Small Trends and Big Cycles in Crude Oil Prices

Xiaoyi Mu and Haichun Ye***

ABSTRACT

We employ an unobserved components model to disentangle the long-term trend from cyclical movements in the price of internationally traded crude oil using data from 1861 to 2010. The in-sample estimation of the model identifies a deterministic quadratic trend and two types of cycles, with the short cycle having a period of 6 years and the long cycle of 29 years. Compared to the large amplitude of the cycles, the growth rate of the long-term trend is small. The out-of-sample forecasting performance of various competing models is compared to that of a "no change" random walk forecast. While the random walk forecast tends to be the most accurate at shorter horizons, it is outperformed by the trend-cycle models at horizons longer than one year. The results provide evidence of predictability in the price of crude oil at long horizons.

Keywords: Oil price, Unobserved components model, Cycle, Trend, Forecast

http://dx.doi.org/10.5547/01956574.36.1.3

1. INTRODUCTION

The continued surge in oil prices since 2001 has renewed the interests in the questions of whether global oil production has peaked or will soon peak, whether the price increase reflects a long-run upward trend arising from "scarcity rent" or simply a cyclical movement, and whether the price of oil is predictable at all (see, for example, Geman [2005], Hamilton [2009], Alquist and Kilian [2010], and Alquist, Kilian and Vigfusson [2012]). In this paper, we employ an unobserved components (UC) model to disentangle the long-run trend from shorter-term cyclical patterns in the price of internationally traded crude oil using data from 1861 to 2010. Our objective is twofold. First, we are interested in testing both the shape of the long-run trend and the patterns of cycles. Second, we will examine the forecastability of oil price at longer horizons, which should matter more to investment decisions as well as policy-making.

A long-standing debate in the natural resource economics literature is whether the price of an exhaustible resource product rises in the long run. The Hotelling (1931) model suggests that the price (net of marginal cost) of an exhaustible resource will rise at the rate of interest rate in a competitive market equilibrium. Modifications of the basic Hotelling model yield a decreasing, flat or U-shaped long-run price path (see Nordhaus [1973], Fisher [1979], Pindyck [1978] and Slade [1982] among others). In particular, Pindyck (1978) and Slade (1982) demonstrate that the price of a non-renewable resource product could initially decline as reserves accumulate or as technological advancement lowers the marginal extraction cost, but over time, as the depletion effect

* Corresponding author. Centre for Energy, Petroleum and Mineral Law and Policy, University of Dundee, Carnegie Building, Dundee, DD1 4HN, UK. E-mail: x.mu@dundee.ac.uk. Tel: +44(0)1382 384843, Fax: +44(0)1382 385854.

** School of Economics, Shanghai University of Finance and Economics, Key Laboratory of Mathematical Economics, Ministry of Education, 777 Guoding Road, Shanghai 200433, China. Email:haichunye@gmail.com.

dominates the effect of technological change, the price could increase, resulting in a U-shaped long-run price path. Krautkraemer (1998) surveys the empirical literature and finds that, while there is little empirical support for the original Hotelling prediction, the conclusion regarding the long-run price trend varies considerably with the sample period and the commodities chosen (see also Slade and Thille [2009] and Livernois [2009]).[1] In this paper, we re-examine the long-run time series properties of oil prices with extended data and establish some stylized facts, which can shed light on new theories about the price of oil over the long run.

We first demonstrate that the real price of crude oil is stationary around a U-shaped long-run trend with a significant structural break in 1973/1974, corroborating some of the earlier findings. We then estimate an unobserved components model to determine the trend and cycle patterns. Two types of cycles stand out, with the short cycle having a period of six years and the long cycle having an average length of 29 years. The amplitude of the long cycle can be as much as 90 percent above or 45 percent below the long-run trend. Compared to the cyclical movements, however, the quadratic trend is small. To our knowledge, no previous studies have quantified the relative importance of long-run trend and cycles of crude oil prices in a unified framework while accounting for the presence of structural breaks.

The major contribution of our work lies in the out-of-sample forecastability of oil price at long horizons. An important literature has developed recently that empirically examines the forecastability of oil prices. At short horizons up to one year, Alquist and Kilian (2010) have shown that a “no change” random walk model tends to predict the monthly price of oil more accurately than the futures prices, reduced-form time series models, and an interest rate based Hotelling model, while Alquist, Kilian and Vigfusson (hereafter AKV, 2012) and Baumeister and Kilian (2012) show that models incorporating lagged values of global real activity and oil inventories can have significant forecasting power. As for long-horizon forecasts, AKV(2012) finds that, while oil futures prices fail to improve forecast accuracy relative to the no-change forecasts, forecasts obtained by adjusting current oil price for expected inflation outperform the no-change forecasts for the nominal price of oil at the horizon of five years. In this study, our out-of-sample forecasting analysis suggests that, while the no-change forecasts tend to be the most accurate at the one-year horizon among the models considered, the trend-cycle models significantly improve the forecast accuracy at horizons of five and ten years. The results hold for both real and nominal oil prices and provide robust evidence that the oil price is predictable at long horizons.

The rest of the paper is set up as follows. Section 2 briefly reviews the empirical literature on the long-run behaviour of oil prices. Section 3 describes the data and modelling approach. In-sample model estimation is presented in Section 4 and out-of-sample forecast accuracy evaluation results in Section 5. Section 6 concludes.

2. THE LITERATURE ON LONG-RUN BEHAVIOUR OF OIL PRICES

The long-run trend of the crude oil price has been empirically studied by a number of authors in the literature. Using data from 1870 to 1978, Slade (1982) finds statistically significant evidence for a U-shaped curve, represented by a quadratic time trend in the prices of crude oil and ten other commodities out of the 12 price series she tested. In an influential paper, Pindyck (1999) shows that the prices of energy products (oil, gas and coal) are mean-reverting to a quadratic trend

1. Another related literature represented by the Prebisch-Singer thesis is whether there is a secular decline in the relative price of primary goods, a survey of which can be found in Bleaney and Greenaway (1993).

line but the rate of reversion is slow, and for practical applications, the random walk assumption is not bad at all. Armed with advancements in time series techniques, a series of papers have tested whether the trend in natural resource commodity prices is deterministic or stochastic. Extending Slade's (1982) data to 1990, Berk and Roberts (1996) find the prices of oil and other commodities are non-stationary based on the Lagrange-Multiplier test and Dickey-Fuller test. Using the same data but with different unit root test techniques, Ahrens and Sharma (1997) conclude that among the 11 commodity price series, the price of oil along with five others is stationary around a deterministic trend. Still with the same data, Lee, List and Strazicich (hereinafter LLS, 2006) employ a Lagrange multiplier test allowing up to two endogenously determined structural breaks and reject the unit root hypothesis for all of the 11 commodities.

Note that all the studies reviewed above focus solely on the trend component of oil prices and ignore the potential cyclical behaviour of oil prices. Dvir and Rogoff (2010) examine changes in persistence and volatility of crude oil prices across three periods from 1861 to 2009 and document striking similarities between the periods of 1861–1878 and 1973–2009. Cashin and McDermott (2002) look at the long-run behaviour of a real commodity price index that consists of non-food agricultural products and industrial metals and document a small downward trend along with the large variability in real commodity prices for the period 1862–1999. Erten and Ocampo (2012) and Zellou and Cuddington (2012) are the two studies that have explicitly modelled the long-run trend and cyclical behaviour of crude oil prices. Erten and Ocampo (2012) use an asymmetric band-pass filter to decompose crude oil prices from 1875 to 2010 into a trend component and super-cycles, defined as periods lasting from 20 to 70 years. They find an upward trend in oil price before the 1920s, a modest downward trend between the 1920s and the 1960s and an upward trend afterwards, and show that the super-cycles in the price of crude oil were rather modest in the early twentieth century and became more pronounced after the 1970s. Applying a similar band-pass filter method to the annual crude oil price data over the period 1861–2010, Zellou and Cuddington (2012) find that there is strong evidence for super-cycles in oil price during the post-WWI period but weak evidence in the pre-WWI period and that the upward long-term trend for the post-WWI period is quite small relative to the super-cycles.

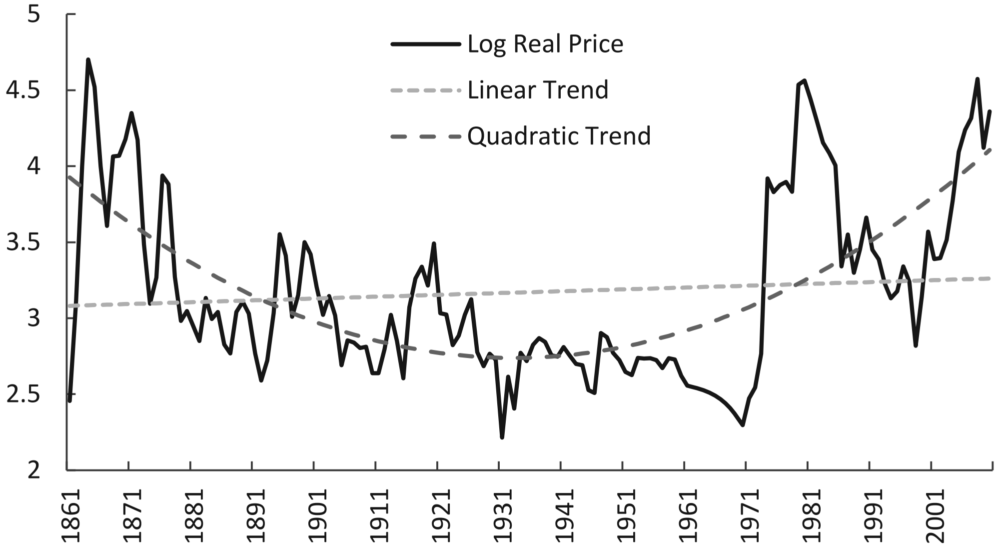
While our conclusions about the trend and cycle behaviour of crude oil price are somewhat similar to those in Erten and Ocampo (2012) and Zellou and Cuddington (2012), this study differs from theirs in several important aspects. First, we take the UC model approach to model the trend and cycles in oil prices. The UC model is more general than the band-pass filters typically used to extract the trend and cycles in oil prices in the sense that the band-pass filters have the length of cycles predefined and can be obtained as special cases of the UC model (Harvey and Trimbur, 2003). Particularly, the model-based approach automatically adapts to the endpoints of a sample and allows us to extract trend and cycles by filters that are optimal and mutually consistent even at the beginning and end of the series. Second, and perhaps more importantly, we focus more on the ability of the trend and cycle models to improve out-of-sample forecast accuracy at long horizons, which has not been done in the two cited super-cycles papers. Last, we allow for structural breaks when modelling the long-run trend and cyclical movements of real oil prices.

3. DATA AND MODELLING APPROACH

Our oil price data are from the BP Statistical Review of the World Energy (2009 and 2010) and are the annual averages of oil prices in the US for 1861–1944, Arabian Light crude oil for 1945–1983 and Brent crude oil for 1984–2010.[2] The switch in the type of crude oils reflects the

2. The data is also used in Hamilton (2011) and Dvir and Rogoff (2010).

Figure 1: Log real oil price and deterministic trend lines: 1861–2010



Notes: The solid line indicates the natural logarithm of annual real oil price over the period 1861–2010. The dotted line indicates the fitted deterministic linear trend. The dashed line indicates the fitted deterministic quadratic trend.

evolution of the oil market. A more detailed description of the data is provided in the appendix. The real price expressed in 2009 dollar values is used in our analysis.

Figure 1 plots the logged real oil price (p), along with its linear and quadratic trend lines, over the full sample period.[3] Even a cursory look reveals several important features. First, there is a quite dramatic change in the trend in the early 1970s, around the time of the first oil shock. The real oil price was generally declining until 1970, after which it does not exhibit a clear trend. Second, the real oil price fluctuates considerably between periods of upswings and downswings, but appears to be reverting to an evolving trend. For example, it rose between the late 1860s and the early 1870s, and then fell from the early to mid-1870s. Since the early 1970s, the real oil price witnessed two waves of strong upswings, from the early 1970s to the early 1980s and from the late 1990s to the late 2000s, and an extended downswing in between. Last, the quadratic trend appears to fit the data much better than the linear one. While the linear trend yields a very weak upward trending price curve, the quadratic trend clearly displays a U-shaped pattern for the real oil price. Moreover, the quadratic trend helps reduce persistence in the data. The half-life of the log real oil price (p) is 5.18 years, drops to 2.64 years for the quadratically detrended p but rises to 5.28 years for the linearly detrended p.

3.1 Testing for Structural Breaks

Considering the potential presence of structural breaks in the data, we begin our analysis with a test of structural breaks. We employ the multiple structural breaks analysis developed by Bai and Perron (1998, 2003) to endogenously determine the number and location of structural

3. A lower case p denotes the logged real oil price and an upper case P denotes the real price in levels.

Table 1: Structural Breaks in Log Real Oil Price: 1861–2010

	(a) Simple Level Breaks Model	(b) Linear Trend With Level Breaks	(c) Quadratic Trend With Level Breaks
Test statistics:			
H_0: No breaks; H_A: At least one break.			
$SupF_T(1)$	4.5334	1.1592	1.9088
$SupF_T(2)$	7.4609**	7.7697**	3.4632
$SupF_T(3)$	9.0601***	507.5711***	6.9938***
UD max	9.0601**	507.5711***	6.938*
WD max	16.4403***	921.0332***	11.3168**
H_0: i breaks; H_A: i + 1 breaks.			
$SupF_T(2\|1)$	6.1205	3.0364	1.7268
$SupF_T(3\|2)$	4.6775	0.0367	0.0000
No. of breaks	1	1	1
Break date	1973	1973	1973

Notes: Up to 3 breaks are allowed in the estimation. The trimming value is set to equal to 0.2 so that each regime has at least 30 observations. Serial correlation in the errors and heterogeneous variances of the residuals across regimes are also allowed in the estimation. The superscripts, ***, **, and *, denote the rejection of the null at the significance levels of 1%, 5% and 10%, respectively.

breaks in the log real price series. Specifically, we estimate the following three *m*-break models, including a simple level shift model, a linear trend model with level shifts and also a quadratic trend model with level shifts:[4]

$$p_t = \delta_j + u_t, \tag{1}$$

$$p_t = \delta_j + \varphi t + u_t, \tag{2}$$

$$p_t = \delta_j + \varphi t + \kappa t^2 + u_t, \tag{3}$$

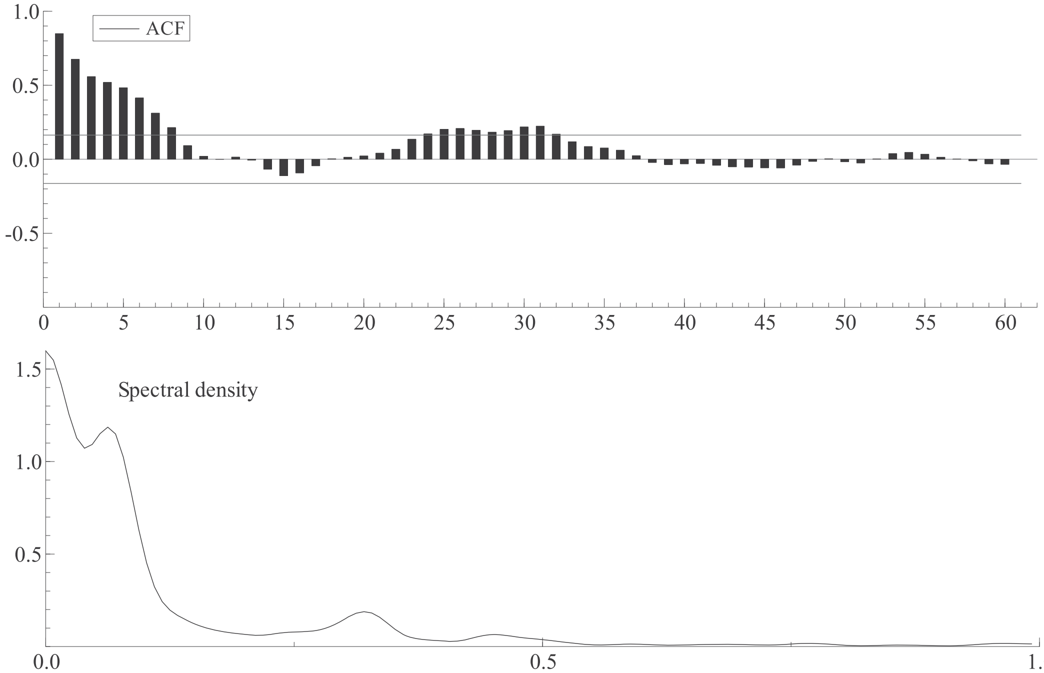
for $t = T_{j-1} + 1, \ldots, T_j$, $j = 1, \ldots, m + 1$. In these models, p_t is the log real price for year t, δ_j is the intercept specific to regime j, t is the deterministic time trend, φ and κ are time-invariant coefficients on the trend terms, and $(T_1, T_2, \ldots, T_m)$ represents the location of structural breaks. The number of breaks is determined by the $SupF_T(l)$ tests combined with the sequential $SupF_T(l + 1 | l)$ test while the location of each break is identified by globally minimizing the sum of squared residuals from the above equations. In the estimation, we allow for up to three breaks. Serial correlation in the errors and heterogeneous variances of the residuals across regimes are also allowed in the estimation.

Table 1 presents the Bai and Perron test results from the simple level shift model, the linear trend with level shifts model and the quadratic trend with level shift models, respectively. In all three cases, the $SupF_T(l)$ tests and the double maximum tests (*UDmax* and *WDmax*) generally reject the null of no break at conventional significance levels, while the sequential $SupF_T(l + 1 | l)$ tests fail to reject one break in favour of two. This thus leads us to conclude that there is only one break in these level-break models. Based on Bai and Perron's global optimization algorithm, we find that this level-break occurred in the year of 1973 in all three level-break models.[5] The identified

4. We also tried to allow for breaks in the intercept and trend slope in the linear and quadratic trend models but did not find significant evidence for shifting slopes in either model.

5. The finding of a level break in 1973 is consistent with the conclusion reached in AKV (2012) based on predictive regressions. We also applied the Quandt-Andrews unknown break point test to the three level shift models and found statistically significant evidence for a level-break in 1974 in all three models.

Figure 2: The autocorrelation and spectral density of the log real oil price



Notes: The upper panel shows the autocorrelation coefficients of the mean-removed log real oil price (*dmlrp*). The lower panel shows the spectral density of the *dmlrp*.

break year clearly points to the first "oil price shock" arising from the Arab oil embargo of 1973–74, which quadrupled the nominal oil price. It is also consistent with the view that the international oil market has changed from being led by multinational oil companies before the oil crisis to being dominated by OPEC, which in turn has been dominated by Saudi Arabia, afterwards.[6] Given the presence of level breaks in the log real oil price series, we then perform Lee and Strazicichi's (2001) LM unit root test with one endogenous break and also Zivot and Andrews (1992) Minimum *t* test to check if the series still has a unit root while allowing for level shifts. Both tests reject the null of a unit root at least at the five percent level, suggesting that the log real oil price series is stationary around a broken trend.[7]

3.2 Detecting Cyclical Movements

The tendency for real oil price to oscillate between upswings and downswings motivates us to formally test for the presence of cycles in the data. Figure 2 presents the correlogram and the estimated spectrum density of the shifting-mean-removed data, respectively. In addition to the

6. For empirical studies on the role of OPEC in the world oil market post-1973, see Alhajji and Huettner (2000) and more recently Almoguera et al. (2011). However, Kilian (2008, 2009) and AKV (2012) emphasize on the role of demand surge and changes in oil price regulation in the US.

7. In the above two unit root tests, only intercept is allowed to shift. The test statistics of the LM one-break test and the Zivot-Andrews test are −3.8696 and −7.0999, respectively. It is also worth noting that, while the one-break LM test identifies 1920 as the break year, the Zivot-Andrews test identifies 1974 as the break year.

significant amount of autocorrelation in the data, the correlogram also indicates the possibility of cyclical movements in the data. The same message appears in the estimated spectral density, but more clearly. The sample spectrum indicates a cycle centred around 30 years. There is also a hint of a subsidiary cycle around six years.[8]

The presence of a cycle against the null hypothesis of no cycles can be formally tested using a Lagrange Multiplier test. As shown in Harvey (1985), the appropriate test statistic is constructed as follows:

$$LM = T \times r_1^2 \tag{4}$$

where T is the number of observations and r_1 is the first-order sample autocorrelation. Under the null hypothesis of no cycles, the test statistic has a chi-square distribution with one degree of freedom. In the case of the shifting-mean-removed data, $LM = 108.12$ and thus the null hypothesis is strongly rejected at one percent significance level.

3.3 The Unobserved Components Model

To jointly capture the long-run trend with structural breaks and also the cyclical behaviour of the oil price, we use a univariate unobserved components model of Harvey (1989) and Koopman et al. (2009). Specifically, the model decomposes the logged real oil price (p) into a time trend (μ_t), two cyclical components (φ_t and γ_t), and an irregular white noise term (ε_t):

$$p_t = \mu_t + \gamma_t + \varphi_t + \delta w_t + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2) \tag{5.1}$$

where p_t is the log real oil price, μ_t the trend component, γ_t a short cycle component, φ_t a long cycle component, w_t a level shift dummy that takes on the value of one for the post-break period (i.e. post-1973) and zero otherwise, and ε_t the white noise, irregular component with a normal and independent distribution of mean zero and variance σ_ε^2. Additionally, if necessary, autoregressive terms can also be included in the model.

The trend component is specified as follows:

$$\mu_t = \mu_{t-1} + \beta_t + \eta_t, \;\; \eta_t \sim NID(0, \sigma_\eta^2) \tag{5.2}$$

$$\Delta^d \beta_t = \xi_t, \;\; \xi_t \sim NID(0, \sigma_\xi^2) \tag{5.3}$$

where β_t is the slope of the trend μ_t, and η_t and ξ_t are normally and independently distributed (NID) errors with mean zero and variance σ_η^2 and σ_ξ^2, respectively. The irregular component ε_t, the level disturbance η_t and slope disturbance ξ_t are assumed to be mutually uncorrelated. The trend component is flexibly specified such that it nests a random walk with drift model, a smooth trend model, and a deterministic trend model. When $\sigma_\eta^2 > 0$ and $\sigma_\xi^2 = 0$, the slope is fixed and the trend reduces to a random walk with drift. When $\sigma_\eta^2 = 0$ and $\sigma_\xi^2 > 0$, it becomes a smooth trend (also known as an integrated random walk trend).[9] If both variances σ_η^2 and σ_ξ^2 are zero, the trend is deterministic.

8. Note that the frequency shown in the lower panel of Figure 2 is scaled by π. So the period is given by dividing two by the scaled frequency. The significant peaks in the spectrum at the scaled frequencies of around 0.06 and 0.3 indicate a cycle of around 30 years and a subsidiary cycle about six years, respectively.

9. Nyblom and Harvey (2001) argue that the smooth trend model is equivalent to the trend obtained by fitting a cubic spline.

The parameter d in equation (5.3) defines the order of the trend. When $d = 1$, the trend is linear. When $d = 2$, it is quadratic.

Based on the visual examination of the correlogram and spectrum of the data and also some preliminary experimentation, we formulate our model with two stochastic cycles, a short business cycle type of around five years and a long cycle lasting over 20 years. Each cyclical component is represented by a mixture of sine and cosine waves:

$$\begin{bmatrix} \varphi_t \\ \varphi_t^* \end{bmatrix} = \rho_\varphi \begin{bmatrix} \cos\lambda_\varphi & \sin\lambda_\varphi \\ -\sin\lambda_\varphi & \cos\lambda_\varphi \end{bmatrix} \begin{bmatrix} \varphi_{t-1} \\ \varphi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_{\varphi,t} \\ \kappa_{\varphi,t}^* \end{bmatrix} \tag{5.4}$$

where ρ_φ is a damping factor such that $|\rho_\varphi| < 1$, λ_φ is the frequency, in radians, between 0 and π, and $\kappa_{\varphi,t}$ and $\kappa^*{}_{\varphi,t}$ are two mutually independent Gaussian white noise disturbances with zero means and common variance $\sigma^2_{\kappa\varphi}$. Thus, this stochastic cycle specification generates a stationary cyclical process with a cycle period of $2\pi/\lambda_\varphi$. A similar specification holds for the short cycle component γ_t.

The model outlined above is very flexible in that it nests a variety of stochastic and deterministic trend and cycle specifications. For example, if only a stochastic level is specified in the model, the model reduces to a simple random walk plus noise model. If the variances of both trend components are restricted to zero and no cycles included, the model becomes a deterministic trend model. In Sections 4 and 5, we shall estimate various candidate models under the framework of the unobserved components model and make comparisons in terms of their in-sample goodness-of-fit and also out-of-sample forecast accuracy. Specifically, we consider the following eight models:

M1: Random walk model (RW),
M2: Random walk with drift model (RWD),
M3: Deterministic linear trend model (LT);
M4: Deterministic quadratic trend model (QT);
M5: Deterministic linear trend model with a level break (LTB);
M6: Deterministic quadratic trend model with a level break (QTB);
M7: Linear trend with cycles and a level break (LTCB);
M8: Quadratic trend with cycles and a level break (QTCB).

Note that the first six models (i.e., M1 through M6) are comparable to those proposed in the literature. To clean out serial correlations in the data, we include two lags of the real oil price in the deterministic trend models (i.e., M3 through M6). In M5 and M6 we also add the level-break dummy to capture the structural break in 1973. Unlike the first six models, M7 and M8 contain two cycle components. In estimating these two trend-cycle models, the variances of the trend components (σ^2_η and σ^2_ξ) are first allowed to be non-zero. It turns out that, in the linear trend-cycle model (*LTCB*), the level component has a non-zero variance while the slope component has a zero variance and that, in the quadratic trend-cycle model (*QTCB*), both the level and slope components have zero variances. That is to say, the trend component of *LTCB* is actually a random walk with drift while that of *QTCB* is simply a deterministic quadratic trend.

For estimation purpose, all above UC models are first converted into state space forms and the unknown parameters of each model are then estimated by maximizing the log-likelihood function via Kalman filter. Once the estimates of the parameters are obtained, the unobserved components are extracted from actual observations using the Kalman filter and the associated

Table 2: The Goodness-of-Fit of Models

	PEV	R^2	AIC	BIC
		In-sample period:1861–1980		
RW	0.0761	0.7406	−2.5424	−2.4960
RWD	0.0758	0.7438	−2.5296	−2.4600
LT	0.0692	0.7660	−2.6206	−2.5509
QT	0.0666	0.7769	−2.6427	−2.5498
LTB	0.0534	0.8211	−2.8635	−2.7706
QTB	0.0530	0.8240	−2.8550	−2.7389
LTCB	0.0453	0.8481	−3.0275	−2.9346
QTCB	***0.0418***	***0.8613***	***−3.0926***	***−2.9765***
		Full-sample period:1861–2010		
RW	0.0732	0.7877	−2.5882	−2.5481
RWD	0.0730	0.7896	−2.5769	−2.5167
LT	0.0677	0.8049	−2.6524	−2.5922
QT	0.0647	0.8149	−2.6852	−2.6049
LTB	0.0555	0.8411	−2.8373	−2.7570
QTB	0.0549	0.8441	−2.8362	−2.7358
LTCB	0.0480	0.8626	−2.9828	−2.9025
QTCB	***0.0449***	***0.8723***	***−3.0357***	***−2.9354***

Notes: For model specifications and abbreviations, see the text. PEV denotes the prediction error variance. AIC is the Akaike information criterion. BIC is the Bayesian information criterion.

smoother.[10] The first 120 sample observations from 1861 to 1980 are used for in-sample model estimation and assessment; and the remaining 30 observations over the period from 1981 to 2010 are reserved for analysing the out-of-sample forecasting accuracy of the models.

4. IN-SAMPLE RESULTS

The upper panel of Table 2 reports the goodness-of-fit measures of the above eight models over the in-sample period (1861–1980). To check the sensitivity of our results with respect to the sample period, we also report the results for the full sample period in the lower panel. The results are largely consistent across the two samples. Several points are worth noting. First, the quadratic trend-cycle model (*QTCB*) produces the least prediction error variance (PEV), the largest R^2 and the smallest values of Akaike information criterion (AIC) and Bayesian information criterion (BIC) of all eight models in both samples. Second, models with a cyclical component (i.e., *LTCB* and *QTCB*) outperform others in all four measures of goodness-of-fit, indicating the importance of incorporating cycles in modelling the long-term time path of the log real oil price. Third, consistent with LLS's (2005) finding that structural breaks are important in explaining natural resource prices, models with structural breaks clearly outperform their counterparts without breaks (i.e., *LTB vs*. *LT* and *QTB vs*. *QT*) in terms of the AIC and BIC. Finally, comparing the quadratic trend models with the linear trend models (*QT* vs. *LT*, *QTB* vs. *LTB*, and *QTCB* vs. *LTCB*), we find that in general the quadratic trend provides a better fit to the data than the linear one in terms of PEV and R^2 except that the deterministic linear trend model with a level break (*LTB*) yields a smaller AIC and BIC than the deterministic quadratic trend model with a level break (*QTB*).[11] Note that the quadratic

10. See, for example, Harvey (1989) and Durbin and Koopman (2001) for details. All the estimation procedures are carried out with the software package STAMP 8.3.

11. The finding that *LTB* is better than *QTB* is not surprising because when the trend is small, a linear trend combined with a structural break provides a parsimonious fit to the data.

Table 3: Estimation Results of the Quadratic Trend-Cycle Model

	In-sample period (1861–1980)	Full-sample period (1861–2010)
Hyperparameters		
Cycle_short ($\sigma^2_{v\gamma}$)	0.0114	0.0117
	(2.651)	(2.117)
Cycle_long ($\sigma^2_{v\varphi}$)	0.0044	0.0051
	(1.019)	(0.915)
Irregular (σ^2_{ε})	0.0043	0.0055
	(1.000)	(1.000)
Short Cycle		
Period ($2\pi/\lambda$)	6.18	6.21
	[5.85,6.52]	[5.85,6.60]
Damping factor (ρ)	0.8947	0.8569
	[0.816, 0.942]	[0.767, 0.916]
Long Cycle		
Period ($2\pi/\lambda$)	27.76	28.81
	[24.89,30.99]	[26.55,31.26]
Damping factor (ρ)	0.9778	0.9843
	[0.944,0.9914]	[0.9612,0.9937]
Coefficients of final state		
Level (μ_t)	2.6788***	2.7461***
	(0.1259)	(0.2466)
Slope (β_t)	0.0024	0.0052
	(0.0047)	(0.0058)
β_t Slope ($\Delta\beta_t$)	0.0002**	0.0001**
	(0.0001)	(0.0001)
Level break 1974	1.2106***	1.0287***
	(0.1859)	(0.1693)
Specification Diagnostics		
R^2	0.8613	0.8723
Rd^2	0.4674	0.4009
Normality test	3.9296	7.9260
DW	1.9497	1.9347

Note: For hyperparameters, q ratios ($\sigma^2_v/\sigma^2_\varepsilon$) are reported in parentheses. Numbers in brackets indicate one standard deviation bands for cycle periods. Root mean standard errors of the final state are reported in parenthesis. *, **, and *** indicate the significance levels of 10%, 5%, and 1%, respectively.

trend fitted here is indeed U-shaped and deterministic, which is consistent with the predictions of Pindyck (1978) and Slade (1982).

Since the quadratic trend-cycle model (*QTCB*) provides the best fit to the data among all the models considered, we report its parameter estimates in Table 3. The left column shows the estimation results from the in-sample period and the right column presents the results from the full sample period. For the in-sample period, the diagnostics shown at the bottom of the table appear largely satisfactory. The estimated coefficients of the final state indicate clear evidence of an increasing quadratic trend and also an upward shift in price level since 1973.

As far as the long and short cycles are concerned, both are stochastic in the sense that their disturbances are found to have positive variances. Intuitively, the cycles in the oil price can be driven by a variety of demand and supply factors including demand shifts, geopolitical events, technological advances, and changes in market structure. Given that the occurrence and influences of these factors have a large degree of randomness, the timing, periodicity and amplitude of the cycles tend to have stochastic properties. While the short cycle has an estimated periodicity of 6.2 years, commensurate with the length of a typical business cycle, the long cycle has an estimated periodicity of about 29 years. The length of the long cycle may be surprising but is nevertheless plausible given the capital intensity, long gestation period, and uncertainties involved in upstream petroleum investment projects. Besides, both cycles are highly persistent, with a damping factor of around 0.85 for the short cycle and a damping factor of about 0.98 for the long cycle. The estimation results obtained from the full sample period are fairly similar.

The left three panels of Figure 3 present the decomposition of the logged real oil price over the full-sample period. As shown in the upper panel, the long-run trend of the real oil price is modestly U-shaped due to its small growth rate. Furthermore, it appears that the real crude oil price would have been on a rising path since the 1970s even in the absence of the structural break in 1973 to 1974, as the minimum was already attained in the late 1960s. The middle panel graphs the short cycle component extracted from the series of log real price. Except for the period before 1880, when the oil industry was still in an embryo stage, the amplitude of short cycles is relatively low, typically in the range of -0.25 to 0.25. Since p is in logarithm, this suggests that the impact of the short cycles is to shift the real oil price away from the trend by about -22 to 28 percent for most of the time. The long cycle exhibited in the lower panel has such a large amplitude that the real oil price is brought to 90 percent above or 45 percent below its long-run trend. If we define a cycle as a period that consists of a "slump," during which the oil price goes from peak to trough, and a "boom," during which the price goes from trough to peak, we can see clearly in Figure 3 five complete cycles from the late 1860s.

5. OUT-OF-SAMPLE FORECAST ACCURACY EVALUATION

The above analysis has shown that the trend-cycle models fit the data better than the simple trend or trend with break models. Whether, and to what extent, the trend-cycle models can improve forecasting performance is still unknown. In this section we further explore how well the trend-cycle models predict the future trajectory of oil price at various forecast horizons by comparing their out-of-sample forecast accuracy. Following Alquist and Kilian (2010), we benchmark the out-of-sample forecast performance of various models against a "no change" random walk forecast.

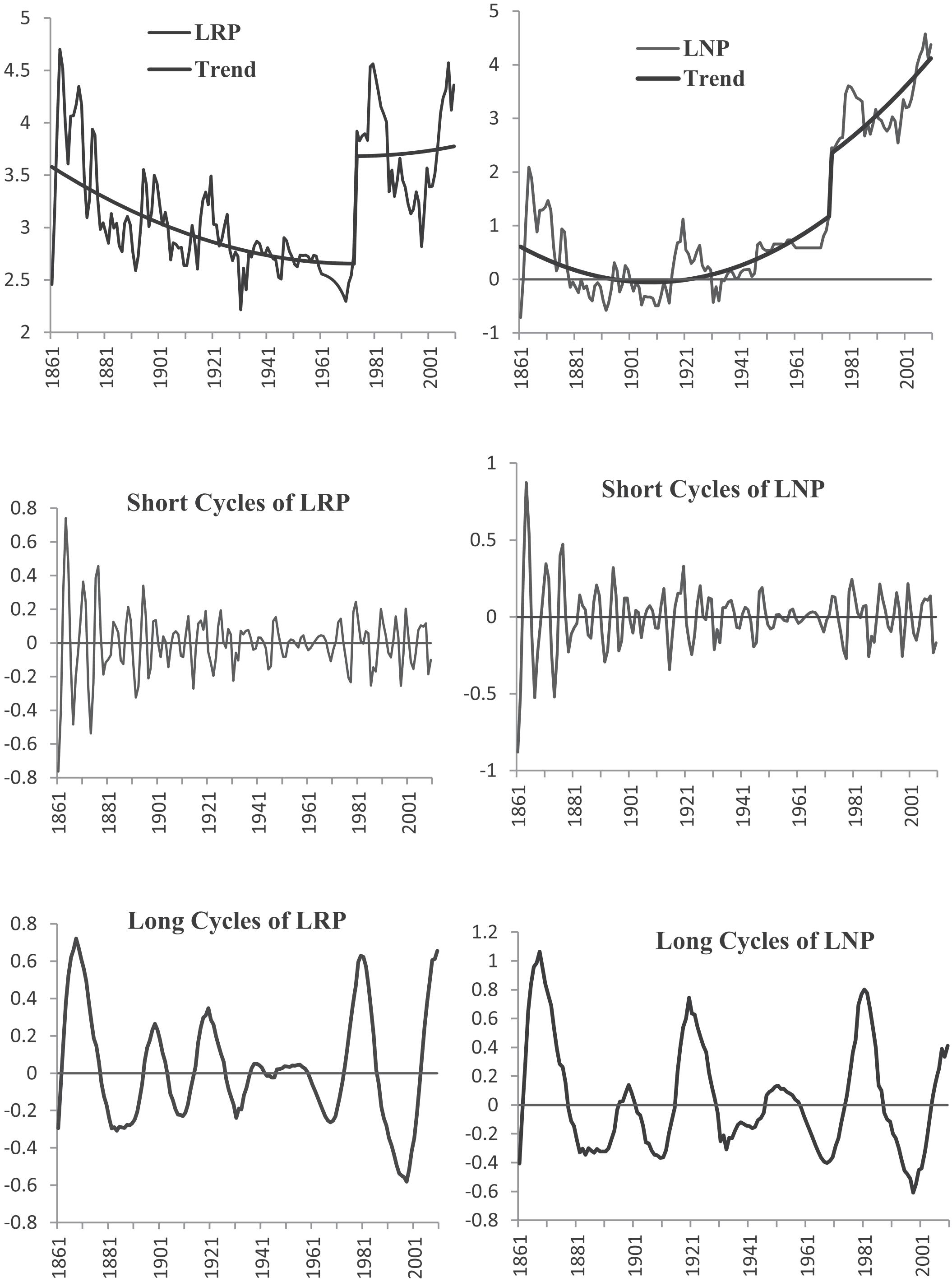
5.1 Models Considered and Evaluation Criteria

In addition to the eight models outlined in Section 3, following Alquist and Kilian (2010), we also include in our comparison list a Hotelling (1931) type model, which predicts that the real oil price rises at the rate of interest rate in the long run, and a "naive" forecast, which postulates that the h-step-ahead forecast of real oil price is simply the average price over the period $t-h+1$ to t.[12] Specifically, the price forecast for the Hotelling model is

M9: $P_{t+h|t} = P_t(1 + i_{t,h})^h$,

12. As demonstrated in Alquist and Kilian (2010), although the crude oil futures with maturities at one through seven years are available, the trading volumes for these long-maturity contracts are thin and the market remains illiquid even in recent years. So we do not consider the forecasting performance of futures market in this study.

Figure 3: The trend-cycle decomposition of logged oil prices over 1861–2010



Notes: LRP—the logged real price of oil; LNP—the logged nominal price of oil. The upper panel shows the logged real(nominal) oil price along with a quadrtic trend interrupted by a level break in 1973, the middle panel shows the short cycles and the lower panel shows the long cycles.

where P_t is the real oil price at t and $i_{t,h}$ refers to the annualized real interest rate at t for a borrowing period of h years. We use the yield rates of treasury bonds at the relevant maturity h, deflated by expected CPI, as a proxy for the interest rates. And the "naïve" forecast is

$$\text{M10: } P_{t+h|t} = \frac{1}{h}\sum_{m=1}^{h} P_{t-m+1}.$$

Note that, by construction, the one-step-ahead naïve forecast is the same as the no change forecast.

To assess the out-of-sample forecast accuracy of the above competing models, we first estimate the competing models for the in-sample period (1861 to 1980) and continuously update the sample, one year at a time, to obtain their respective recursive h-step-ahead forecast errors for the out-of-sample period. The forecast performance of these models is then evaluated by means of mean squared forecast error (MSFE) and mean absolute forecast error (MAFE). Specifically, we use the random walk model as the benchmark model and conduct the Diebold and Mariano (1995) test for the null of equal accuracy against the alternative that the candidate forecast is more accurate than the random walk forecast. In addition, we also compute the success ratio (i.e., the proportion of correctly predicted signs of oil price changes) and perform the directional accuracy test of Pesaran and Timmermann (2009) which allows for serial correlation to see whether a candidate model forecast has significant predictability on the signs of oil price changes. From a practical point of view, the directional accuracy test is perhaps more related to investment decisions than the MSFE or MAFE-based forecast accuracy tests.

5.2 Forecasting the Real Price of Oil

Table 4 reports the forecast performance of various candidate models against the random walk forecast at horizons of one, five, and ten years. To make our results more comparable with those in the oil price forecasting literature, we first obtain the forecasts of the logged real oil prices, convert them back to the level forecasts of real oil prices, and then use these level forecasts to calculate MSFEs and MAFEs.[13] For ease of comparison, we only present the actual values of MSFE and MAFE for the random walk model and express the MSFEs and MAFEs of other competing models in ratios relative to the random walk forecast.

At the horizon of one year, only the linear and quadratic trend models produce slightly lower MSFE than the random walk forecast while all other models have larger MSFE than the random walk forecast. In terms of MAFE, the two deterministic trend models (*LT* and *QT*) along with their level-break counterparts (*LTB* and *QTB*) have smaller MAFE than the random walk forecast, but the differences are mostly marginal. Furthermore, the Diebold-Mariano statistics suggest that none of these candidate models outperform the random walk model regardless of whether the loss function is quadratic or absolute. The result agrees with the findings of Alquist and Kilian (2010) that the "no change" forecast tends to be the most accurate at one year horizon.[14]According to the directional accuracy test, while most models have success ratios around 50 percent, only the two trend-cycle models achieve significant, marginal at the 10 percent level though, predictive success.

13. The results from forecast accuracy comparisons are virtually the same if we compute MSFEs and MAFEs based on the forecasts of logged prices.

14. It should be noted that Alquist and Kilian (2010), AKV (2012) and Baumeister and Kilian (2012) all forecast the monthly or quarterly prices rather than annual prices.

Table 4: *h*-step-ahead Recursive Forecast Error Diagnostics for the Real Price of Oil

	$h=1$			$h=5$			$h=10$		
	MSFE	MAFE	Success ratio	MSFE	MAFE	Success ratio	MSFE	MAFE	Success ratio
RW	139.842	9.018	n.a.	726.761	20.380	n.a.	1318.937	29.068	n.a.
RWD	1.037	1.017	0.500	1.126	1.054	0.500	1.249	1.125	0.476
	(0.906)	(0.812)	(n.a.)	(0.884)	(0.894)	(n.a.)	(0.925)	(0.967)	(n.a.)
LT	0.996	0.975	0.500	0.710	0.737	0.577	0.623*	0.652*	0.571
	(0.489)	(0.398)	(0.864)	(0.189)	(0.116)	(0.793)	(0.07)	(0.058)	(0.706)
QT	0.947	0.909	0.600	0.504**	0.754	0.692*	0.445**	0.724	0.762***
	(0.318)	(0.178)	(0.701)	(0.039)	(0.102)	(0.079)	(0.021)	(0.109)	(0.010)
LTB	1.089	0.992	0.500	0.736	0.861	0.577	0.520**	0.701**	0.762
	(0.675)	(0.473)	(0.227)	(0.193)	(0.243)	(0.259)	(0.017)	(0.046)	(0.168)
QTB	1.073	0.978	0.533	0.854	1.008	0.577	0.676*	0.902	0.667
	(0.684)	(0.430)	(0.277)	(0.288)	(0.517)	(0.259)	(0.062)	(0.273)	(0.798)
LTCB	1.106	1.053	0.467*	0.524**	0.727**	0.808	0.282***	0.488***	0.810
	(0.845)	(0.797)	(0.070)	(0.011)	(0.016)	(0.800)	(0.002)	(0.004)	(0.435)
QTCB	1.054	1.010	0.533*	0.516**	0.745*	0.846	0.326**	0.610*	0.762***
	(0.702)	(0.549)	(0.072)	(0.028)	(0.078)	(0.852)	(0.034)	(0.100)	(0.010)
Hotelling	1.122	1.062	0.380	1.650	1.231	0.520	2.588	1.534	0.524
	(0.905)	(0.887)	(n.a.)	(0.917)	(0.948)	(n.a.)	(0.927)	(0.953)	(n.a.)
Naïve	1.000	1.000	n.a.	1.284	1.243	0.308	0.966	1.09	0.571
	(0.500)	(0.500)		(0.910)	(0.952)	(0.741)	(0.451)	(0.659)	(1.00)

Notes: Except the benchmark random walk model for which the actual losses are reported, the MSFE and MAFE results for all other models are presented as the ratios to the random walk forecast. The p-values in parenthesis under the MSFE and MAFE are based on pairwise Diebold and Mariano (1995) tests for the null of a random walk model. The success ratio is the proportion of times that the forecast correctly predict the sign of the change. The p-value of the sign test is based on Pesaran and Timmermann (2009), which is not defined when a relevant model predicts the change is always positive such as the case of RWD. *, **, and *** indicate the significance levels of 10%, 5%, and 1%, respectively.

At the five-year horizon, the two trend-cycle models significantly improve the forecast accuracy, in terms of MSFEs and MAFEs, over the random walk model. They are able to reduce MSFEs and MAFEs relative to the random walk forecasts by about 50 percent and 30 percent, respectively. As for other candidate models, we notice that, while the deterministic trend model (with or without the level break) generally produce smaller MSFEs and MAFEs than the random walk model, the Hotelling model and the naive forecasts actually yield larger MSFEs and MAFEs than the random walk model. As far as the success ratios are concerned, while the forecasts from all other models have success ratios below 70%, the two trend-cycle models can successfully predict the signs of oil price changes over 80% of time.

At the ten-year horizon, we again find that the forecasts from the two trend-cycle models perform significantly better than the random walk forecasts with respect to both MSFEs and MAFEs. As a matter of fact, the MSFEs and MAFEs associated with the two trend-cycle models are the smallest among all the models considered here. Compared with the random walk model, they can achieve up to 70 and 50 percent accuracy gains in terms of MSFEs and MAFEs, respectively. It is also worth mentioning that, while the Hotelling model and the naive forecasts still perform poorly in accuracy, all the deterministic trend models yield smaller MSFEs and MAFEs than the random walk model, and in some cases their accuracy gains are statistically significant. In terms of the success ratios, again, the two trend-cycle models can correctly predict the signs of oil price changes around 80% of time.

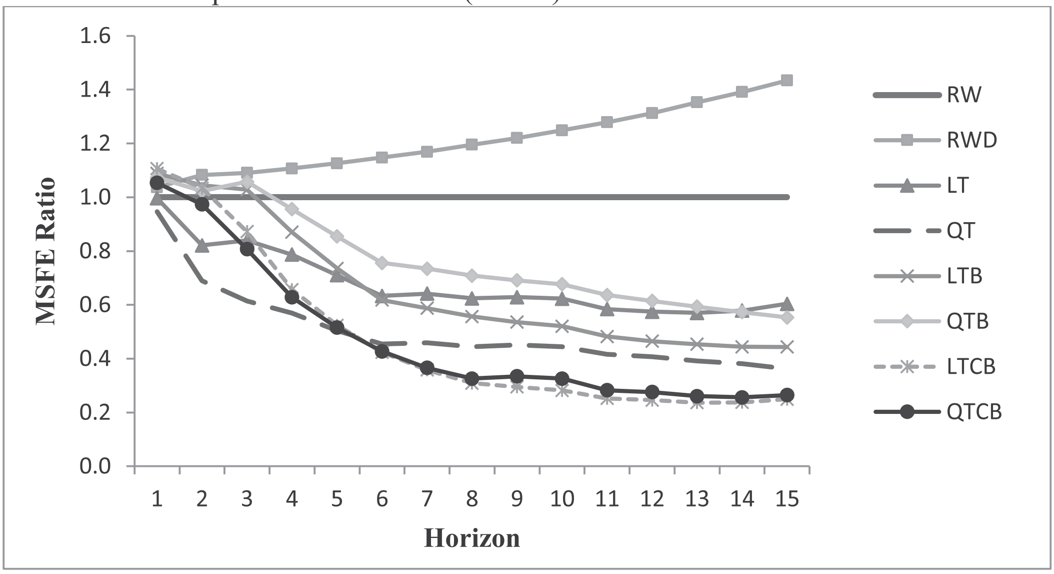
Compared to the forecast accuracy gains at shorter horizons reported in the literature, the MSFE and MAFE reductions of the trend-cycle models at long horizons are considerably large.[15] A closer examination of the time-series properties of real oil prices reveals that the accuracy gains uncovered here may be attributed to several factors. First, cycles matter. As we demonstrated in Section 4, there is clear evidence of cycles in crude oil prices with short cycle lasting over 6 years and long cycle around 29 years, respectively. Clearly, the trend-cycle models, which take into consideration the cyclical behaviour of oil prices, can do a better job in capturing this intrinsic property of real oil prices and are hence able to improve forecast accuracy at long horizons. Second, trend rules in the long run. Note, in Table 4, that all the models with a trend component, including both the trend-cycle models and the deterministic trend models (with and without level breaks), generate lower MSFEs and MAFEs at horizons of five and ten years, and that the MSFE and MAFE ratios become smaller as the forecasting horizon lengthens. These are consistent with the fact that the oil price is a slow mean reversion process and, as shocks dissipate over time, the price eventually reverts to its long-run trend. Since the trend plays such an important role in the long-run evolution of oil prices, forecasting models with the trend incorporated are expected to achieve accuracy gains at long horizons. Lastly, by construction, the low frequency, annual data used in this study is less noisy than the monthly or quarterly data used in previous studies. We suspect this may also contribute to the large accuracy gain here.

To see how the forecasting performance changes with forecasting horizons, in Figure 4 we plot the MSFE and MAFE metrics from various models (M2 through M8), normalized by the random walk model, against the forecasting horizon. At horizons of one and two years, there is no significant difference between competing models. As the forecasting horizon lengthens, all models but *RWD* start to produce lower MSFEs and MAFEs than the random walk model. The two trend-

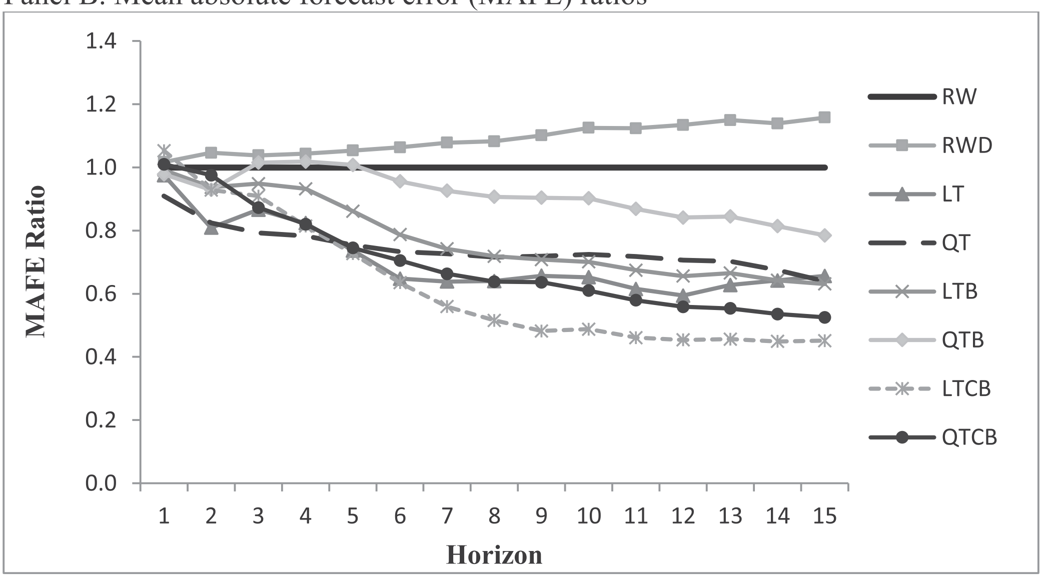
15. For example, AKV(2012) report that some of their VAR models can achieve up to 20 percent accuracy gains at horizons of 1 and 3 months. Baumeister and Kilian (2012) show that the MSFEs from their real-time forecasts can be as much as 25% lower than that of the no-change forecasts for 1- and 3-month ahead.

Figure 4: Forecast performance at different forecast horizons

Panel A. Mean squared forecast error (MSFE) ratios



Panel B. Mean absolute forecast error (MAFE) ratios



Notes: The horizontal axis denotes the forecasting horizons. The vertical axis shows the ratios of the loss functions (MSFE and MAFE) from all other models to the that of the random walk (*RW*) model, which is normalized to one at all horizons. The solid line indicates the RW model. The line with squares indicates the random walk with drift (*RWD*) model. The line with triangles indicates the deterministic linear trend (*LT*) model. The dashed line indicates the determinisitc quadtratic trend (*QT*) model. The line with crosses indicates the determinisitc linear trend with a level break (*LTB*) model. The line with diamonds indicates the determinisitc quadratic trend with a level break (*QTB*) model. The broken line with stars indicates the linear trend-cycle (*LTCB*) model. The line with bubbles indicates the quadratic trend-cycle (*QTCB*) model.

cycle models are found to have better forecasting performance than other models starting from the horizon of five years, with the linear trend-cycle model forecasts slightly better than the forecasts from the quadratic trend-cycle model. Interestingly, the superiority of the trend-cycle forecasts relative to the random walk forecast appears to level off at the ten-year horizon. The pattern is clearly consistent with the slow mean reversion rate reported in the earlier section.

Since the out-of-sample forecasting performance of a model can vary with the forecasting window, we further examine the manner in which the MSFE and MAFE for each model change as the starting year for the out-of-sample period is varied, beginning in 1981. Figures 5 and 6 graph respectively the MSFE and the MAFE for each model as a function of the starting year. For example, the MSFE associated with 1991 means that the MSFE value is calculated using the out-of-sample period from 1991 to 2010. A careful examination of these graphs reveals several interesting points. First, consistent with the result presented in Table 4, the random walk forecast is among the best, if not the best, at the horizon of one year. However, forecasts based on the quadratic trend-cycle model is not much different from the RW forecast and even becomes slightly better than the RW forecast starting from the late 1990s. Second, at longer horizons of five and ten years, the two trend-cycle models are in general the best performers on both MSFE and MAFE metrics, indicating the importance of accounting for the cyclical movements in predicting oil prices in the longer term. Interestingly, while the two trend-cycle models perform almost equally well in forecast accuracy before 1993, the linear trend and cycle model (*LTCB*) is clearly outperformed by the quadratic trend and cycle model (*QTCB*) after 1993. This probably reflects the continued oil price increase since 1998.

5.3 Forecasting the Nominal Price of Oil

In this section, we further examine whether the trend-cycle models have superior predicting power over the nominal price of oil. We re-estimate the eight statistical models (M1 through M8) in Section 4 for logged nominal oil prices over the in-sample period and then form the recursive forecasts for the nominal prices over the evaluation period.[16] As noted in AKV(2012), adjusting the current spot price of oil by expected inflation results in statistically significant improvements in forecast accuracy at the five-year horizon. Therefore, we include that model, which we refer to as the AKV model, in the race. Specifically, the h-horizon forecast of the nominal price at time t in the AKV model is

$$P_{t+h|t} = P_t(1 + \pi_{t,h})^h,$$

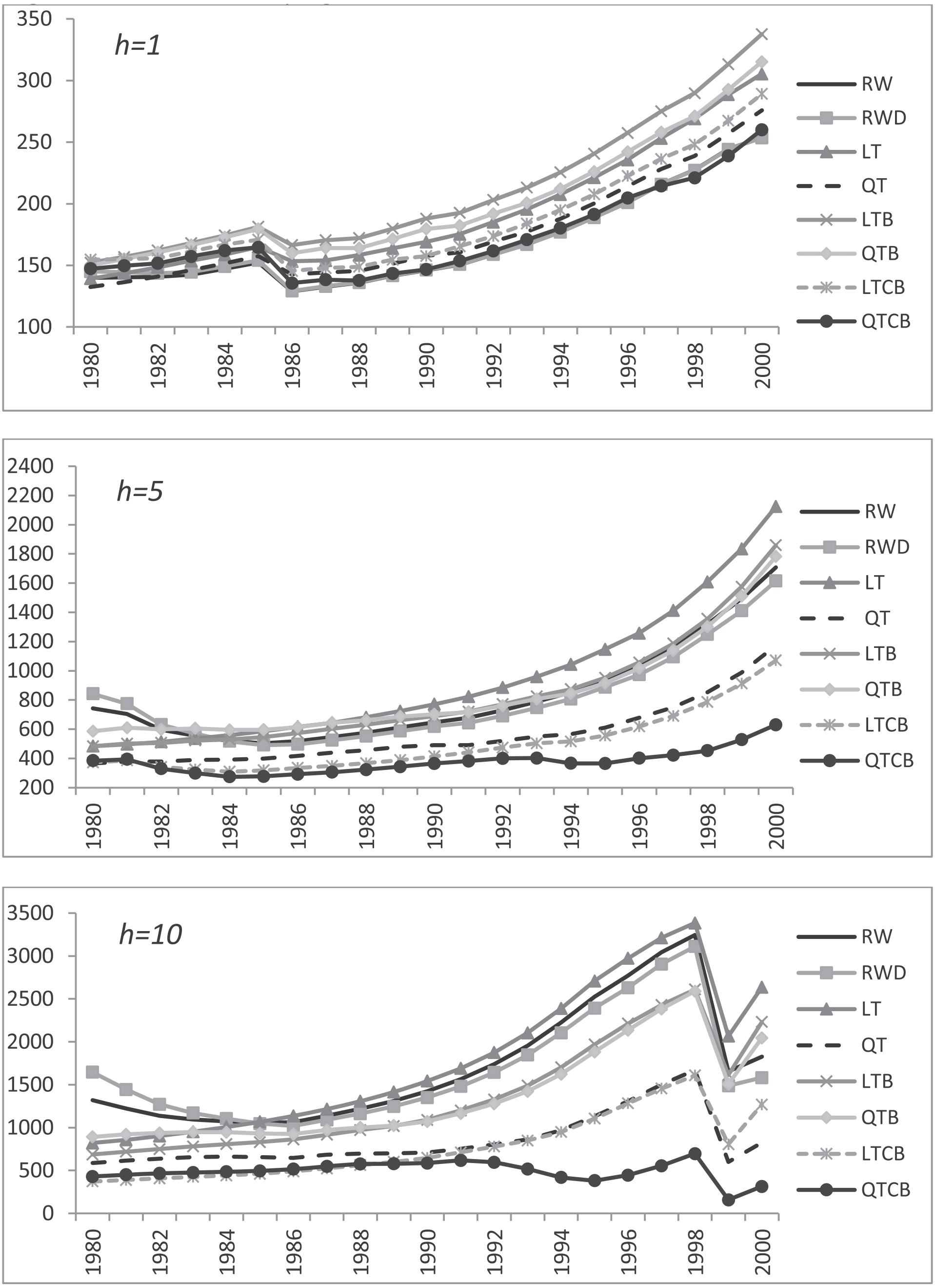
where π is the expected inflation. Following AKV (2012), we use the 1-year and 10-year median CPI inflation forecasts from the Survey of Professional Forecasters (SPF) as measures of expected inflation.[17]Since the 1-year CPI inflation expectation series is not available before 1981, our evaluation period in this section begins in 1982.[18] For completeness, we also include the Hotelling

16. We also tested for structural breaks in the logged nominal price series and found significant evidence for a level break in 1973. To save space, the results for structural break tests and in-sample estimations are not reported but available from the authors upon request.

17. The data was obtained from the Federal Reserve Bank of Philadelphia. The SPF inflation forecast data is quarterly and available for 1-year and 10-year horizons. We use the fourth quarter survey of each year to form our forecasts. The 10-year SPF forecast (properly scaled) is used for both the 5-year and 10-year horizons. Since the 10-year SPF CPI inflation forecasts are not available until 1991Q1, we extend the 10-year CPI inflation expectation series with the10-year inflation forecasts from the Blue Chip Economic Indicators,

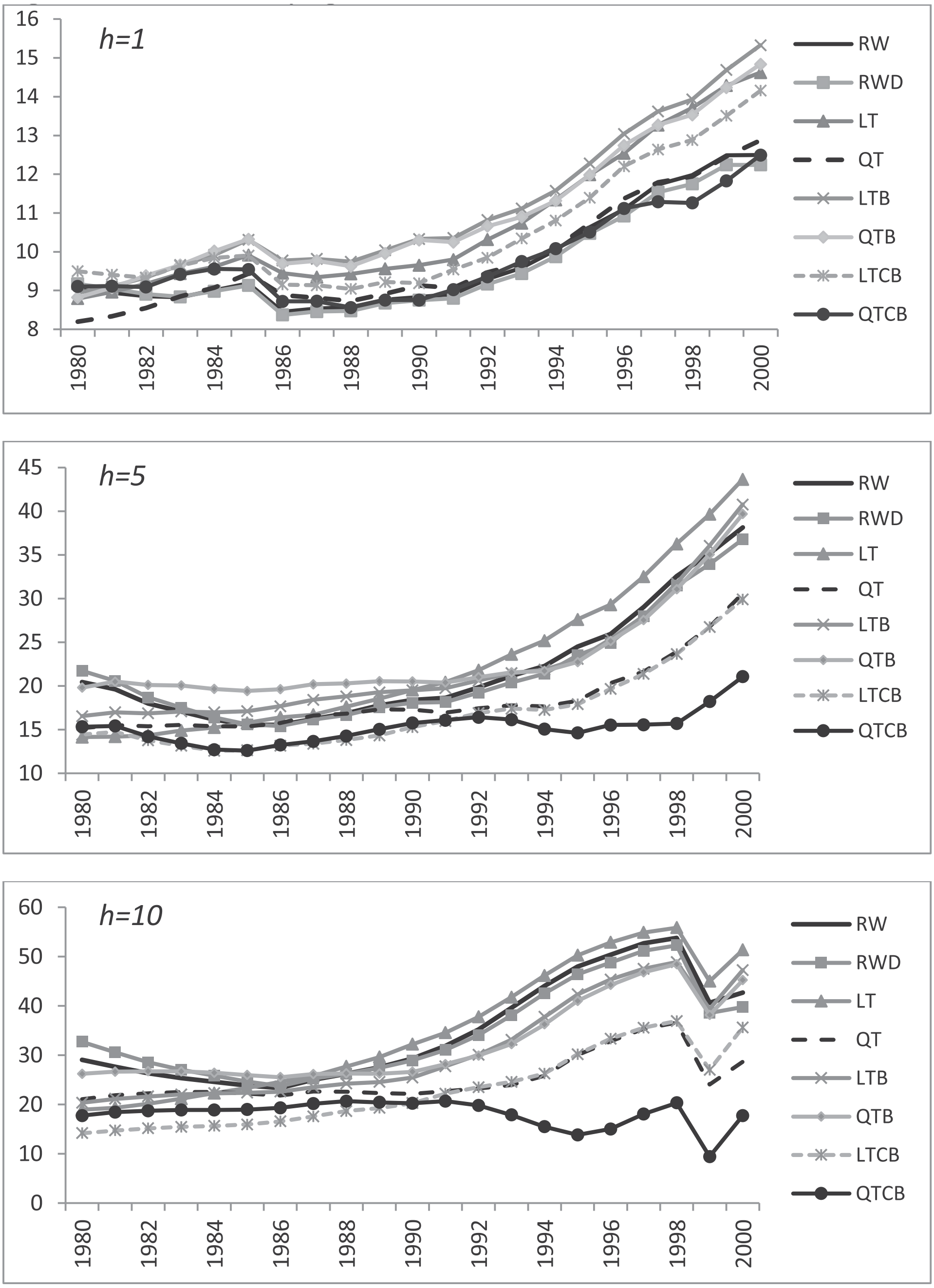
18. We also compared models' forecast accuracy, excluding the AKV model, using the evaluation period 1981–2010 and obtained very similar results.

Figure 5: MSFEs for varying forecast windows



Notes: The horizontal axis indicates the starting year of the forecast evaluation period. The vertical axis shows mean squared forecast error (MSFE) for varying forecasting windows.The solid line indicates the RW model. The line with squares indicates the random walk with drift (*RWD*) model. The line with triangles indicates the deterministic linear trend (*LT*) model. The dashed line indicates the determinisitc quadtratic trend (*QT*) model. The line with crosses indicates the determinisitc linear trend with a level break (*LTB*) model. The line with diamonds indicates the determinisitc quadratic trend with a level break (*QTB*) model. The line with stars indicates the linear trend UC model. The line with bubbles indicates the quadratic trend UC model.

Figure 6: MAFEs for varying forecast windows



Notes: The horizontal axis indicates the starting year of the forecast evaluation period. The vertical axis shows mean absolute forecast error (MAFE) for varying forecasting windows.The solid line indicates the *RW* model. The line with squares indicates the random walk with drift (*RWD*) model. The line with triangles indicates the deterministic linear trend (*LT*) model. The dashed line indicates the determinisitc quadtratic trend (*QT*) model. The line with crosses indicates the determinisitc linear trend with a level break (*LTB*) model. The line with diamonds indicates the determinisitc quadratic trend with a level break (*QTB*) model. The line with stars indicates the linear trend–cycle (*LTCB*) model. The line with bubbles indicates the quadratic trend-cycle (*QTCB*) model.

model and the naïve forecasts, both of which are similarly defined as in the case of real price forecasts except that the Hotelling forecast is now based on nominal interest rates.[19]

Table 5 assesses the forecast accuracy of various forecasting models against the benchmark of a random walk forecast for horizons of 1, 5 and 10 years.[20] There are several interesting features worth noting. First, the AKV model yields no better forecasts than the random walk forecasts for the nominal oil prices at all horizons. As a matter of fact, the MSFEs and MAFEs of the AKV model are larger than those of the RW model, which is in sharp contrast with the AKV (2012) finding that it can achieve about 15 percent MSFE reduction relative to the no-change forecast at the five-year horizon. The result appears to be driven by the difference in evaluation periods. If we restrict our evaluation period to 1998–2009, roughly the same evaluation period as in AKV(2012), we also find that the five-year-ahead MSFEs from the AKV model is exactly 15 percent lower than that of the no-change forecast.[21] As the oil price has been generally increasing during this period, it is not surprising that the AKV model predicts the nominal oil price better than the no-change forecast. However, once we include the late 1980s and 1990s during which the oil price was generally declining, the conclusion changes.

Second, the ranking of the UC-based models (M3–M8) is largely consistent with that of real price forecasts. While at the horizon of one year, no models can beat the random walk model, the two trend-cycle models produce significantly smaller MSFEs and MAFEs than the random walk forecasts at the horizons of five and ten years. Furthermore, models with a quadratic trend always outperform their linear counterparts (e.g. *QT vs LT, QTB vs LTB*, and *QTCB vs. LTCB*) at all horizons. Between the two trend-cycle models, the *QTCB*, instead of the *LTCB*, becomes the champion of all models for the nominal price and the magnitude of the accuracy gains relative to the random walk forecast is considerable. The MSFE (MAFE) reductions of the *QTCB* model are as much as 68 (35) percent and 82 (57) percent at the five and ten-year horizons, respectively. To understand why this model performs so well, we plot the trend and cycle decomposition of the logged nominal price over the full sample period in the right panels of Figure 3. Similar to the logged real price (*LRP*), the logged nominal price (*LNP*) also exhibits clear evidence of cyclicity. A striking difference between the two top panels is that the upward trend after the level-break of 1973 in the nominal price is much more pronounced than in the real price. Indeed, although not reported, the estimates of the slope parameter from the *QTCB* model of the logged nominal price are 0.044 for the in-sample period of 1861–1980 and 0.060 for the full sample, both of which are more than 10 times larger than the statistically insignificant slope parameter estimates reported in Table 3 for the real price. When the trend is apparent, a model capturing the increasing trend would naturally result in improvements in forecasting accuracy compared to the no-change forecast.

Lastly, while the *RWD* and naïve forecasts sometimes produce lower MSFEs and MAFEs than the random walk forecasts, the reductions are mostly marginal and not statistically significant. The Hotelling predictions of nominal prices are always worse than random walk forecasts. These results are generally consistent with those reported in Table 4 for real prices of oil.

To summarize, we show that, while the random walk forecasts dominate all others at short horizons, the trend-cycle models can do better jobs in predicting oil prices over longer horizons.

19. While not reported for the sake of brevity, we also formed the nominal price forecasts by adjusting the real price forecasts with the CPI inflation expectations. In doing so, we obtained very similar results.

20. We also evaluated the predictive accuracy of various forecasting models against the benchmark of the AKV model and obtained very similar results.

21. Although not reported, we also compared the forecasting performance of UC-based models to that of the AKV model during this period and find that the QT, QTB, LTCB and QTCB all have smaller MSPEs than the AKV model.

Table 5: *h*-step-ahead Recursive Forecast Error Diagnostics for the Nominal Price of Oil

	$h=1$			$h=5$			$h=10$		
	MSFE	MAFE	Success ratio	MSFE	MAFE	Success ratio	MSFE	MAFE	Success ratio
RW	111.527	7.055	n.a.	496.303	15.139	n.a.	976.234	21.938	n.a
AKV	1.011	1.018	0.517	1.033	1.071	0.640	1.281	1.194	0.650
	(0.490)	(0.668)	(n.a.)	(0.524)	(0.745)	(n.a.)	(0.696)	(0.850)	(n.a.)
RWD	1.004	0.995	0.517	0.900	0.981	0.640	0.940	1.043	0.650
	(0.526)	(0.445)	(n.a.)	(0.211)	(0.410)	(n.a.)	(0.339)	(0.641)	(n.a.)
LT	1.176	1.059	0.586	0.967	0.929	0.480	1.007	1.000	0.400
	(0.491)	(0.809)	(0.466)	(0.632)	(0.240)	(0.596)	(0.969)	(0.501)	(0.786)
QT	1.151	0.945	0.586	0.422*	0.678*	0.840**	0.371*	0.702	0.750
	(0.712)	(0.317)	(0.304)	(0.052)	(0.093)	(0.020)	(0.089)	(0.203)	(1.000)
LTB	1.204	1.083	0.517	1.171	0.981	0.440	1.043	0.969	0.450*
	(0.864)	(0.896)	(0.390)	(0.845)	(0.443)	(0.120)	(0.788)	(0.363)	(0.079)
QTB	1.050	0.984	0.586	0.620*	0.869	0.800	0.480*	0.836	0.750
	(0.662)	(0.414)	(0.132)	(0.063)	(0.224)	(1.00)	(0.084)	(0.266)	(1.000)
LTCB	1.140	1.057	0.483	0.589**	0.789**	0.760	0.431**	0.579***	0.850
	(0.821)	(0.763)	(0.226)	(0.042)	(0.037)	(0.168)	(0.023)	(0.002)	(0.428)
QTCB	1.156	0.958	0.552	0.320*	0.657*	0.680*	0.176**	0.431**	0.900***
	(0.716)	(0.376)	(0.201)	(0.069)	(0.081)	(0.061)	(0.048)	(0.034)	(0.009)
Hotelling	1.023	1.055	0.517	1.332	1.270	0.640	2.316	1.660	0.650
	(0.629)	(0.768)	(n.a.)	(0.790)	(0.870)	(n.a.)	(0.872)	(0.883)	(n.a.)
Naïve	1.000	1.000	n.a.	1.283	1.136	0.423	0.986	0.960	0.762
	(0.500)	(0.500)		(0.943)	(0.943)	(0.862)	(0.415)	(0.325)	(0.145)

Notes: Except the benchmark random walk model for which the actual losses are reported, the MSFE and MAFE results for all other models are presented as the ratios to the random walk forecast. The p-values in parenthesis under the MSFE and MAFE are based on pairwise Diebold and Mariano (1995) tests for the null of a random walk model. The success ratio is the proportion of times that the forecast correctly predict the sign of the change. The p-value of the sign test is based on Pesaran and Timmermann (2009), which is not defined when a relevant model predicts the change is always positive such as the case of RWD. *, **, and *** indicate the significance levels of 10%, 5%, and 1%, respectively.

Particularly, the superiority of the trend-cycle models relative to the random walk forecast increases in forecasting horizons. These results are robust to the choice of forecasting windows and also consistent for both nominal and real prices.

6. DISCUSSION AND CONCLUSION

In this paper, we employ an unobserved components model to disentangle the long-term trend from cyclical movements in international benchmark crude oil prices using annual data from 1861 to 2010. The estimation of the model identifies a slowly evolving U-shaped long-term trend interrupted by a level break in 1973 and two types of cycles, with the short cycle having a period of six years and the long cycle having an average length of 29 years. There are two main contributions contained in this paper. First, we complement the existing literature by documenting the patterns of the cycles and demonstrating the importance of incorporating cycles in understanding the time path of oil prices. With regard to the long-run trend, we use longer and more recent data to show that models with a deterministic quadratic trend provide a better fit to the data than their linear counterparts, corroborating to some of the earlier findings. Compared to the amplitudes of the cycles, however, the growth rate of the long-run trend is small. The second and more important contribution lies in the out-of-sample forecast comparisons. We show that, while the random walk forecasts tend to be the most accurate at shorter horizons among the models considered, they are outperformed by models that combine a long-run trend and cycles at horizons of five to ten years. The trend-cycle models can not only produce more accurate forecasts but also have better capability to predict the signs of oil price changes. The result is robust to the choice of forecasting windows, the loss functions and whether the oil price is specified in nominal or real terms. Given that the models considered in this paper are reduced-form in nature, it is possible that models incorporating market structure and economic fundamentals are even better equipped to capture the trajectory of oil prices. Nonetheless, our results provide clear evidence of predictability in oil prices at longer horizons and complement to the existing literature.

There are a number of areas for future research. First, although this paper has provided empirical evidence on the existence of cycles in crude oil prices, we have not explored the theoretical implications. A structural model explaining the causes and consequences of the long cycles could provide further insight into the temporal behaviour of commodity prices. Second, given the estimated length of the short cycles, it could be interesting to investigate the lead-lag relation between the cycle of oil prices and the usual business cycles.

APPENDIX: THE OIL PRICE DATA

The oil price in our analysis consists of the annual averages of oil prices in the US for the period 1861–1944, Arabian Light crude oil for 1945–1983, and Brent crude oil for 1984–2010. The switch of the type of crude oils reflects the evolution of the oil market. The modern oil industry generally dates back to 1859 when the first commercial oil well was drilled by Edwin Drake in Titusville, Pennsylvania.[22] From its inception until the end of the World War II, the US had the largest oil industry and since the crude oil was freely traded, it is reasonable to use the US average price.[23] Moreover, since the "As Is Agreement" signed in 1928, the price for internationally traded

22. See Dahl (2003) pp. 146–152 for a brief history of the oil market.
23. Until the end of World War II, the US supplied roughly two-thirds of crude oil in the world (Yergin, 1991).

oil is quoted on the basis of oil price in the Gulf of Mexico, known as the Gulf Basing Point System (Stevens, 2010). In this sense, the US domestic price also represents the price for the world oil market.

After World War II, the "centre of the gravity" of the oil industry shifted from the US to the Middle East when Saudi Arabia became a major exporter (Yergin, 1991). With growing international oil trade between the Middle East and the West, the Arab Light became the benchmark crude oil until the mid-1980s. With the development of crude oil futures trading in London and New York in the 1980s, the Brent– a crude oil produced in the North Sea – and the West Texas Intermediate crude oil have become two widely used benchmark oil in international trade. Thanks to its accessibility as a seaborne crude, geographic proximity to the market, excellent liquidity and stability, Brent is used as a pricing benchmark in over two thirds of the crude oil trade in the world (ICE, 2012). Of particular note, the Brent and Arabia Light are of similar quality and traded at similar prices.

ACKNOWLEDGMENTS

We thank the two anonymous referees, the journal editor, and seminar participants at University of Dundee and the SIRE Energy and Environmental Economics workshop for helpful comments and discussions. This research project is partially sponsored by Shanghai Pudong Program. All errors remain our own.

REFERENCES

Alhajji, A F; and David Huettner (2000), "OPEC and World Crude Oil Markets from 1973 to 1994: Cartel, Oligopoly or Competitive?" *The Energy Journal* 21(3): 31–60. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol21-No3-2.

Almoguera, Pedro A.; Douglas, Christopher C.; Herrera, Ana Maria (2011). "Testing for the Cartel in OPEC: Non-cooperative Collusion or Just Non-cooperative?"*Oxford Review of Economic Policy*, Spring 27(1): 144–68.

Alquist, Ron and Lutz 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.

Alquist, Ron; Lutz Kilian and Robert Vigfusson (2012). "Forecasting the price of Oil", forthcoming in G. Elliott and A. Timmermann (eds), *Handbook of Economic Forecasting*, 2 Amsterdam: North-Holland.

Ashley, Richard (2003). "Statistically Significant Forecasting Improvements: How Much Out-of-Sample Data is Likely Necessary?" *International Journal of Forecasting* 19: 229–39. http://dx.doi.org/10.1016/S0169-2070(01)00139-X.

Bai, Jushan and Pierre Perron (1998). "Testing for and Estimation of Multiple Structural Changes," *Econometrica* 66: 47–79. http://dx.doi.org/10.2307/2998540.

Baumeister, C. and L. Kilian (2012) "Real-Time Forecasts of the Real Price of Oil," *Journal of Business and Economic Statistics* 30(2): 326–336. http://dx.doi.org/10.1080/07350015.2011.648859.

Berk, Peter and Michael Roberts (1996). "Natural resource prices: will they ever turn up?"*Journal of Environmental Economics and Management* 31: 65–78. http://dx.doi.org/10.1006/jeem.1996.0032.

Bleaney, M. F and D. Greenaway (1993). "Long-Run Trends in the Relative Price of Primary Commodities and in the Terms of Trade of Developing Countries" *Oxford Economic Papers, New Series* 45(3): 349–363.

BP (various years) "Statistical Review of the World Energy," published annually at www.bp.com/statisticalreview.

Cashin, Paul.and C. J. McDermott (2002). "The long-run behaviour of commodity prices: small trends and big variability," IMF Staff Papers, 49(2): 175–199.

Dahl, Carol A. (2004). *International Energy Markets: Understanding Pricing, Policies, and Profit*, Tulsa, OK: Pennwell Press, March.

Diebold, Francis and Roberto Mariano (1995), "Comparing predictive accuracy," *Journal of Business and Economic Statistics* 13(3): 253–263.

Durbin, J. and S.J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford.

Dvir, E. and K.S. Rogoff (2010). "The Three Epochs of Oil" NBER working paper 14927.

Energy Information Administration (EIA, 2012) (http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n = PET&s = RCRR01NUS_1&f = A).

Erten, Bilgeand José Antonio Ocampo (2012) "Super-cycles of commodity prices since the mid-nineteenth century" *UN-DESA working paper No. 110*.

Fisher, A. C. (1979). "Measures of resource scarcity," in *"Scarcity and Growth Reconsidered,"* (Smith ed.), Resources for the Future, Johns Hopkins Press, Baltimore.

Geman, Helyette (2005). "Energy commodity prices: Is mean-reversion dead?" *Journal of Alternative Investments*, Fall 2005, pp. 31–45. http://dx.doi.org/10.3905/jai.2005.591576.

Hamilton, James (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, James (2011). "Historical Oil Shocks" *NBER Working Paper No. 16790*

Harvey, Andrew C. (1989). *Forecasting. Structural Time Series Models and the Kalman Filter*, Cambridge University Press: Cambridge.

Harvey, Andrew C. and Thomas M. Trimbur (2003). "General model-based filters for extracting cycles and trends in economic time series" *Review of Economics and Statistics* 85(2): 244–255. http://dx.doi.org/10.1162/003465303765299774.

Hotelling, H. (1931). "The economics of exhaustible resources," *Journal of Political Economy* 39: 137–175. http://dx.doi.org/10.1086/254195.

Inter Continental Exchange (the ICE), (2012). "ICE Brent Crude Oil: FAQ" https://www.theice.com/publicdocs/futures/ICE_Brent_FAQ.pdf

Koopman, S.J., A. C. Harvey, J. A. Doornik and N. Shephard (2009). *Structural Time Series Analyser, Modeller and Predictor*, Timberlake Consultants: London.

Krautkraemer, Jeffrey A. (1998). "Nonrenewable Resource Scarcity," *Journal of Economic Literature* 36 (4): 2065–2107

Lee, Junsoo and Mark C. Strazicich (2001). "Break point estimation and spurious rejections with endogenous unit root tests," *Oxford Bulletin of Economics and Statistics* 63: 535–558. http://dx.doi.org/10.1111/1468-0084.00234.

Lee, Junsoo and Mark C. Strazicich (2003). "Minimum LM unit root test with two structural breaks," *Review of Economics and Statistics* 85(4): 1082–1089. http://dx.doi.org/10.1162/003465303772815961.

Lee, Junsoo; John List and Mark 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.

Livernois, John. (2009). "On the empirical significance of the Hotelling Rule," *Review of Environmental Economics and Policy* 3(1): 22–41. http://dx.doi.org/10.1093/reep/ren017.

Nordhaus, William D. (1973). "The Allocation of Energy Resources," *Brookings Papers on Economic Activity* 1973(3): 529–576. http://dx.doi.org/10.2307/2534202.

Nyblom, Jukka and Harvey, Andrew (2001). "Testing against smooth stochastic trends" *Journal of Applied Econometrics* 16: 415–429. http://dx.doi.org/10.1002/jae.604.

Pesaran, M Hashem and Allan Timmermann (2009). "Testing dependence among serially correlated multicategory variables," *Journal of the American Statistical Association* 104(485): 325–337.

Pindyck, R.S. (1978). "The optimal exploration and production of non-renewable resources" *Journal of Political Economy* 86: 841–861. http://dx.doi.org/10.1086/260714.

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.

Slade, M.E. (1982). "Trends in natural resource commodity prices: an analysis of the time domain." *Journal of Environmental Economics and Management* 9: 122–137. http://dx.doi.org/10.1016/0095-0696(82)90017-1.

Slade, M.E. and Henry Thille (2009). "Wither Hotelling: Tests of the Theory of Exhaustible Resources," *Annual Review of Resource Economics* 1: 239–260. http://dx.doi.org/10.1146/annurev.resource.050708.144223.

Smith, James (2009). "World Oil: Market or Mayhem?" *Journal of Economic Perspectives* 23(3): 145–164. http://dx.doi.org/10.1257/jep.23.3.145.

Stevens, Paul (2010) "The history of oil," *POLINARES working paper No.3*.

Yergin, Daniel (1991). "*The Prize: The Epic Quest for Oil, Money, and Power.*" New York: Simon & Schuster.

Zellou, Abdel and John T. Cuddington (2012). "Is There Evidence of Super Cycles in Crude Oil Prices?" *SPE Economics and Management* (forthcoming).

Zivot, Eric and Donald W.K. Andrews (1992). "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis," *Journal of Business & Economic Statistics* 10(3): 251–70.