Two-Factor Models for Forecasting of Futures Prices

Jun S. Han[1], Nino Kordzakhia[1], Pavel V. Shevchenko[2], and Stefan Trück[2]

[1] School of Mathematical and Physical Sciences, Macquarie University
jun.han1@hdr.mq.edu.au
[2] Department of Actuarial Studies and Business Analytics, Macquarie University

Abstract. We analyse the h-step forecasts of futures prices, using the extended two-factor model from [7] that is based on the Schwartz-Smith model from [10]. In the extended model, the short and long-term factors represent correlated Ornstein-Uhlenbeck processes and will be jointly estimated, along with model parameters, using the Kalman filter through the maximum likelihood method. Furthermore, we assume that the error terms in the measurement equation system are interdependent and serially correlated. A comparative analysis of the reduced-form and full models from [7] and [10], respectively, has been carried out using the simulated futures prices.

1 Introduction

Forecasting financial markets presents an irresistible challenge that continues to captivate the attention of investors, analysts and researchers. Forecasting algorithms serve as powerful tools in investment decision-making, such as resource allocation, risk management, and strategic planning.

The two-factor model introduced in [10] is often used for pricing in various commodity markets. A multi-factor model framework was overviewed in [4] with the factors following correlated Ornstein-Uhlenbeck processes. This framework has been extended to incorporate seasonal components, as presented in [11] and [8]. Furthermore, the paper [3] extended the model to encompass long-dated commodity prices with correlated stochastic volatilities and interest rates, and [1] extended the model to incorporate observable regression factors.

As for calibrating to real data is concerned, in the context of multi-factor models, it is typically assumed that measurement errors are independent. However, in [6], the authors argued for the necessity of modelling the dependence of measurement errors. Additionally, the study in [7] further extended to incorporate serial correlations of marginal measurement errors in the Kalman filter and then estimated through the parameter estimation procedure. One-step-ahead forecasting under the reduced model setup has been widely studied for comparative analyses, using futures prices from different markets. Examples include [9], [4] and, [5].

In this paper, we employ the two-factor model framework from [7] to study its relative forecasting performance. Through a comparison with an existing reduced-form model in [10], we observe that the two-factor model in [7] demonstrates notable out-of-sample forecast accuracy.

This paper is organised as follows. In Section 2, we provide the description of the two-factor model. In Section 3, we discuss forecasting results using the simulated futures prices. Lastly, in Section 4, we provide a concluding overview of the results presented in this paper.

2 Models

Consider the risk-neutral probability space denoted by $(\Omega, \mathcal{F}, \mathbb{Q})$. The risk-neutral dynamics of the spot price S_t is defined as $\ln S_t = \chi_t + \xi_t$, where χ_t and ξ_t are unobservable variables that represent short-term and long-term factors, respectively, with the following stochastic differential equations:

$$\begin{cases} d\chi_t = (-\kappa\chi_t - \lambda_\chi)dt + \sigma_\chi dW_t^\chi, \\ d\xi_t = (\mu_\xi - \lambda_\xi - \gamma\xi_t)dt + \sigma_\xi dW_t^\xi, \\ \mathbb{E}^{\mathbb{Q}}[dW_t^\chi dW_t^\xi] = \rho_{\chi\xi}dt, \end{cases} \tag{1}$$

where, for χ_t and ξ_t, respectively, $\kappa, \gamma \in \mathbb{R}^+$ are the rates of the mean-reversion, λ_χ, $\lambda_\xi \in \mathbb{R}$ are deterministic unknown risk premia, σ_χ, $\sigma_\xi \in \mathbb{R}^+$ are volatility parameters. $\mu_\xi \in \mathbb{R}$ is the mean of the long-term factor, and $\rho_{\chi\xi}$ is the correlation coefficient of the two correlated standard Brownian processes W_t^χ and W_t^ξ under $\mathbb{Q}$.

The risk-neutral futures price at time t, maturing in T years, is calculated as $F_{t,T} = \mathbb{E}^{\mathbb{Q}}[S_T|\mathcal{F}_t]$. Then, in discrete time, the evolution of futures prices is given by the following linear state space model:

$$\mathbf{x}_t = \mathbf{c} + \mathbf{G}\mathbf{x}_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}), \tag{2}$$

$$\mathbf{y}_t = \mathbf{d}_t + \mathbf{B}_t\mathbf{x}_t + \mathbf{v}_t, \tag{3}$$

$$\mathbf{v}_t = \Phi_p(\boldsymbol{L})\mathbf{v}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_\varepsilon). \tag{4}$$

where, assuming N futures contracts with maturities $T_1 < T_2 < \ldots < T_N$,

$$\begin{aligned} \mathbf{x}_t &= \begin{pmatrix} \chi_t \\ \xi_t \end{pmatrix}, \ \mathbf{c} = \begin{pmatrix} 0 \\ \frac{\mu_\xi(1-e^{-\gamma\Delta t})}{\gamma} \end{pmatrix}, \ \mathbf{G} = \begin{pmatrix} e^{-\kappa\Delta t} & 0 \\ 0 & e^{-\gamma\Delta t} \end{pmatrix}, \\ \mathbf{y}_t &= \left(\ln F_{t,T_1}, \ln F_{t,T_2}, \ldots, \ln F_{t,T_N}\right)', \\ \mathbf{d}_t &= \left(A(T_1-t), A(T_2-t), \ldots, A(T_N-t)\right)', \\ \mathbf{B}_t &= \begin{pmatrix} e^{-\kappa(T_1-t)} & \ldots & e^{-\kappa(T_N-t)} \\ e^{-\gamma(T_1-t)} & \ldots & e^{-\gamma(T_N-t)} \end{pmatrix}', \end{aligned}$$

with $\Phi_p(\boldsymbol{L}) = \sum_{r=1}^{p} \boldsymbol{\phi}_r \boldsymbol{L}^r$ being a function of the lag operator, Δt being the time difference in years between discretised time instances $t-1$ and t, and the function $A(T-t)$ is defined as

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

The covariance matrices of error terms $\mathbf{w}_t$ and $\boldsymbol{\varepsilon}_t$ are defined as

$$\mathbf{W} = \begin{pmatrix} \frac{1-e^{-2\kappa\Delta t}}{2\kappa}\sigma_\chi^2 & \frac{1-e^{-(\kappa+\gamma)\Delta t}}{\kappa+\gamma}\sigma_\chi\sigma_\xi\rho_{\chi\xi} \\ \frac{1-e^{-(\kappa+\gamma)\Delta t}}{\kappa+\gamma}\sigma_\chi\sigma_\xi\rho_{\chi\xi} & \frac{1-e^{-2\gamma\Delta t}}{2\gamma}\sigma_\xi^2 \end{pmatrix}, \mathbf{V}_\varepsilon = \begin{pmatrix} s_{11}^2 & s_{12} & \dots & s_{1N} \\ s_{12} & s_{22}^2 & \dots & s_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ s_{1N} & s_{2N} & \dots & s_{NN}^2 \end{pmatrix}.$$

Model 1 assumes $\mathbf{v}_t$ to be serially uncorrelated, with $\mathbf{V}_\varepsilon$ being a diagonal matrix that consists of variances of measurement errors. **Model 2** considers the full model as described above. Now, applying the lag polynomial $\boldsymbol{I} - \Phi_p(\boldsymbol{L})$ to (3), where $\boldsymbol{I}$ is the identity matrix, the measurement equation is expressed as

$$\mathbf{y}_t = \mathbf{d}_t^*(p) + \mathbf{B}_t^*(p)\mathbf{x}_t + \mathbf{v}_t^*(p), \tag{5}$$

$$\mathbf{d}_t^*(p) = \sum_{r=1}^{p}\boldsymbol{\phi}_r\mathbf{y}_{t-r} + \mathbf{d}_t - \sum_{r=1}^{p}\boldsymbol{\phi}_r\mathbf{d}_{t-r} + \sum_{r=1}^{p}\boldsymbol{\phi}_r\mathbf{B}_{t-r}\left[\sum_{k=1}^{r}(\mathbf{G}^{-1})^k\right]\mathbf{c}, \tag{6}$$

$$\mathbf{B}_t^*(p) = \mathbf{B}_t - \sum_{r=1}^{p}\boldsymbol{\phi}_r\mathbf{B}_{t-r}(\mathbf{G}^{-1})^r, \tag{7}$$

$$\mathbf{v}_t^*(p) = \sum_{r=1}^{p}\boldsymbol{\phi}_r\mathbf{B}_{t-r}\left[\sum_{k=1}^{r}(\mathbf{G}^{-1})^{r-k+1}\mathbf{w}_{t-k+1}\right] + \boldsymbol{\varepsilon}_t. \tag{8}$$

For model parameter estimations, we adopt the Kalman filtering technique with the maximum likelihood estimation method. The optimal Kalman gain matrix is

$$\begin{aligned}
\mathbf{K}_t^* &= (\mathbf{P}_{t|t-1}[\mathbf{B}_t^*(p)]' + \mathbf{C}_t^*(p))(\mathbf{L}_{t|t-1}^* + \mathbf{B}_t^*(p)\mathbf{C}_t^*(p) + [\mathbf{C}_t^*(p)]'[\mathbf{B}_t^*(p)]')^{-1}, \\
\mathbf{P}_{t|t-1} &= Var[\mathbf{x}_t|\mathfrak{I}_{t-1}], \\
\mathbf{C}_t^*(p) &= Cov(\mathbf{x}_t, \mathbf{v}_t^*(p)|\mathfrak{I}_{t-1}) = \mathbf{W}\left(\sum_{r=1}^{p}\boldsymbol{\phi}_r\mathbf{B}_{t-r}(\mathbf{G}^{-1})^r\right)' \\
\mathbf{L}_{t|t-1}^* &= Var[\mathbf{e}_{t|t-1}^*] \\
&= \mathbf{B}_t^*(p)\mathbf{P}_{t|t-1}\mathbf{B}_t^{*\prime}(p) + \mathbf{V}_{t|t-1}^*(p) + \mathbf{B}_t^*(p)\mathbf{C}_t^*(p) + [\mathbf{C}_t^*(p)]'[\mathbf{B}_t^*(p)]', \\
\mathbf{V}_{t|t-1}^*(p) &= Var[\mathbf{v}_t^*(p)|\mathfrak{I}_{t-1}] \\
&= \left[\sum_{r=1}^{p}\boldsymbol{\phi}_r\mathbf{B}_{t-r}(\mathbf{G}^{-1})^r\right]\mathbf{W}\left[\sum_{r=1}^{p}\boldsymbol{\phi}_r\mathbf{B}_{t-r}(\mathbf{G}^{-1})^r\right]' + \mathbf{V}_\varepsilon.
\end{aligned}$$

Now, assuming the prediction error $\mathbf{e}^*_{t|t-1} \sim \mathcal{N}(\mathbf{0}_N, \mathbf{L}^*_{t|t-1})$, we construct the marginalised likelihood based on the prediction errors

$$\ell_t(\boldsymbol{\psi}, \mathbf{V}_\varepsilon, \boldsymbol{\phi}; \mathbf{y}_t) = -\frac{1}{2}\ln(\det(\mathbf{L}^*_{t|t-1})) - \frac{1}{2}(\mathbf{e}^*_{t|t-1})'(\mathbf{L}^*_{t|t-1})^{-1}\mathbf{e}^*_{t|t-1},$$

where $\boldsymbol{\psi} = (\kappa, \sigma_\chi, \lambda_\chi, \gamma, \mu_\xi, \sigma_\xi, \lambda_\xi, \rho_{\chi\xi})$ is the set of model parameters and $\det(\cdot)$ denotes the determinant of a matrix. The sum of the marginalised likelihoods for $t > p$ is the likelihood function to be maximised for the estimation of the model parameters. That is,

$$\ell(\boldsymbol{\psi}, \mathbf{V}_\varepsilon, \boldsymbol{\phi}; \mathbf{y}_t) = -\frac{1}{2}\sum_{t=p+1}^{n}\left[\ln(\det(\mathbf{L}^*_{t|t-1})) + (\mathbf{e}^*_{t|t-1})'(\mathbf{L}^*_{t|t-1})^{-1}\mathbf{e}^*_{t|t-1}\right], \quad (9)$$

and (9) is to be maximised subject to the constraint $\kappa \geq \gamma$, as in [2].

3 Simulation Study: Forecasting of Futures Prices

In this section, we assess the forecasting ability of the two models using parameters in [7] to simulate the logarithm of futures prices, with sample sizes $n = 500, 1000$ for $N = 10$ contracts. This takes into account the serial correlation and inter-correlation between prices of different futures contracts. To examine the out-of-sample predictive ability, we consider the root mean square error (RMSE) and the continuous ranked probability score (CRPS). The RMSE is useful for evaluating point estimates of forecasts, while the CRPS, approximated using the pinball loss (PL), is used for the evaluation of the full predictive distribution based on quantiles. Equations (10) and (11) are used to compute the scores.

$$\mathrm{RMSE} = \sqrt{\frac{1}{h}\left(\sum_{t=1}^{h}(y_t - \hat{y}_t)^2\right)}, \quad (10)$$

$$\mathrm{PL}(\hat{Q}_t(p), y_t, p) = \begin{cases} p(y_t - \hat{Q}_t(p)) & y_t \geq \hat{Q}_t(p) \\ (1-p)(\hat{Q}_t(p) - y_t) & y_t < \hat{Q}_t(p) \end{cases},$$

$$\mathrm{CRPS} = 2\int_0^1 \mathrm{PL}(\hat{Q}_t(p), y_t, p)dp, \quad (11)$$

where y_t and $\hat{y}_t$ represents the actual and price forecasts, respectively, and $\hat{Q}_t(p)$ represents quantile forecasts with probability p.

For our forecasting approach, we employ the simulation of a large number of future paths for the logarithm of prices. The average at each time point is taken within the forecasting horizon, while we also use quantiles to construct the confidence interval. Our process begins by fitting the model to the training data to obtain estimated parameters, as well as estimated state variables and

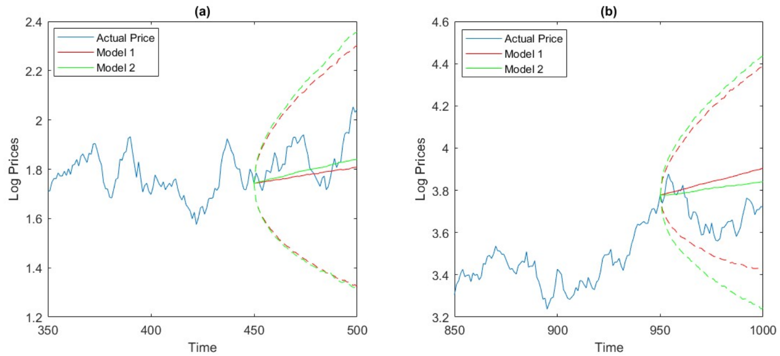

Fig. 1: Logarithms of futures prices along with their forecasts using Models 1 and 2 for the first contract with (a) $n = 500$, and (b) $n = 1,000$. The bold line shows the mean, and the dotted line represents the 95% confidence interval.

Table 1: RMSE and CRPS for $n = 500, 1,000$ and $N = 10$

	$n = 500$				$n = 1,000$			
	RMSE		CRPS		RMSE		CRPS	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
C_1	0.1056	0.0929	0.0618	0.0589	0.1815	0.1486	0.1020	0.0851
C_2	0.0993	0.0893	0.0595	0.0580	0.1947	0.1809	0.1115	0.1015
C_3	0.1097	0.0965	0.0638	0.0603	0.1637	0.1615	0.0920	0.0910
C_4	0.1008	0.0906	0.0598	0.0580	0.1938	0.1981	0.1115	0.1109
C_5	0.1094	0.0974	0.0643	0.0610	0.1743	0.1826	0.0987	0.1022
C_6	0.1093	0.0953	0.0648	0.0602	0.1811	0.1900	0.1027	0.1065
C_7	0.1077	0.0940	0.0632	0.0588	0.1828	0.1880	0.1040	0.1053
C_8	0.1031	0.0881	0.0609	0.0567	0.1775	0.1761	0.1000	0.0981
C_9	0.1114	0.0944	0.0647	0.0586	0.1811	0.1696	0.1023	0.0948
C_{10}	0.1060	0.0926	0.0619	0.0582	0.1697	0.1474	0.0958	0.0838

fitted residuals. Subsequently, we proceed to simulate sample paths for the state variables $\mathbf{x}_t$ using (2) (and $\mathbf{v}_t$ using (8) for Model 2). These simulations enable us to obtain futures paths of $\mathbf{y}_t$ using (3).

Figure 1 visually illustrates the performance of forecasts for the futures prices for each model, while Table 1 shows the summary of RMSE and CRPS metrics. Model 2 achieves lower RMSE than Model 1 in most contracts, having the path of the expected forecast to be much closer to the actual path. The result was consistent when CRPS was compared, indicating that the full model showed better out-of-sample forecasts expressed as probability distributions. Both models showed a widening of the confidence intervals, which correctly reflects the increasing uncertainty associated with a longer forecast horizon.

4 Conclusion

In this study, we conducted a comparative analysis of the out-of-sample forecast accuracy of the two-factor modelling approach widely used for futures price data calibration. We explored both its reduced form as presented in [10] and the full model as detailed in [7]. For examination of the relative performance of the two models, we have used the simulated futures price data with different sample sizes. In summary, our analysis provided valuable insights into the accuracy and forecasting capabilities of the two models in the context of the term structure of futures prices. The full two-factor model has demonstrated a higher level of accuracy for almost all contracts for the two different data sets.

A `MATLAB` code, developed by the first author, can be accessed on GitHub[3], along with a simulated sample dataset.

References

1. Ames, M., Bagnarosa, G., Matsui, T., Peters, G.W., Shevchenko, P.V.: Which risk factors drive oil futures price curves? Energy Economics **87**, 104676 (2020)
2. Binkowski, K., He, P., Kordzakhia, N., Shevchenko, P.: On the parameter estimation in the Schwartz-Smith's two-factor model. In: *Research School on Statistics and Data Science*. Springer, pp. 226-237 (2019)
3. Cheng, B., Nikitopoulos, C.S., Schlögl, E.: Pricing of long-dated commodity derivatives: Do stochastic interest rates matter? Journal of Banking & finance **95**, 148–166 (2018)
4. Cortazar, G., Millard, C., Ortega, H., Schwartz, E.S.: Commodity price forecasts, futures prices, and pricing models. Management Science **65**(9), 4141–4155 (2019)
5. Costa, A.B.R., Ferreira, P.C.G., Gaglianone, W.P., Guillén, O.T.C., Issler, J.V., Lin, Y.: Machine learning and oil price point and density forecasting. Energy Economics **102**, 105494 (2021)
6. Han, J.S., Kordzakhia, N., Shevchenko, P.V., Trück, S.: On correlated measurement errors in the schwartz–smith two-factor model. Dependence Modeling **10**(1), 108–122 (2022)
7. Han, J.S., Kordzakhia, N., Shevchenko, P.V., Trück, S.: On autoregressive measurement errors in a two-factor model. 2021-22 MATRIX Annals (2023)
8. Peters, G.W., Briers, M., Shevchenko, P., Doucet, A.: Calibration and filtering for multi factor commodity models with seasonality: incorporating panel data from futures contracts. Methodology and Computing in Applied Probability **15**, 841–874 (2013)
9. Prokopczuk, M.: Pricing and hedging in the freight futures market. Journal of Futures Markets **31**(5), 440–464 (2011)
10. Schwartz, E., Smith, J.E.: Short-term variations and long-term dynamics in commodity prices. *Management Science* **46**(7), 893–911 (2000)
11. Sørensen, C.: Modeling seasonality in agricultural commodity futures. Journal of Futures Markets: Futures, Options, and Other Derivative Products **22**(5), 393–426 (2002)

[3] `https://github.com/Junee1992/EUA-Futures-Pricing/tree/main/MAF2024`