0.0108	0.0138	**0.0063**	**0.0078**	0.0104
GARCH-M(1,1)	0.0079	**0.0105**	0.0140	0.0066	**0.0078**	0.0108
EGARCH(1,1)	0.0081	0.0110	**0.0135**	0.0066	0.0080	0.0101
GARCH-*t*(1,1)	0.0079	0.0109	0.0137	0.0065	0.0080	0.0101
ARCH(1) with jumps	0.0080	0.0109	0.0136	0.0065	0.0083	0.0102
GARCH(1,1) with jumps	**0.0078**	0.0107	0.0137	**0.0063**	0.0080	0.0103
GARCH-*t*(1,1) with jumps	0.0079	0.0108	0.0137	**0.0063**	**0.0078**	0.0103
Monthly Frequency	2 months Futures			4 months Futures		
Forecast Horizon	1 year	3 years	5 years	1 year	3 years	5 years
Schwartz & Smith	0.0074	0.0090	**0.0111**	0.0062	0.0075	0.0095
GARCH(1,1)	**0.0070**	0.0088	0.0116	**0.0060**	**0.0073**	0.0097
GARCH-M(1,1)	0.0072	0.0087	0.0116	0.0063	0.0074	0.0102
EGARCH(1,1)	0.0073	0.0089	0.0114	0.0066	0.0080	**0.0094**
GARCH-*t*(1,1)	0.0072	0.0089	0.0114	0.0063	0.0076	**0.0094**
ARCH(1) with jumps	0.0073	0.0088	0.0113	0.0066	0.0080	**0.0094**
GARCH(1,1) with jumps	0.0071	**0.0085**	0.0115	0.0061	**0.0073**	0.0096
GARCH-*t*(1,1) with jumps	0.0071	0.0098	0.0115	**0.0060**	0.0074	0.0096

Note: Numbers shown are MSPE; numbers in bold refer to the model which minimizes the forecast error measure for a given futures series and a given forecast horizon.

Table 4: Relative Oil Price Forecast Errors, Monthly Data

Monthly Frequency	1 month Futures			3 month Futures		
Forecast Horizon	1 year	3 years	5 years	1 year	3 years	5 year
Schwartz & Smith	1.10	1.08	0.98	1.06	1.02	0.96
GARCH-M(1,1)	1.00	0.97	1.02	1.04	1.02	1.03
EGARCH(1,1)	1.03	1.01	0.98	1.04	1.02	0.97
GARCH-*t*(1,1)	1.01	0.99	0.99	1.04	1.02	0.97
ARCH(1) with jumps	1.01	1.00	0.99	1.04	1.06	0.98
GARCH(1,1) with jumps	1.00	0.98	1.00	1.00	1.03	0.99
GARCH-*t*(1,1) with jumps	1.01	1.00	0.99	1.01	1.00	0.98
Monthly Frequency	2 months Futures			4 months Futures		
Forecast Horizon	1 year	3 years	5 years	1 year	3 years	5 years
Schwartz & Smith	1.06	1.02	0.95	1.06	1.02	0.97
GARCH-M(1,1)	1.02	0.99	0.99	1.06	1.01	1.04
EGARCH(1,1)	1.04	1.02	0.98	1.11	1.08	0.97
GARCH-*t*(1,1)	1.02	1.01	0.98	1.05	1.03	0.97
ARCH(1) with jumps	1.05	1.00	0.97	1.05	1.03	0.97
GARCH(1,1) with jumps	1.01	0.97	0.99	1.02	0.99	0.98
GARCH-*t*(1,1) with jumps	1.01	1.12	0.99	1.01	1.01	0.99

Note: Numbers shown are the MSPE of a given model relative the MSPE of the corresponding GARCH(1,1).

Table 5: Averages of Parameter Estimates for Select Models

	Schwartz & Smith							GARCH(1,1)			
Hzn	κ	λ_χ	σ_χ	μ_ξ	λ_ξ	σ_ξ	$\rho_{\chi\xi}$	μ	α_0	α_1	ϕ
Weekly Frequency											
1Yr	0.1654	0.2104	1.0012	0.0012	0.0496	0.6365	−0.1301	−0.04	0.0001	0.854	0.12
5Yr	0.1601	0.2237	1.0011	0.0011	0.0364	0.6500	−0.1438	−0.04	0.0001	0.845	0.12
Monthly Frequency											
1Yr	0.2083	0.3048	1.0069	0.0069	0.1769	0.7131	0.0483	−0.05	0.0016	0.557	0.23
5Yr	0.1988	0.3211	1.0069	0.0069	0.1597	0.7293	0.0332	−0.05	0.0017	0.541	0.22