Journal of Business & Economics
Vol.5 No.1 (Jan-June 2013) pp.1-23

RANGE-BASED PRICE FORECASTS AND A TRADING STRATEGY FOR CORN AND SOYBEANS FUTURES

Brian S.F. Chan[*]
Alan T.K. Wan[**]

Abstract

The high volatility of food prices over the past decade has made price forecasting increasingly important to policy makers and market participants alike. Food price forecasts are undertaken on a regular basis by various government agencies, and there is appreciable evidence that these forecasts have implications for government food policies. It is noted that most existing studies on food price forecasts are based on periodic averages or close-to-close price data. On the other hand, considerable literature has accumulated over the past few years regarding the use of range-based forecasting methods. One such method is based on the observation that movements in the daily high and low prices are tied up in the long run by a condition closely approximated by the daily price range. This paper applies range-based method to forecasting the daily high and low prices of corn and soybeans futures. It is found that this approach offers significant advantages over the traditional ARIMA and random walk methods in terms of out-of-sample forecast accuracy. Another attraction of this method is that it is very easy to implement. While there are many avenues in which the high and low price forecasts can be put to use, as one application this study develops a

[*] Brian S.F. Chan, Department of Management Sciences, College of Business, City University of Hong Kong, Hong Kong.
[**] Alan T.K.Wan, Department of Management Sciences, College of Business, City University of Hong Kong, Hong Kong. Corresponding author email address: Alan.Wan@cityu.edu.hk

Acknowledgements: An earlier version of this paper was presented at the World Business Institute Conference in Business, Economics and Finance, Madrid, September 2013. We thank the referee and participants at this conference for their helpful comments and suggestions.

trading strategy of corn and soybeans futures that makes use of these price forecasts. This strategy generally yields very reasonable profits, and its success depends in part on the accuracy of the price forecasts produced by the underlying model.

Key Words: Annualised returns, food, forecast accuracy, high, low, Vector error correction.

1. Introduction

Agri-food trade is an important economic activity worldwide. It also plays a major role in generating employment for many countries. In a deregulated food market, information about food commodity prices is critical to market participants, in part to help them manage price risk. This information is also important to policy makers to help them make better decisions. Since the beginning of this century, prices of nearly all food commodities have experienced a major boom, accompanied by higher price volatility that has lasted longer than before. The 2008 global food price crisis has triggered widespread concern over the volatility experienced in food prices, and many experts are of the view that the world has entered a new regime characterized by unstable food prices in the international market. This uncertainty accentuates the need to forecast prices accurately for better decision making. Food price forecasts are undertaken regularly by various government agencies such as the Food and Agriculture Organization (FAO) of the United Nations, and the Economic Research Service (ERS) of the U.S. Department of Agriculture (USDA). These forecasts have become increasingly important to food policy makers as they provide an important signal about the changing structure of food and agricultural economies. Identifiable users of the ERS food price forecasts include the USDA's Chief Economist and Secretary's Office, the Federal Reserve Board, and the U.S. Congress. The U.S. President's annual budget for designing food and agricultural programs, including the Food Stamp Program, also makes use of the food price forecasts produced by the ERS. A succinct summary and a critical appraisal of ERS' forecasting methodology can be found in Joutz, Trost, Hallahan, Clauson and Denbaly (2000).

Existing empirical analysis of agriculture and food typically defaults to average price data over a time period or end-of-period price data. This is the case with all the papers cited in Allen's (1994) comprehensive survey of agricultural price forecasts. The same is also evident in the more recent empirical studies. For instance, the food demand model for Tunisia developed by Dhehibi and Gil (2003) is based on annual consumer price data obtained by monthly averages; the price interdependency study of energy and agricultural products by Ciaian and Kancs (2011) is based on monthly end-of-period data; and the econometric model for the U.S. dairy industry developed recently by Mosheim (2012) is based on average quarterly data. As well, food price forecasts produced by the ERS are based on quarterly econometric, Box-Jenkins, and Vector Autoregressive Models (Joutz, Trost, Hallahan, Clauson and Denbaly, 2000).

On the other hand, stemming from the work of Parkinson (1977, 1980), there has been substantial evidence showing that for many price variables, the *range* over a given period also contains important information, particularly in relation to price volatility. The price range is the difference between the highest and lowest prices of the period. It has long been a popular tool in technical trading analysis. For example, the *candlestick chart* uses information about the range of price movement over a period, and *the average true range* uses the high and low prices of current and previous periods to provide an indication of price volatility. Subsequent to the work of Parkinson came a stream of studies on range-based forecasting models of volatility. Some better known examples include Gallant, Hsu and Tauchen (1999) and Alizadeh, Brandt and Diebold (2002), who developed range-based models for estimating stochastic volatility. Chou (2005) introduced the conditional autoregressive range model that combines range analysis and GARCH innovations, while Brandt and Jones (2006) proposed the range-based exponential GARCH model. Of particular relevance to this paper are the recent studies by Cheung, Cheung and Wan (2009) and He and Wan (2009). The former study showed that for many stock indices, the daily high and low are cointegrated, and the error correction term is closely approximated by the daily range of the stock index. The authors of this study also found evidence of the vector error correction model (VECM) that exploits the interaction between the daily high, daily low and the range significantly outperforming the univariate ARIMA model in out-of-sample forecasts. He and Wan (2009) found similar results when modelling the exchange rate of the U.S. dollar versus the Pound sterling and the Yen.

The existence of cointegration between the highest and lowest prices implies that these two extreme quantities are tied up by a long run equilibrium relationship. Although it is possible to deviate from the long run position in the short run, these deviations tend to disappear with the passage of time as a result of the tendency to move back to the equilibrium. Cointegration analysis has been used extensively in agricultural and food research. See Gutierrez, Westerlund and Erickson (2007), Ciaian and Kancs (2011), and Serra, Zilberman and Gil (2011) for recent examples. Perhaps the most salient feature of the results of Cheung, Cheung and Wan (2009) and He and Wan (2009) is that the price range is found to be a close proxy for the error correction term implied by the cointegrating relationship. This is significant as it hints at the potential gain of a simple substitution of the range for the entity capturing the long run behaviour in forecasting. Since the implementation is easy, this methodology holds promise, especially for practitioners. To the best of our knowledge, this approach has not been utilised in agricultural price forecasting, and the aim of this paper is to take steps in this direction.

In this paper, two agri-food commodities are focused; corn and soybeans. These two commodities are considered major commodities due to the wide range of products, in which they are used, and their large trading volumes and turnovers in the world market. Because of their importance, they have been the subject of investigation in many agricultural and food studies. Some recent examples include Irwin, Good and Martines-Filho (2006), Bekkerman, Goodwin and Piggott (2008), Bernard and Bernard (2010), Yu and Babcock (2010), and Urcola and Irwin (2011). Corn and soybeans are primarily used as food and feed, but they are also used for the production of ethanol fuel and as industrial materials. Exports from the U.S., Argentina and Brazil make up the overwhelming supply of corn and soybeans in the world market, with China and Japan being their major importers. Corn and soybeans are usually traded as futures contracts at exchanges like the Chicago Board of Trade (CBOT). They are measured in bushels with one contract covering 5000 bushels. Corn futures are delivered in March, May, July, September and December, while soybeans futures are delivered in January, March, May, July, August, September and November of every year. Huge price movements have been observed for corn and soybeans futures in recent years, making planning very difficult for market participants. It is noted that between 2007 and 2011, corn (soybeans) futures price per contract averaged $476.5 ($1,097.5), reaching a record high of $787 ($1,658) in June 2011 (July 2008) and a low of $293.5 ($653.5) in December 2008 (January 2007) (See http://wikiposit.org/p?futures). The rapidity and intensity of these price fluctuations call for an in-depth analysis of the underlying price behaviour. This paper shows that some valuable insights of the price movements can be obtained from the daily range of the prices. More specifically, it is found that the price range contains information that can lead to more accurate forecasts of the future high and low prices, and these improved forecasts have the potential to offer a better guide to policy makers, traders and other market participants.

The remainder of this paper is structured as follows. Section 2 contains a description of the data, along with results confirming the existence of long run equilibrium between the highs and lows of the daily prices of each of corn and soybeans futures; in particular, special attention is paid to the role of the daily range in the long run relationship. Section 3 compares the out-of-sample forecasting accuracy of the VECM that takes into account this long run relationship with the univariate ARIMA and random walk models with respect to some common error summary measures. As an application, in Section 4 a trading strategy of corn and soybean futures is developed that makes use of the forecasts of the daily high and low prices. It is found that this trading strategy generally yields very good profits over the evaluation period considered. Some concluding remarks are presented in Section 5.

2. Data Source and Modelling Methodology

Our analysis is based on the daily high and low price data of the corn futures contract CZ10 and the soybeans contract SX10 at the CBOT. These two contracts are selected as the basis of this study because of their large trading volumes compared with other contracts over similar period. The tradable period is 6th February 2007 – 14th December 2010 for CZ10, and 1st November 2007 – 12th November 2010 for SX10. These periods contain 975 and 758 trading days respectively, but after discounting the days in which no trading was recorded, these numbers reduce to 973 and 752. In all cases, the first 400 records of the data series is used to estimate the model parameters and the remaining observations for ex-post forecast comparisons. All our data are downloaded from Bloomberg.

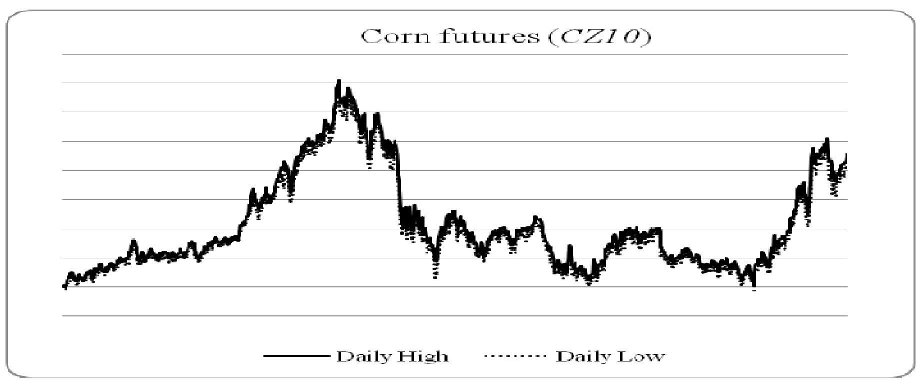

Fig. 1 Daily high and low of Corn futures CZ10

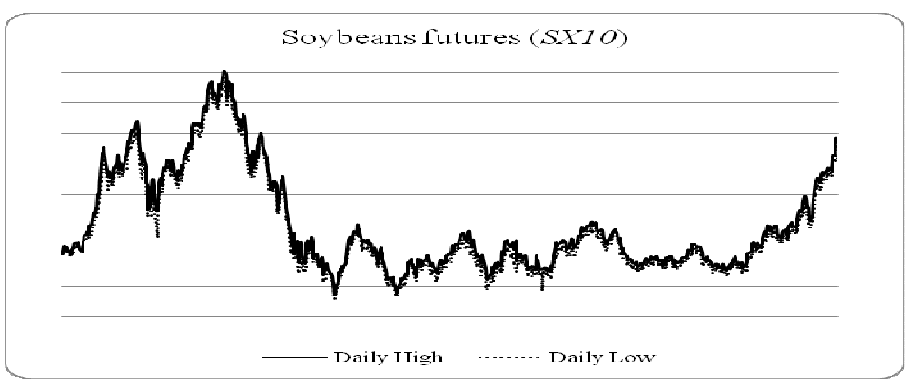

Fig. 2 Daily high and low of Soybeans futures SX10

Daily high and low of corn futures prices are denoted as CH_t and CL_t respectively. The corresponding prices of soybeans are denoted as SH_t and SL_t Figures 1 and 2 present the plots of these series. None of the series look stationary. The plots also reveal a persistent tendency of co-movement between each pair of high and low, suggesting the possibility of cointegration. Table 1 presents the Augmented Dickey-Fuller (ADF) test results for the orders of integration of these series. The Schwartz Bayesian Criterion (SBC) is used to determine the lag-lengths of the Dickey-Fuller regressions. The test results confirm that all four high/low series are non-stationary but their first differences are stationary. Accordingly, each of the four high/low series is first-order integrated (I(1)).

Table 1
Augmented Dickey-Fuller (ADF) Test Results

	Level		First difference	
	t-statistic for testing $\gamma = 0$	Lag length (L)	t-statistic for testing $\gamma = 0$	Lag length (L)
CH_t	-0.91[a]	1	-8.67[b]	5
CL_t	-0.97[a]	1	-7.53[b]	8
CR_t	-5.34[b]	4		
SH_t	-1.22[a]	1	-9.04[b]	3
SL_t	-1.33[a]	1	-5.95[b]	8
SR_t	-4.41[b]	6		

Notes: Let Y_t be one of CH_t, CL_t CR_t, SH_t, SL_t and SR_t. Our ADF tests are based on the following Dickey-Fuller regressions with and without drift/trend terms, and adopt the sequential procedure of Dolado, Jenkinson and Sosvilla-Rivero (1990) (see Giles, Giles and McCann, 1993 for a description of this procedure) to determine the order of integration:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{j=1}^{L} \beta_j \Delta Y_{t-j} + \varepsilon_t \tag{8}$$

$$\Delta Y_t = \alpha + \gamma Y_{t-1} + \sum_{j=1}^{L} \beta_j \Delta Y_{t-j} + \varepsilon_t \tag{9}$$

$$\Delta Y_t = \gamma Y_{t-1} + \sum_{j=1}^{L} \beta_j \Delta Y_{t-j} + \varepsilon_t \qquad (10)$$

[a] Cannot reject $\gamma = 0$ and $\gamma = \beta = 0$ in (8); cannot reject $\gamma = 0$, but reject $\gamma = \alpha = 0$ in (9); t-statistic compared with standard normal critical value; reported t-statistic and the associated lag-length are for testing $\gamma = 0$ in (9).

[b] Reject $\gamma = 0$ in (8); reported t-statistic and the associated lag-length are for testing $\gamma = 0$ in (8).

Next test for the presence of cointegration is tested between high-low pair. Table 2 reports the testing results of Johansen's (1988, 1991) Trace and Maximal-eigenvalue tests. Again, the SBC is used to choose the lag parameter. For each high-low pair, the tests convincingly reject the null hypothesis of no cointegration, but do not reject the null of a single cointegrating relationship. Given that each series is individually I(1), the test results imply that for each pair of high and low, a vector of coefficients exists to form a stationary linear combination of the two variables. That is, the high and its corresponding low series have the same stochastic trend that drives them individually to wander randomly over time, and an appropriate linear combination of the high and low can eliminate the effects of the common

Table 2
Cointegration Test Results

Corn futures			
H_o	Maximal Eigen value statistic	Trace statistic	Lag length
r = 0	56.88*	57.92*	4
r = 1	1.04	1.04	4
Cointegrating vector		(1, -1.048)	
Soybeans futures			
H_o	Maximal eigenvalue statistic	Trace statistic	Lag length
r = 0	44.93*	45.97*	4
r = 1	1.37	1.33	4
Cointegrating vector		(1, -1.024)	

Notes: r = 0 and r = 1 represent the null of no and one cointegrating vector respectively; the asterisk * indicates that the test is rejected at the 5% level of significance.

stochastic trend.

Table 2 also reveals that the normalized cointegrating equations for the two pairs of highs and lows are:

$$CH_t\text{-}1.048CL_t = CZ_t \tag{1}$$

and

$$SH_t\text{-}1.024SL_t = SZ_t, \tag{2}$$

where CZ_t and SZ_t are the error correction terms, reflecting the extent to which the equilibrium is not met. These estimated cointegrating equations, which capture the empirical long run relationship, suggest that the daily high and low tend to move almost on a one-to-one basis. Note that (1) and (2) are very close to

$$CH_t\text{-}CL_t = CR_t \tag{3}$$

and

$$SH_t\text{-}SL_t = SR_t \tag{4}$$

respectively, where CR_t and SR_t are the daily price ranges of the two commodities. Indeed, ADF test results in Table 1 show that CR_t and SR_t are I(0), indicating that they are reasonable proxies for the error correction terms. Thus, although the daily high and low of each commodity are non-stationary, their I(1) behaviour exactly offsets each other over time, so that the difference between them (i.e., the range) is stationary. In the balance of this paper, CR_t and SR_t are treated as the error correction terms. This has the advantage of reducing the computational burden in forecasting.

The error correction representation (Engle and Granger, 1987) is the most common approach to situations that incorporate both the information relating to the long run equilibrium and short run dynamics. Using the range variable CR_t as a proxy for the error correction term CZ_t the VECM of the daily high and low prices of corn futures may be represented by the following two equations:

$$\Delta CH_t = \mu_1 + \sum_{i=1}^{l} \Gamma_{1i} \Delta CH_{t-i} + \sum_{i=1}^{l} \Phi_{1i} \Delta CL_{t-i} + \alpha_1 CR_{t-1} + \varepsilon_{1,t} \tag{5}$$

and

$$\Delta CL_t = \mu_2 + \sum_{i=1}^{l} \Gamma_{2i} \Delta CH_{t-i} + \sum_{i=1}^{l} \Phi_{2i} \Delta CL_{t-i} + \alpha_2 CR_{t-1} + \varepsilon_{2,t}\ , \tag{6}$$

where $\mu_1, \mu_2, \Gamma_{1i}\text{'}s, \Gamma_{2i}\text{'}s,\ \Phi_{1i}\text{'}s,\ \Phi_{2i}\text{'}s,\ \alpha_1$ and α_2 are parameters to be estimated, and l is the lag order of the short run dynamics. The VECM equations for soybeans futures are similarly specified with ΔSH_t replacing ΔSL_t replacing ΔCL_t and SR_t replacing CR_t in (5) and (6). In the VECM equations, α_1 and α_2 measure the proportion of last period's disequilibrium that is corrected for in the current period, and $\Gamma_{1i}\text{'}s$ and $\Gamma_{2i}\text{'}s$ measure the extent to which the highs and lows respond to the short run transitory effects. Again, we use the SBC to determine the lag length l. It is found that for the VECMs of both corn and soybeans futures, the choice of $l = 3$ minimizes the SBC.

Table 3 gives the estimation results. The Q-statistics indicate the lag specification adequately captures the short run dynamics. The results also show that for corn futures, the range variable is significant in the high equation, but for soybeans futures, it is significant in the low equation. The significance of the range variable indicates that the range variable proxies the error correction term well, and is consistent with the cointegration results. In both VECMs, the coefficient estimate of the range variable is negative in the other lag variables they are mostly positive. For instance, consider the daily

Table 3
VECM Specifications Based on Equations (5) and (6) and Estimated Coefficients

Corn futures					
ΔCH_t			ΔCL_t		
	Coefficient Estimate	t - statistic		Coefficient Estimate	t - Statistic
μ_1	1.68869	2.44	μ_2	0.08058	0.12
Γ_{11}	-0.12819	-1.48	Γ_{21}	0.38808	4.73
Γ_{12}	-0.12042	-1.39	Γ_{22}	0.22544	2.75
Γ_{13}	0.02955	0.38	Γ_{23}	0.12113	1.66
Φ_{11}	0.26777	2.94	Φ_{21}	-0.17816	-2.07
Φ_{12}	0.10575	1.20	Φ_{22}	-0.23228	-2.78
Φ_{13}	0.07371	0.94	Φ_{23}	-0.16996	-2.30
α_1	-0.12909	-2.17	α_2	0.04126	0.69
	Q - statistic	p - value		Q – statistic	p – value
Q(6)	4.39	0.63		6.14	0.41
Q(12)	9.17	0.69		11.42	0.49
Q(18)	12.73	0.81		14.28	0.71
Soybeans futures					
ΔSH_t			ΔSL_t		
	Coefficient estimate	t - statistic		Coefficient estimate	t - statistic
μ_1	2.49711	0.88	μ_2	-7.00019	-2.47
Γ_{11}	-0.32212	-2.75	Γ_{21}	0.15964	1.36
Γ_{12}	-0.30839	-2.85	Γ_{22}	0.00913	0.08
Γ_{13}	-0.16603	-1.96	Γ_{23}	0.00808	0.10
Φ_{11}	0.51694	4.48	Φ_{21}	0.09420	0.82
Φ_{12}	0.25383	2.36	Φ_{22}	-0.02468	-0.23
Φ_{13}	0.12196	1.43	Φ_{23}	-0.07195	-0.84
α_1	-0.11643	-1.03	α_2	0.28904	2.56
	Q - statistic	p - value		Q – statistic	p - value
Q(6)	10.22	0.12		1.22	0.98
Q(12)	13.93	0.31		13.14	0.36
Q(18)	24.92	0.13		25.39	

high equation of corn futures, where the coefficient estimates of the lagged daily high differences are mostly negative and the lagged daily low differences are all positive. The negative coefficients are indicative of the presence of regressive behaviour. Higher daily highs tend to regress to a lower level, and lower daily highs tend to return to a higher level. On the other hand, the positive coefficients of the lagged daily low differences suggest certain spill over effects. Higher (Lower) daily lows lead to higher (lower) daily highs.

3. Forecast Performance Comparisons

To evaluate the out-of-sample forecasting ability of the VECMs developed in the last section, we compare the high and low forecasts generated by these VECMs with those obtained from two alternative methods. The first is based on the ARIMA specifications of the highs and lows. Now, let *Yt* be one of CH_t, CL_t, SH_t and SL_t. The ARIMA(p,d,q) model is given by

$$\phi_p(B)\Delta^d Y_t = \delta + \theta_q(B)\varepsilon_t \qquad (7)$$

Where $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - ..... - \phi_p B^p$,

$\theta_p(B) = 1 - \theta_1 B - \theta_2 B^2 - ..... - \theta_q B^q$, B is the backward shift operator, and δ is the drift term.

Again, the SBC is used to determine the lag order of each of the ARIMA processes. Table 4 provides the chosen ARIMA specifications and estimates of the ARIMA coefficients in the models. The second is based on the naive random walk models of highs and lows. As indicated in Section 3, the ex-post forecast period is 8 September 2008 – 14 December 2010 for corn futures, and 23rd June, 2009 – 12th November, 2010 for soybeans futures. We consider forecast horizons of *h*=1, 2, 3 and 6 days. These forecasts are computed using a dynamic recursive procedure by which after every *h*-period ahead forecast, the models are re-estimated based on the same specifications as those given in Tables 3 and 4, but with the number of observations

Table 4
ARIMA Specifications Based on Equation (7) and Estimated Coefficients

	Corn futures			
	CH_t: ARIMA(0,1,3)		CL_t: ARMIA(0,1,6)	
	Coefficient Estimate	t - statistic	Coefficient estimate	t – statistic
δ	0.59308	1.42	0.58016	1.97
Θ_1	-0.02636	-0.53	-0.12321	-2.45
Θ_2	-0.04364	-0.87	0.04028	0.79
Θ_3	-0.11975	-2.39	0.06499	1.28
Θ_4	-		-0.03994	-0.77
Θ_5	-		-0.01005	-0.19
Θ_6	-		0.18649	3.62
	Q – statistic	p - value	Q - statistic	p - value
Q(6)	4.97	0.18	-	-
Q(12)	9.82	0.37	5.66	0.46
Q(18)	13.37	0.57	9.29	0.68
	Soybeans futures			
	SH_t: ARMIA(0,1,6)		SLt: ARIMA(8,1,1)	
	Coefficient estimate	t - statistic	Coefficient estimate	t – statistic
δ	-0.19361	-0.15	-0.34574	-0.17
Θ_1	-0.09348	-1.86	-0.89451	-13.16
Θ_2	0.02423	0.48	-	-
Θ_3	0.03406	0.67	-	-
Θ_4	-0.09504	-1.89	-	-
Θ_5	0.07568	1.49	-	-
Θ_6	-0.12728	-2.51	-	-
$\emptyset_1$	-	-	-0.71840	-8.84
$\emptyset_2$	-	-	0.14982	2.38
$\emptyset_3$	-	-	-0.08211	-1.31
$\emptyset_4$	-	-	-0.05017	-0.80
$\emptyset_5$	-	-	0.03254	0.52
$\emptyset_6$	-	-	-0.02596	-0.41
$\emptyset_7$	-	-	0.08082	1.29
$\emptyset_8$	-	-	0.17619	3.38
	Q – statistic	p - value	Q - statistic	p – value
Q(6)	-	-	-	-
Q(12)	4.64	0.59	5.49	0.14
Q(18)	14.16	0.29	12.30	0.20

increased by 1. Also, in computing the forecasts when h>1, it is assumed that the actual highs and lows that lie in the forecast period are unknown, and their values are replaced by their forecasts generated from previous rounds if these values are needed as explanatory variables in the forecasting equation.

Comparisons of the mean-absolute and root mean-squared forecast errors of the highs and lows are presented in Table 5. To facilitate readability, the best and worst forecasts in each case are flagged by a † and an * respectively. A clear picture emerges from the results. In terms of root mean-squared forecast errors, the two VECMs always out-perform their corresponding ARIMA and random walk counterparts that ignore the interaction between

Table 5
Out-of-Sample Forecast Performance

Corn futures		Random Walk		ARIMA		VECM	
		CH_t	CL_t	CH_t	CL_t	CH_t	CL_t
h=1	MAFE	5.9754	5.6875	6.0573 *	5.9094 *	5.7418 †	5.5599†
	RMSFE	8.4073	8.1154	8.4542 *	8.3791 *	7.9622 †	7.8091†
h=2	MAFE	10.3178*	10.0167*	6.2256	5.9256	5.8730†	5.5888†
	RMSFE	27.7758*	27.3392*	8.6516	8.3491	8.0153†	7.8589†
h=3	MAFE	13.3996*	13.3319*	6.3445	6.1574	5.9162†	5.6357†
	RMSFE	38.0835*	37.5217*	8.7404	8.5910	8.0872†	7.7531†
h=6	MAFE	20.4683*	20.8732*	6.9373	7.0524	6.4448†	6.0712†
	RMSFE	59.1697*	58.4996*	9.2286	9.7544	8.3770†	8.4124†
Soybeans futures		Random Walk		ARIMA		VECM	
		SH_t	SL_t	SH_t	SL_t	SH_t	SL_t
h=1	MAFE	8.9510	8.7152	9.1162*	8.8791*	8.7869†	8.4726†
	RMSFE	13.1445	12.9353	13.3975*	13.1741*	12.9062†	11.7221†
h=2	MAFE	16.1527*	16.4581*	9.0371	8.8401†	8.8855†	8.8629
	RMSFE	54.1140*	53.2380*	13.2494	13.0025	13.1617†	12.0355†
h=3	MAFE	22.4851*	22.2734*	9.1096	9.2888†	8.9613†	9.4549
	RMSFE	75.3788*	73.8381*	13.3904	13.5817	13.3370†	12.5539†
h=6	MAFE	38.3445*	38.2386*	9.4946	9.3363†	9.2782†	10.9219
	RMSFE	116.9397*	114.2642*	13.6351†	13.4766†	13.6966	13.8473†

Notes: MAFE = mean-absolute forecast errors; RMSFE = root mean-squared forecast errors

the highs and lows, and the accuracy of the VECMs relative to their competitors generally improves as *h* increases. For example, in the case of corn futures, when *h*= 1, the forecast of the high price produced by the

VECM outperforms the next best forecast in root mean-squared error terms by 4.07 percent, and this percentage increases to 6.00 percent, 7.24 percent and 7.64 percent when increases to 2, 3 and 6 respectively. Except for the case of h= 1, the random walk model always produces the worst forecasts in root mean-squared error terms. These conclusions generally carry over to comparisons based on mean-absolute forecast errors, except that the ARIMA model can occasionally have a slight edge over the VECM when predicting the daily low prices of soybeans futures.

4. An Application: A Trading Strategy Based on High and Low Forecasts

The preceding section shows that the VECM that accounts for the information contained in the price range often works well relative to traditional models and could deliver more reliable forecasts of prices. It is hoped that this will increase the appeal of the range-based approach in agricultural and food research to benefit investors, traders and policy makers. Here, as *one* application of range-based forecasting, a trading strategy of corn and soybeans futures is proposed that makes use of the out-of-sample forecasts of high and low. The spirit of this strategy is similar to the stock trading rule that uses signals from barrier options developed by Cheung, Cheung, He and Wan (2010) but it is also different from that rule.

Trading rule in this study works as follows. Take, for example, the case of corn futures. Let the opening and closing prices of corn futures on day t be CO_t and CC_t respectively, and the forecasted high and low of corn futures for day $t+h$, formed on day t, be PCH_{t+h} and PCL_{t+h} respectively. The following steps summarize our trading rule:

Step 1: On a given day t, if $PCH_{t+h}- CO_t > CO_t - PCL_{t+h}$ then generate a "buy alert signal".

Step 2: Counting from day t, if the "buy alert signal" persists for $m\ (\geq 1)$ consecutive days, then buy the corn futures on day $t+m-1$ using the closing price of that day.

Step 3: After buying the futures, on another day *s*, if PCH_{s+h}-CO_s< CO_s-PCL_{s+h} , then generate a "sell alert signal".

Step 4: Counting from day *s*, if the "sell alert signal" persists for m (≥ 1) consecutive days, then sell the corn futures on day *s+m-1* using the closing price of that day.

The rationale of this strategy is based on the general observation that market participants tend to over-react to price movements, and this over-reaction often drive prices further up or down, depending on the market condition at the time. The success of this strategy depends heavily on the accuracy of forecasts as well as market behaviour. The condition under Step 1 of the trading rule will arise if CO_t, day *t's* opening price, either is below PCL_{t+h} or lies within the interval bounded by PCL_{t+h} and PCH_{t+h} and closer to PCL_{t+h} than to PCH_{t+h}. Assuming that the forecasts formed on day *t* are an accurate reflection of the market situation on day *t+h*, the price will more likely move up than down in the next *h* days subsequent to day *t*. If the same condition is observed for *m* consecutive days, then it is a reasonable belief that the market has formed an upward trend, and the investor should act according to Step 2 and enter the market. It is expected that the rising price will continue until a counter-reaction appears, which will ultimately lead to the emergence of a sell signal. In the latter case, a converse argument may be used to explain Steps 3 and 4.

Of course, forecasting the market can never be certain, and it must be reiterated that the accuracy of the high and low forecasts plays a dominating role in the strategy. Given that the above strategy is by no means foolproof, and its success is largely an empirical issue, we simulate the buy/sell actions of corn and soybeans futures according to the above steps over the forecast periods considered in Section 3. In our experiment, we set m, the waiting period, to 1, 2 and 3 days, and h, the forecast horizons, to 1, 2, 3 and 6 days, resulting in 12 scenarios for the trading of each commodity futures, and 24 scenarios in total. Also, it is not required that the investor sells the futures contract he/she bought previously before he/she can enter the market again, i.e., it is assumed that the investor acts on a continuous mode. If there are unsold futures contracts at the end of the evaluation period, these futures will not be counted in our profit calculation. When calculating the return resulting from each trade, a one-way 0.1 percent deduction is included to mimic transaction cost. As every trade has a different holding period, the return of each trade is annualised in order to facilitate comparisons.

Table 6 reports the average, best, worst and standard deviation of annualised return, the total number of trades, and percentages of trades with positive return for corn and soybeans futures based on our proposed strategy. The annualised return is calculated as follows. Let C_p and C_{p+j} ($j \geq 1$) be the futures' closing price on the buying and selling days respectively. The actual percentage return of a given trade, net of transaction cost, is given by $R = \left((C_{p+j} - C_p) / C_p\right) \times 100\% - 0.1\%$. The annualised return of this trade is $AR = (R / j) \times 250$, where the factor 250 mimics the number of trading days in a year. We also report a z-test for testing if the average annualised return is significantly different from zero. The results show that the average annualised returns based on our strategy are always in positive territories. The average annualised return from trading corn futures ranges from 9.28 percent to 67.17 percent, while that from trading soybeans futures ranges from 14.40 percent to 82.50 percent. The z-tests also show that in 20 out of the total 24 scenarios, the average annualised return differs significantly from zero. It is also apparent that for the majority of scenarios considered, trades based on our proposed strategy yield positive returns more frequently than negative returns. Generally speaking, the percentages of trades with positive returns are higher when *m=3* than when *m=1* or 2. For example, with corn futures, when *m=3* and *h=6* nearly all of the 295 trades result in profits, with the average profit being 42.56 percent in annualised terms. For soybeans futures, when *m=3 and h=1* over 82 percent of trades result in profits, with the average being 77.32 percent in annualised terms. The less favourable outcomes associated with *m=1* or 2 compared with *m=3* are perhaps not

Table 6
Summary of Trading Strategy Results

	h=1	h=2	h=3	h=6
Corn futures				
m=1				
Average annualised return	67.17%	59.04%	56.84%	34.25%
Worst annualised return	-1330%	-1330%	-1330%	-1064%
Best annualised return	1480%	1480%	1480%	1459%
% of trades with positive returns	53.53%	53.20%	57.36%	50.37%
Std. of annualised returns	392.97%	361.82%	339.50%	278.48%
No. of trades	325	359	387	395
z-stat	3.0819	3.0922	3.2936	2.4447
m=2				
Average annualised return	36.53%	36.44%	9.28%	30.77%
Worst annualised return	-554.58%	-554.58%	-554.58%	-554.58%
Best annualised return	972.77%	763.06%	763.06%	763.06%
% of trades with positive returns	61.14%	65.43%	51.16%	63.81%
Std. of annualised returns	217.54%	187.46%	169.98%	137.32%
No. of trades	175	217	258	351
z-stat	2.2219	2.8638	0.8771	4.1981
m=3				
Average annualised return	55.12%	35.46%	49.33%	42.56%
Worst annualised return	-136.56%	-136.56%	-136.56%	-139.46%
Best annualised return	720.10%	720.10%	720.10%	720.10%
% of trades with positive returns	70.96%	57.36%	77.84%	93.89%
Std. of annualised returns	118.28%	101.03%	84.53%	67.94%
No. of trades	93	129	176	295

z-stat	4.4939	3.9865	7.7431	10.7609
Soybeans futures				
m=1				
Average annualised return	45.11%	14.40%	24.96%	36.78%
Worst annualised return	-1382%	-1382%	-1382%	-1288%
Best annualised return	1106%	1164%	1164%	1618%
% of trades with positive returns	56.31%	49.68%	51.49%	46.84%
Std. of annualised returns	300.76%	319.36%	336.00%	344.22%
%No. of trades	190	159	134	111
z-stat	2.0677	0.5689	0.8602	1.1259
m=2				
Average annualised return	76.51%	82.50%	54.88%	56.24%
Worst annualised return	-186.43%	-490.20%	-490.20%	-490.20%
Best annualised return	455.93%	426.21%	418.13%	418.13%
% of trades with positive returns	71.42%	70.76%	63.63%	62.79%
Std. of annualised returns	131.93%	138.02%	123.85%	137.35%
No. of trades	91	65	55	43
z-stat	5.5321	4.8195	3.2864	2.6849
m=3				
Average annualised return	77.32%	77.16%	69.17%	79.75%
Worst annualised return	-95.70%	-95.70%	-95.70%	-95.70%
Best annualised return	281.81%	241.89%	276.09%	208.99%
% of trades with positive returns	82.05%	73.07%	70.37%	76.19%
Std. of annualised returns	87.95%	99.69%	112.84%	100.02%
No. of trades	39	26	27	21
z-stat	5.4899	3.9468	3.1853	3.6536

surprising because too short a waiting period can result in substantial noises in trading. Although our strategy can also result in individual trades with negative returns, the losses, both in terms of percentage magnitude and the frequency of occurrences, are generally more than compensated for by the gains.

Overall, the application of the proposed strategy to real data appears to be quite encouraging. Clearly, the success of this strategy depends in part on the accuracy of the VECM forecasts. As a comparison, we have also simulated the same buy/sell actions based on ARIMA forecasts that ignore the interaction between the high, low and the range. The results across all dimensions are generally inferior to those based on VECM forecasts. In particular, the trading strategy based on ARMA forecasts generally result in smaller percentages of trades with positive returns, and can sometimes deliver negative average annualised returns. The detailed findings are not reported here for brevity, but they are available upon request from the authors.

5. Conclusion

The recent financial forecasting literature has shown that range-based models are more informative and often yield predictions that are superior to those obtained from traditional forecasting methods. Building heavily on this existing literature, an attempt has been made to construct price forecasting models of corn and soybeans futures that exploit the long run interaction between the daily high and low prices and the price range. Specifically, we find that for both of these commodity futures, the daily high and low prices obey a long run equilibrium relationship, with the disequilibrium response in each period being well proxied by the daily price range. This leads to a VECM with a cointegrating vector restriction resembling the price range. Our results show that this approach leads to increased forecast accuracy compared to the traditional ARIMA models. To further demonstrate the usefulness of the methodology, we develop a trading strategy of corn and soybeans futures based on the high and low forecasts that proves to be effective in generating profits. This is just one among the many applications in which the range-based forecasts can be put to use. To the best of our knowledge, this is the first study of agricultural price forecasting using range-based methods, and it is hoped that the current study can help increase the awareness and the appeal of this approach among agricultural and food policy researchers.

References

Alizadeh, S., Brandt, M.W., and Diebold, F.X. (2002). Range-Based Estimation of Stochastic Volatility Models. *Journal of Finance, 57,* 1047 – 1091.

Allen, P.G. (1994). Economic Forecasting in Agriculture. *International Journal of Forecasting, 10*, 81 – 135.

Bekkerman, A., Goodwin, B.K. and Piggott, N.E. (2008). Spatio-temporal Risk and Severity Analysis of Soybean Rust in the United States. *Journal of Agricultural and Resource Economics, 33,* 311 – 331.

Bernard, J.C., and Bernard, D.J. (2010). Comparing Parts with the Whole: Willingness to Pay for Pesticide-Free, Non-GM, and Organic Potatoes and Sweet Corn. *Journal of Agricultural and Resource Economics, 35*, 457 – 475.

Brandt, M., and Jones, C. (2006). Volatility Forecasting with Range-Based EGARCH Models. *Journal of Business and Economic Statistics, 24*, 470 – 486.

Cheung, Y.L., Cheung, Y.W., and Wan, A.T.K. (2009). A High-Low Model of Daily Stock Price Ranges. *Journal of Forecasting, 28,* 103 – 119.

Cheung, Y.L., Cheung, Y.W., He, A.W.W., and Wan, A.T.K. (2010). A Trading Strategy Based on Callable Bull/Bear Contracts. *Pacific Basin Finance Journal, 18,* 186 – 198.

Chou, R.Y. (2005). Forecasting Financial Volatilities with Extreme Values: The Conditional Autoregressive Range (CARR) Model. *Journal of Money, Credit and Banking, 37*, 561 – 582.

Ciaian, P., and Kancs, D. (2011). Food, Energy and Environment: Is Bioenergy the Missing Link? *Food Policy, 36*, 571 – 580.

Dhehibi, B., and Gil, J.M. (2003). Forecasting Food Demand in Tunisia Under Alternative Pricing Policies. *Food Policy, 28*, 167 – 186.

Dolado, J.J., Jenkinson, T., and Sosvilla-Rivero, S. (1990). Cointegration and Unit Roots. *Journal of Economic Surveys, 4*, 249 – 273.

Engle, R.F., and Granger, C.W.J. (1987). Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica, 55*, 251 – 276.

Gallant, A.R., Hsu, C.T., and Tauchen, G. (1999). Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance. *Review of Economics and Statistics, 81*, 617 – 631.

Giles, D.E.A., Giles, J.A., and McCann, E. (1993). Causality, Unit Roots and Export-Led Growth: The New Zealand Experience. *Journal of International Trade and Economic Development, 1,* 195 – 217.

Gutierrez, L., Westerlund, J., and Erickson, K. (2007). Farmland Prices, Structural Breaks and Panel Data. *European Review of Agricultural Economics, 34*, 161 – 179.

He, A.W.W., and Wan, A.T.K. (2009). Predicting Daily Highs and Lows of Exchange Rate: A Cointegration Analysis. *Journal of Applied Statistics, 36*, 1191 – 1204.

Irwin, S.H., Good, D.L., and Martines-Filho, J. (2006). The Performance of Agricultural Market Advisory Services in Corn and Soybeans. *American Journal of Agricultural Economics, 88*, 162 – 181.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control, 12*, 231 – 254.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica, 59*, 1551 – 1580.

Joutz, F.L., Trost, R.P., Hallaham, C., Clauson, A., and Denbaly, M. (2000). *Retail Food Price Forecasting at ERS: The Process, Methodology and Performance from 1984 to 1997* (Technical Bulletin No. 1885). Washington: USA, Economic Research Service, U.S. Department of Agriculture.

Mosheim, R. (2012). A Quarterly Econometric Model for Short-Term Forecasting of the U.S. Dairy Industry (Technical Bulletin No. 1932). Washington: USA, Economic Research Service, U.S. Department of Agriculture.

Parkinson, M. (1977). Option Pricing: The American Put. *Journal of Business, 50*, 21 – 36.

Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *Journal of Business, 53*, 61 - 65.

Serra, T., Zilberman, and Gil, J. (2011). Price Volatility in Ethanol Markets. *European Review of Agricultural Economics, 38*, 259 – 280.

Urcola, H. A., and Irwin, S. H. (2011). Are Agricultural Options Overpriced? *Journal of Agricultural and Resource Economics, 36*, 63 – 77.

Yu, T., and Babcock, B.A. (2010). Are U.S. Corn and Soybeans Becoming More Drought Tolerant? *American Journal of Agricultural Economics, 92*, 1310 – 1323.