purposes. Nowadays, there are two price benchmarks for crude oil, the Brent and the WTI (West Texas Intermediate). In this paper, we will focus on the WTI, which is also known as Texas light sweet (relatively low-density and low sulphur content). WTI is mainly traded in North America, and its futures contracts are mainly traded on the New York Mercantile Exchange (linked to the CME group). On a typical trading day about 1 million contracts are traded at the CME, the majority being futures and 10% of contracts being options. All futures on WTI traded on the CME have physical settlement, which means that these futures are tied to actual delivery when they mature, and the place of settlement is Cushing, Oklahoma.

The pricing of futures and options on commodities in a modern theoretical framework has been considered at least since Black (1976). Ramaswamy and Sundaresan (1985) extended Black's ideas to value American style call options on futures contracts. Further, Shastri and Tandon (1986) suggested a way to price both American type call and put options on the Standard and Poor 500 index and the West German Mark futures contracts. Schwartz (1997) and Hilliard and Reis (1998) shifted the attention to the incorporation of a stochastic convenience yield into the modeling. The convenience yield takes account of certain costs and benefits that apply to holding the asset, and as Schwartz (1997) demonstrated, allowing this model parameter to follow an Ornstein–Uhlenbeck process provides the model with enough flexibility to produce a reasonable large class of shapes for the forward-curves. The Schwartz (1997) model is typically calibrated to futures prices by using the Kalman-Filter approach, taking account of the fact that the convenience yield and possibly the spot-price of the commodity are not directly observable.[1,2] Hilliard and Reis (1998) presented a formula that prices a European call option on a commodity in the Schwartz (1997) framework.

Our paper expands on Schwartz (1997) and Hilliard and Reis (1998) as such as that we estimate the implied spot price and convenience yield of the underlying commodity from options rather than futures. More specifically, we use the extended Kalman-Filter and prices of European call options on WTI crude oil futures to estimate the Schwartz (1997) model. The motivation for this lies in the fact that option prices carry far more information on the volatility structure of the underlying asset than futures do. The implied volatility smile as well as the existence and strong use among practitioners of local and stochastic volatility models provide plenty of evidence for this statement.

The remainder of the paper is organized as follows. The mathematical framework will be presented in the second section, while the econometric methodology used to estimate the model will follow in the third section. In the fourth section, the data and general assumptions for this paper will be described in detail. Empirical results and conclusions will be discussed in the last section.

2 Mathematical model

Let us recall from the Schwartz (1997) two-factor model that the dynamics of the commodity

With positive correlation $\rho = dZ_1 dZ_2$, this model generates a mean reversion effect in spot and convenience yield. This is typically present in commodity prices and explained in Schwartz (1997). For pricing of commodity contracts, the dynamics under the pricing measure is relevant. Following Schwartz (1997), we have

$$dS/S = (r - \delta)dt + \sigma_1 d\hat{Z}_1$$
$$d\delta = (\kappa(\alpha - \delta) - \lambda)dt \quad .$$

then, it is not hard to obtain the following:

$$dF/F = \sigma_F d\hat{Z}_F,$$

where[5]

$$\sigma_F^2(\sigma_1, \sigma_2, \rho, \kappa, \tau) = \sigma_1^2 + \sigma_2^2 H(\tau)^2 - 2\rho\sigma_1\sigma_2 H(\tau).$$

This representation enabled Hilliard and Reis (1998) to provide an explicit expression for the price of a European call option in Schwartz' (1997) model:

$$C(t, T_1, T) = P(t, T_1)[F(t, T)N(d_1) - KN(d_2)],$$

where $d_1 = \frac{\ln(F(t,T)/K) + 0.5\upsilon^2}{\upsilon}$, $d_2 = d_1 - \upsilon$. Here t is the current time; T_1 is the time of maturity of the call option on the futures contract; T is the time of maturity of the underlying futures contract, with $T_1 \leq T$; $F(t, T)$ is the futures price at time t for maturity at time T; $C(t, T_1, T)$ is the price of a call with maturity T_1 on the underlying futures contract with the maturity T at time t; $P(t, T_1)$ is the price of a zero-coupon bond with maturity T_1 and K is the strike price of the European call option. Further υ^2 can be computed as:[6]

$$\upsilon^2 = \sigma_1^2(T_1 - t) - \frac{2\sigma_1\sigma_2\rho}{\kappa}\left[(T_1 - t) - \frac{e^{-\kappa(T-T_1)} - e^{-\kappa(T-t)}}{\kappa}\right]$$
$$+ \frac{\sigma_2^2}{\kappa}\left[(T_1 - t) - \frac{2}{\kappa}\left(e^{-\kappa(T-T_1)} - e^{-\kappa(T-t)}\right) + \frac{1}{2\kappa}\left(e^{-2\kappa(T-T_1)} - e^{-2\kappa(T-t)}\right)\right].$$

For our empirical analysis it is convenient to express Hilliard and Reis' (1998) option pricing formula in the notation of Schwartz, as we will consider simultaneously options with different strikes and maturities at the same time. This reflects on choosing current time $t = 0$ at each instance and T respectively T_1 becoming the time to maturity. By doing this, we eliminate t from the notation and obtain the following expression for the price of the European call option above

$$C(S, \delta, T_1, T) = P(t, T_1)[F(S, \delta, T)N(d_1) - KN(d_2)],$$

where $d_1 = \frac{\ln(F(S,\delta,T)/K) + 0.5\upsilon^2}{\upsilon}$ and $d_2 = d_1 - \upsilon$. Following in the spirit of Schwartz (1997), we will in the following assume that option prices are directly observable, while the state variables S and δ are hidden.[7] To get hold of these hidden state variables we will employ filtering theory.

3 The extended Kalman Filter algorithm

The classical Kalman Filter algorithm requires a linear setup and is hence not applicable to our options based framework. Instead, we will employ the extended Kalman Filter. The principles of the Kalman Filter algorithm and the extended Kalman Filter algorithm are the same, except that the latter involves linearization of some key functionals and approximations.

---

[5] Compare Schwartz (1998) and Hilliard and Reis (1998).

[6] Solving the integrals: $\upsilon^2(t, T_1, T) = \int_t^{T_1} \sigma_2^2 dw + \int_t^{T_1}[\frac{\sigma_2}{\kappa}(1 - e^{-\kappa(T-w)})]^2 dw - \int_t^{T_1} \frac{2\sigma_1\sigma_2\rho}{\kappa}(1 - e^{-\kappa(T-w)})dw$, υ can be calculated.

[7] Alternative assumptions on observability and its effects on calibration are discussed in Ewald et al. (2017).

Before discussing the extended Kalman Filter, in order to introduce some notations, the basic principles behind filtering technology are explained in the following: Using Bayes theory, filters can use the information about current observations to predict the values of unobservable variables at the current time point, and then update the information and forecast the situation at the next time point (compare Pasricha 2006).

To be more specific, the process of filtering can be described as follows: In a state space model, we consider two parts, the state variable x_k for $k = 1, 2, \ldots, K$ and the observations z_k for $k = 1, 2, \ldots, K$, where K is the count of the observations of the time variable. Both are possibly multi-dimensional. Normally, progression is determined by the state dynamics, $x_k = f_k(x_{k-1}, v_k)$, where x_k and x_{k-1} are the respective states at time points k and $k-1$ and v_k is the noise effecting the dynamics. Then x_k is following a first-order Markov process and $x_k \mid x_{k-1} \sim p_{x_k \mid x_{k-1}}(x_k \mid x_{k-1})$. The relationship between state variable(s) and the observations can be described as $z_k = h_k(x_k, n_k)$, where x_k is the state variable(s) and n_k is the measurement noise at time k.

We denote with $z_{1:k}$ the history of observations from the start of the time series until time k. We assume that the observations are conditionally independently given x_k. For $k \geq 1$ the expression $p(x_k \mid x_{k-1})$ presents the state transition probability. Essentially, the recursive filter consist of two steps: The first step is referred to as the prediction step, while the second step is called the update step. Schematically this comes down to

$$p(x_{k-1} \mid z_{1:k-1}) \rightarrow p(x_k \mid z_{1:k-1}) \quad \text{prediction step}$$
$$p(x_k \mid z_{1:k-1}), z_k \rightarrow p(x_k \mid z_{1:k}) \quad \text{update step.}$$

As for the prediction step, assume the probability density function $p(x_{k-1} \mid z_{1:k-1})$ is available at time point $k-1$, using the Chapman–Kolmogorov equation, the prior probability of the state at time k can be expressed as:

$$p(x_k \mid z_{1:k-1}) = \int p(x_k \mid x_{k-1}) p(x_{k-1} \mid z_{1:k-1}) dx_{k-1}.$$

For the update step, the posterior probability density function is

$$p(x_k \mid z_{1:k}) = \frac{p(z_k \mid x_k) p(x_k \mid z_{1:k-1})}{p(z_k \mid z_{1:k-1})},$$

where

$$p(z_k \mid z_{1:k-1}) = \int p(z_k \mid x_k) p(x_k \mid z_{1:k-1}) dx_k.$$

The classical Kalman filter requires the state transition function as well as the measurement function to be linear and the noise terms to be normal distributed. Only in this case, prediction and update steps reflect conditional expectations and variances only. In the case of the extended Kalman filter, a local linearization of the aforementioned equations is carried out and the relevant conditional densities $p(x_{k-1} \mid z_{1:k-1})$, $p(x_k \mid z_{1:k-1})$ and $p(x_k \mid z_{1:k})$ are approximated by normal densities as follows:

$$p(x_{k-1} \mid z_{1:k-1}) \approx N(x_{k-1}; m_{k-1|k-1}, P_{k-1|k-1}),$$
$$p(x_k \mid z_{1:k-1}) \approx N(x_k; m_{k|k-1}, P_{k|k-1}),$$

where $m_{k-1|k-1}$ respectively $m_{k|k-1}$ and $P_{k-1|k-1}$ respectively $P_{k|k-1}$ reflect expectations and variances of the relevant conditional distributions in approximation. We have

$$p(x_k \mid \underset{\sim}{\mathbf{z}}_{k}) \approx N(x_k; m_{k|k}, P_{k|k})$$
$$m_{k|k-1} = f_k(m_{k-1|k-1})$$
$$P_{k|k-1} = Q_{k-1} + \widehat{F}_k P_{k-1|k-1} \widehat{F}_k^T$$
$$m_{k|k} = m_{k|k-1} + K_k(\underset{\sim}{\mathbf{z}}_{k} - h_k(m_{k|k-1})),$$

as well as

$$P_{k|k} = P_{k|k-1} - K_k \widehat{H}_k P_{k|k-1} = (I - K_k \widehat{H}_k) P_{k|k-1},$$

where

$$K_k = P_{k|k-1} \widehat{H}_k^T \left(\widehat{H}_k P_{k|k-1} \widehat{H}_k^T + R_k \right)^{-1}$$

is known as gain, and

$$\widehat{F}_k = \frac{df_k(x)}{}$$

Table 1 The estimated parameters

Parameters	κ	α	λ	σ_1	σ_2	ρ
K = 97.5	0.8584	0.0661	0.0738	0.2593	0.4791	0.7642
K = 100	0.8135	0.0557	0.0537	0.2624	0.4818	0.7582
K = 102.5	0.7863	0.0821	0.0621	0.2713	0.5041	0.7687
Futures	0.7031	0.0259	0.0497	0.5592	0.1645	0.5189

2015, respectively. Prices for the corresponding futures are collected at the same time. The risk free rate has been obtained from treasury notes[8] maturing in the same months as the options, i.e. April 2014, May 2014, Aug 2014, Nov 2014, and March 2015, respectively.[9]

5 Empirical results and conclusion

Estimated parameters from the various European calls are shown in Table 1. Overall, the

Fig. 1 Estimated spot price from calls and futures contracts. (Color figure online)
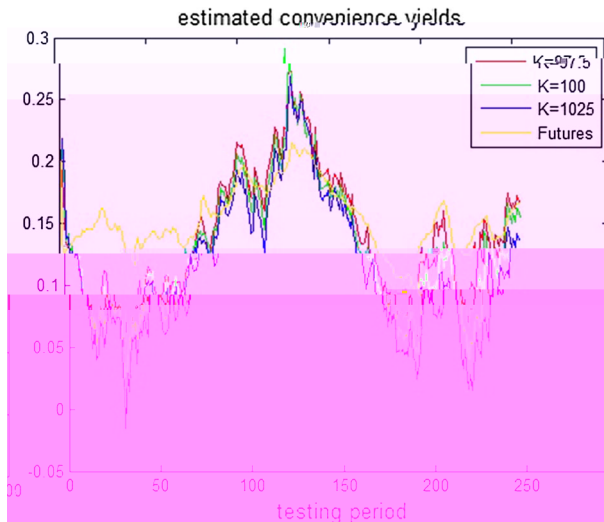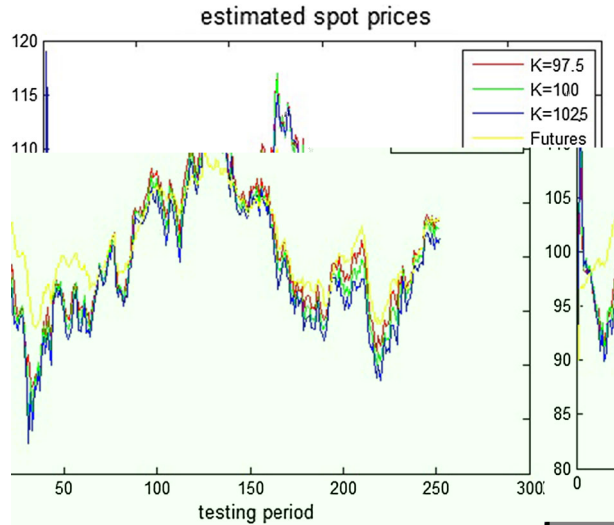
Fig. 2 Estimated convenience yield from calls and futures contracts

classical Kalman Filter. The convenience yields corresponding to the three different strikes are all relatively closely aligned, fluctuating between 0 and about 0.3. Convenience yields corresponding to lower strikes appear to be slightly elevated. Prominently, at the beginning of the sample period the estimated convenience yields based on strikes $K = 100$ and $K = 102.5$ drop to about 0, while this peg does not appear in the call with strike $K = 97.5$. As for the estimated spot, the convenience yield obtained from the associated futures prices differs markedly, showing less extreme movements and less fluctuation.

Figures 3, 4, and 5 show the term structures of the European calls, i.e. prices for increasing maturities, at different dates during the sample period. In all three pictures, the horizontal axis represents the number of months to maturity and the vertical axis represents the prices
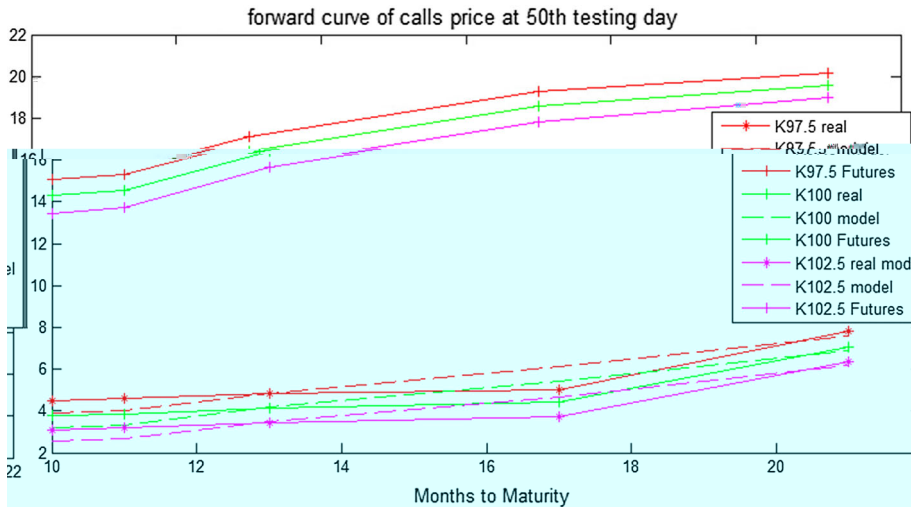
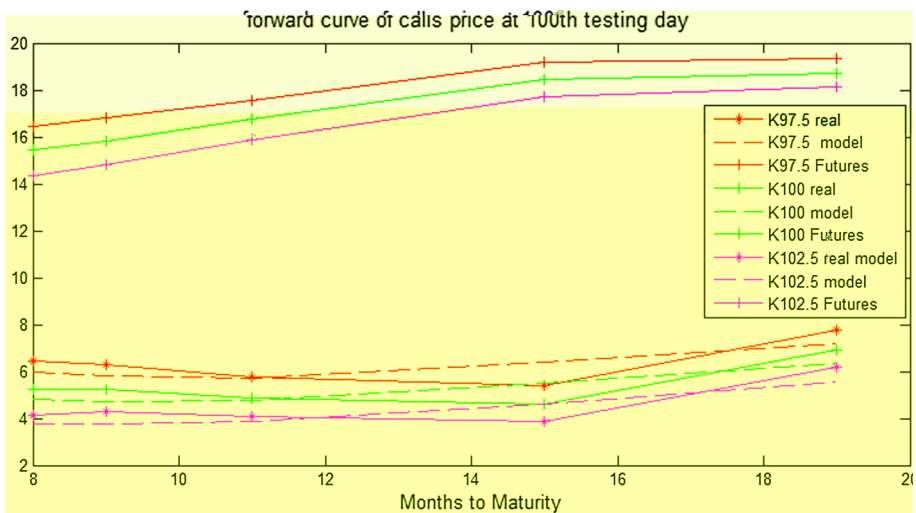**Fig. 3** The term structures of the European calls at 50th testing day

Fig. 4 The term structures of the European calls at 100th testing day

of the calls. As shown in the figures, the dotted lines represent the model-estimated prices corresponding to different calls based on different strike prices, where model calibration has been undertaken via the extended Kalman filter and option data as described earlier. The solid lines represent the observed prices corresponding to the different calls based on different strike prices. It is not hard to observe that the model-estimated prices are close to the observed prices of the calls in all three cases, whatever the strike prices are, the fit is excellent. However, in addition to these, the lines with plus signs lying far above the observed option prices represent the model prices obtained from the Schwartz (1997) model, calibrated to observed futures prices via the classical Kalman filter. As can be seen very clearly, this calibration performs poorly, providing clear evidence, that if the main objective is to price
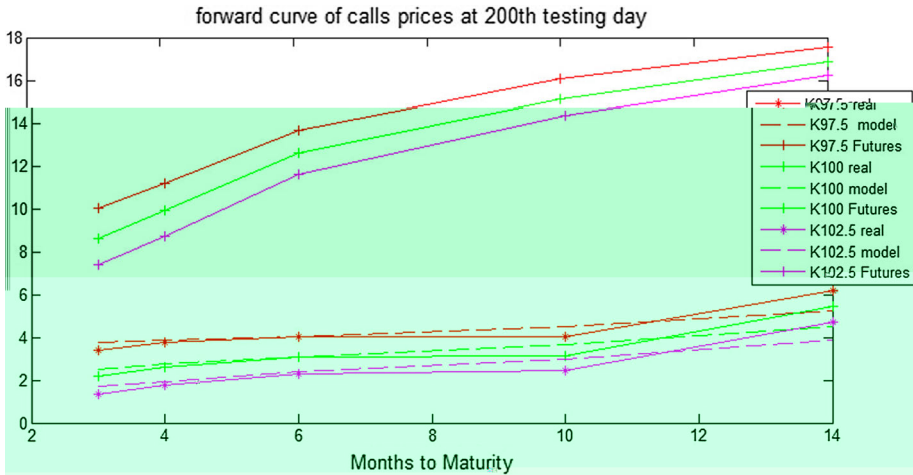
Fig. 5 The term structures of the European calls at 200th testing day
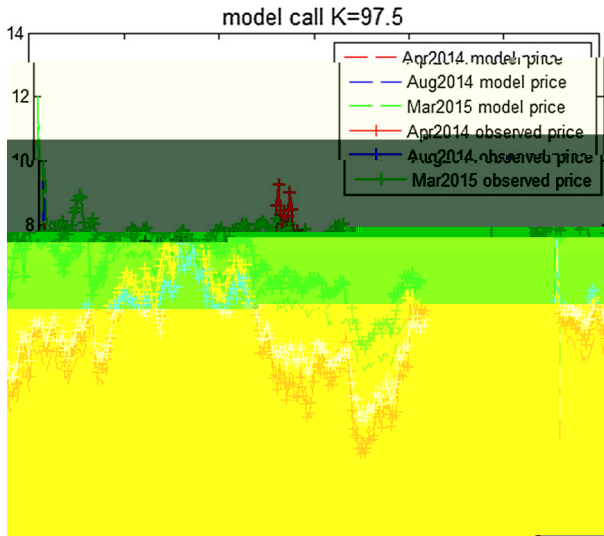

Fig. 6 The observed and model prices of European calls with the strike price of K = 97.5

options in the Schwartz ([1997]) model, then the approach presented in our paper should be followed, rather than the standard Schwartz ([1997]) approach.

Figures [6], [7] and [8] show model and observed prices over the sample period for nine calls maturing in April 2014, August 2014 and March 2015 with strike prices $K = 97.5$, $K = 100$ and $K = 102.5$. As can be seen, fit is excellent, except towards the end of the sample period for the option maturing in March 2015.

Figure [9] shows that all four parameter sets, i.e. those from three options with strikes $K = 97.5$, $K = 100$, and $K = 102.5$ as well as from the underlying futures contracts, produce a very good fit for pricing futures contracts.
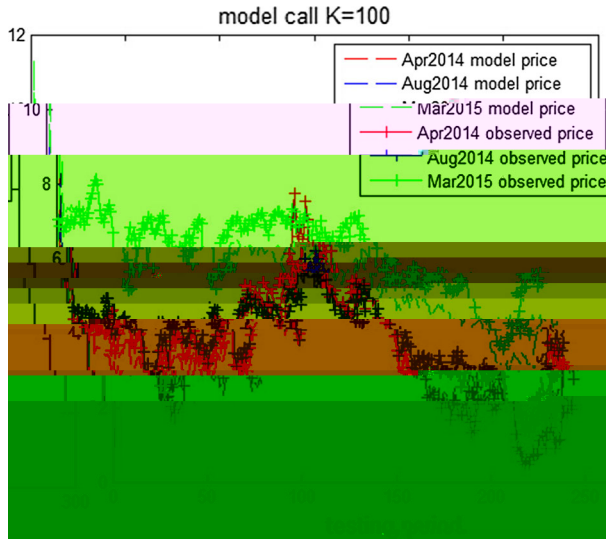
Fig. 7 The observed and model prices of European calls with the strike price of K = 100
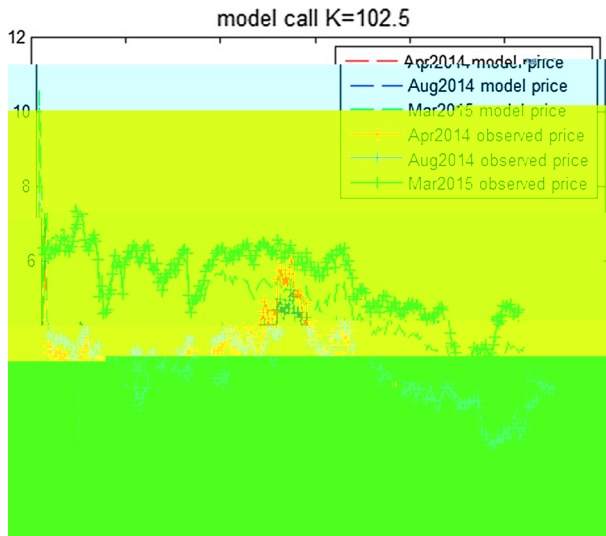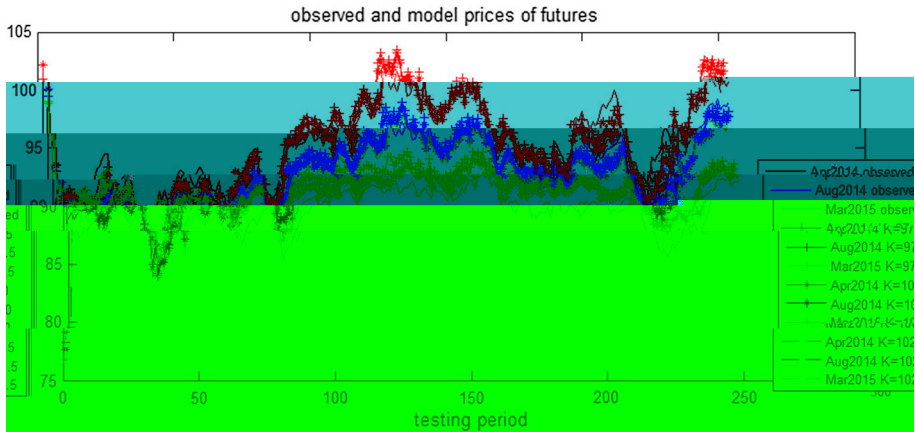

Fig. 8 The observed and model prices of European calls with the strike price of K = 102.5

6 Conclusions

Using the extended Kalman filter, we have estimated the Schwartz (1997) two-factor model by means of historical prices for options with differing maturities and strikes. We found that some parameter estimates significantly differ from those obtained via the classical Schwartz (1997) approach, which involves using the classical Kalman filter and futures prices. We then demonstrated, that while the parameter sets obtained from the options are also able to provide a good fit when pricing futures, the opposite does not hold. Using the parameters obtained via the classical Schwartz (1997) approach to price options produces a significantly

Fig. 9 The observed futures prices and the model-estimated futures prices based on the different strike prices of European style calls

poorer fit than our approach. Hence, when the objective is to price both futures and options simultaneously within the Schwartz (1997) framework, our recommendation is to fit the model to options rather than futures by using the extended Kalman filter.

References

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics*, *3*, 167–179.
Chen, J., & Ewald, C.-O. (2017). Pricing commodity futures options in the Schwartz multi factor model with stochastic volatility: An asymptotic method. *International Review of Financial Analysis*, *52*, 144–151. https://doi.org/10.1016/j.irfa.2017.05.002.
Ewald, C. O., Zhang, A., & Zong, Z. (2017). *On the consistency of two benchmark approaches to calibrate the Schwartz (1997) two-factor model for commodity futures*, working paper.
Harvey, A. C. (1989). *Forecasting structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.
Hilliard, J. E., & Reis, J. (1998). Valuation of commodity futures and options under stochastic convenience yields, interest rates and jump diffusion in the spot. *The Journal of Financial and Quantitative Analysis*, *33*(1), 61–86.
Pasricha, G. K. (2006). *Kalman Filter and its economic applications*. University of California, Santa Cruz, CA 95064. http://mpra.ub.uni-muenchen.de/22734/.
Ramaswamy, K., & Sundaresan, S. M. (1985). The valuation of option on futures contracts. *Journal of Finance*, *40*(5), 1319–1340.
Schwartz, E. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, *52*(3), 923–973.
Schwartz, E. (1998). Valuing long-term commodity assets. *Financial Management*, *27*(1), 57–66.
Shastri, K., & Tandon, K. (1986). An empirical test of a valuation model for American options on futures contracts. *Journal of Financial and Quantitative Analysis*, *21*(4), 377–392.