INTERNATIONAL JOURNAL OF FINANCE AND ECONOMICS
Int. J. Fin. Econ. **12**: 1–20 (2007)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/ijfe.303

AN EMPIRICAL MODEL OF DAILY HIGHS AND LOWS

YIN-WONG CHEUNG[a,b,*,†]
[a]*University of California, Santa Cruz, USA*
[b]*University of Hong Kong, Hong Kong*

ABSTRACT

We construct an empirical model for daily highs and daily lows of US stock indexes based on the intuition that highs and lows do not drift apart over time. Our empirical results show that daily highs and lows of three main US stock price indexes are cointegrated. Data on openings, closings, and trading volume are found to offer incremental explanatory power for variations in highs and lows within the VECM framework. With all these variables, the augmented VECM models explain 40–50% of variations in daily highs and lows. The generalized impulse response analysis shows that the responses of daily highs and daily lows to the shocks depend on whether data on openings, closings, and trading volume are included in the analysis. Copyright © 2007 John Wiley & Sons, Ltd.

JEL CODE: C32; G10

KEY WORDS: High; low open; close; trading volume; VECM model

1. INTRODUCTION

Stock price behaviour is quite intensively examined. While price data on open, high, low, and close are available, studies on stock returns and volatility usually employed only close-to-close return data. Indeed, the studies based on close-to-close return data outnumber those based on the other three price variables by a wide margin. Do the data on closings contain more information about price dynamics than the other three variables? Seemingly, the answer is not a definite yes. The high and the low, for instance, correspond to the prices at which the excess demand is changing its direction—the information that is not reflected by data on closing prices. Also, the price range, given by the difference of the high and the low, contains useful information on return volatility. In the seminal study Parkinson (1980) shows that the price range is a more efficient volatility estimator than, for example, the variance estimator based on close-to-close return data under certain assumptions.[1] Thus, there is no apparent reason to ignore information on the other three price variables in studying stock price behaviour.

Recently, there is a revived interest in studying the price range variable. In addition to examining its stochastic properties, some recent studies use the price range to model intertemporal volatility behaviour and, thus, incorporate it in various GARCH and stochastic volatility models to construct conditional or local variance estimators.[2] Mok *et al.* (2000), on the other hand, directly use data on highs and lows to test whether the S&P 500 and Hang Seng indexes follow a random walk specification. Overall, there is still a relatively small number of studies on the high and the low.[3]

The current exercise offers an exploratory analysis of the empirical properties of highs and lows. There are several reasons for analysing the high and the low. First, it is conceivable that data on highs and lows contain information that is not included in, say, the closings. For instance, the high and the low are the

*Correspondence to: Yin-Wong Cheung, Department of Economics, E2, University of California, Santa Cruz, CA 95064, USA.
† E-mail: cheung@ucsc.edu

turning points of the underlying price series, while the close is (usually) not. Further the high and the low can be used to construct other variables of interest such as the price range. In our exercise, it is shown that modelling the range using only its own history may be inferior to a model that jointly describes the behaviour of highs and lows.

Second, the pricing of some derivatives requires information on the high and the low. For example, exotic options such as the knock-out (knock-in) options and lookback options are constructed based on the highest price (or the lowest price) during an agreed upon period.[4]

Third, the high and the low are key components of some technical trading techniques.[5] For example, the price channel strategy initiates a buy (sell) when the price closes above (below) the upper (lower) channel constructed from daily highs and lows. Support and resistance levels are price levels at which there is a possible reverse of the trend. A breakthrough of these levels is considered as an important trading signal. In addition, highs and lows are used in forming trading techniques such as candlestick charts and stochastic oscillators.

The motivation of the empirical model of highs and lows used in the current study is quite intuitive. For equity markets in developed countries such as the US, stock prices exhibit stochastic trends and are typically characterized by $I(1)$ processes. Daily highs and lows, however, do not appear to drift apart from each other too far over time. If one assumes there is a stochastic trend underlying the stock price data generating process, both the high and low are likely to be driven by the same stochastic trend. If this is the case, then the high and the low can individually drift around without an anchor but their differences should not diverge over time. Thus, highs and lows may follow a cointegration relationship.

To explore the idea, we consider three main US stock indexes: the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index and formally test whether their highs and lows are cointegrated. To anticipate the results, the test corroborates the notion of cointegration between daily highs and daily lows. The vector error correction model derived from the cointegration relationship is extended to include other explanatory variables including opening prices, closing prices, and trading volumes. The responses of the high and the low to shocks are analysed in the presence of different groups of explanatory variables.

2. PRELIMINARY ANALYSES

In this study, we consider three main US stock indexes—the Dow Jones Industrial index (DJ), the NASDAQ index (NQ), and the S&P 500 index (SP). Daily data on opens, highs, lows, closes, and trading volume from 2 January 1990 to 31 December 2004 were downloaded from the Yahoo! Finance and Bloomberg L.P. websites. As a preliminary analysis, a modified Dickey–Fuller test known as the ADF-GLS test (Elliott *et al.*, 1996) is used to test for stationarity. The ADF-GLS test is shown to be approximately uniformly most powerful invariant. Let Y_t be a generic notation of a stock index's daily open (O_t), daily high (H_t), daily low (L_t), daily close (C_t), and daily trading volume (V_t) series, in logarithms. The price range R_t defined by $H_t - L_t$ is also considered. The ADF-GLS test that allows for a linear time trend is based on the following regression:

$$(1-L)Y_t^{\tau} = \alpha_0 Y_{t-1}^{\tau} + \sum_{k=1}^{p} \alpha_k (1-L) Y_{t-k}^{\tau} + \varepsilon_t \tag{1}$$

where L is the lag operator, Y_t^{τ} is the locally detrended process under the local alternative of $\bar{\alpha}$ and is given by $Y_t^{\tau} = Y_t - \tilde{\gamma}' z_t$ with $z_t = (1, t)'$. $\tilde{\gamma}$ is the least squares regression coefficient of $\tilde{Y}_t$ on $\tilde{z}_t$, where $(\tilde{Y}_1, \tilde{Y}_2, \ldots, \tilde{Y}_T) = (Y_1, (1-\bar{\alpha}L)Y_2, \ldots, (1-\bar{\alpha}L)Y_T)$, $(\tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_T) = (\tilde{z}_1, (1-\bar{\alpha}L)\tilde{z}_2, \ldots, (1-\bar{\alpha}L)z_T)$, and ε_t is the error term. The local alternative $\bar{\alpha}$ is defined by $\bar{\alpha} = 1 + \bar{c}/\mathrm{T}$ for which $\bar{c}$ is set to -13.5. The Bayesian information criterion is used to determine p, the lag parameter. If the estimated residuals do not pass the diagnostic test, then the lag parameter is increased until they do pass. The unit root hypothesis is rejected when the ADF-GLS test statistic, which is given by the usual t-statistic for $a_0 = 0$ against the alternative of $a_0 < 0$, is significant.[6]

DOI: 10.1002/ijfe

The test results are given in Table 1. The unit-root null hypothesis is not rejected by the four price series (O_t, H_t, L_t, and C_t) but is rejected by the range and trading volume data. To compare these results with those commonly reported in the literature, Table 1 also presents the results obtained from the augmented Dickey–Fuller test results, which is based on the regression $\Delta Y_t = \delta + \beta t + \gamma Y_{t-1} + \Sigma_{j=1}^{p} \beta_j \Delta Y_{t-j} + \varepsilon_t$. Both sets of test results suggest that O_t, H_t, L_t, and C_t are $I(1)$ variables and V_t and R_t are $I(0)$ variables. As indicated by the Q-statistics, the lag structures used to conduct these tests adequately capture the intertemporal dynamics.

We have to address the stationarity issue of trading volume before we proceed to the next stage of analysis. The detrending method used to achieve stationarity depends on data characteristics. While the trading volume does not contain a stochastic trend given by an $I(1)$ process, it has a significant deterministic trend component. Thus, we removed the estimated trend from trading volume data. Henceforth, V_t refers to the detrended volume data. The degrees of association between the stationary variables ΔO_t, ΔH_t, ΔL_t, ΔC_t, V_t, and R_t are presented in Table 2. The changes in opens, highs, and lows have a high correlation coefficient that ranges from 0.57 to 0.80 across the three stock indexes. ΔC_t tends to have a low correlation with ΔO_t but a high correlation with ΔH_t and ΔL_t. For the three index series, the trading volume has a small correlation coefficient with the changes in prices but a relatively large one with the range. The large correlation between trading volume and range may be driven by their association with volatility. Among the four price variables, the range has the largest correlation coefficient with changes in the low followed by changes in the opening. In the subsequent sections, a dynamic and multivariate setting is used to investigate the intertemporal properties of changes in highs and lows.

Table 1. Unit root tests

	ADF-GLS				ADF			
	STAT	LAG	Q5	Q10	STAT	LAG	Q5	Q10
(*A*) *The Dow Jones Industrial index*								
O	−1.352	8	0.007	1.422	−1.249	8	0.007	1.383
H	−1.332	3	5.663	7.909	−1.225	3	5.657	7.910
L	−1.513	6	0.069	5.640	−1.432	6	0.070	5.648
C	−1.53	1	6.455	14.612	−1.447	1	6.483	14.642
V	−7.106*	11	0.899	12.937	−9.588*	10	0.646	11.395
R	−5.043*	10	0.324	6.950	−6.253*	10	0.170	3.858
(*B*) *The NASDAQ index*								
O	−1.235	6	0.043	11.818	−1.210	6	0.044	11.857
H	−1.235	5	0.733	3.359	−1.202	5	0.749	3.425
L	−1.245	7	0.024	5.698	−1.223	7	0.024	5.672
C	−1.264	3	2.572	7.610	−1.232	3	2.572	7.633
V	−4.556*	13	1.257	11.019	−5.281*	10	1.916	15.072
R	−6.413*	10	1.127	12.329	−7.122*	10	0.847	9.226
(*C*) *The S&P 500 index*								
O	−1.071	8	0.025	1.149	−0.977	8	0.025	1.13
H	−1.184	2	2.611	9.948	−1.099	2	2.657	10.035
L	−1.156	10	0.006	0.229	−1.065	10	0.005	0.235
C	−1.085	8	0.014	0.660	−0.98	8	0.014	0.631
V	−5.081*	10	1.368	13.592	−5.548*	10	1.168	11.301
R	−4.993*	11	0.716	13.543	−8.109*	10	0.382	8.664

Note: The table reports results of applying the ADF-GLS and ADF tests to daily open (*O*), daily high (*H*), daily low (*L*), daily close (*C*), daily trading volume (*V*), and daily price range (*R*) series. Panels A, B, and C give results for the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index, respectively. 'ADF-GLS' and 'ADF', the ADF-GLS and ADF test results; 'STAT', the test statistics, 'LAG', the lag parameters used in the test procedures; 'Q5' and 'Q10', the Box-Ljung Q-statistics calculated from the first 5 and 10 estimated residual autocorrelations. None of the Q-statistics is significant. '*' The rejection of the unit root null hypothesis at the 5% level.

DOI: 10.1002/ijfe

3. AN EMPIRICAL MODEL

As previously stated, an empirical model for highs and lows is built based on the intuition that these two variables are interlinked and driven by some common dynamic factors. Results in the previous section show that the high and the low are $I(1)$ variables. Thus, the cointegration technique is used to investigate their dynamic interactions.

3.1. Cointegration test

The Johansen procedure is used to formally test for cointegration. Let $\mathbf{X}_t$ be a 2×1 vector containing the daily high and daily low series of a stock index (that is, $\mathbf{X}_t \equiv (H_t, L_t)'$) and has a $(p+1)$th-order autoregressive representation:

$$\mathbf{X}_t = \mu + \sum_{i=1}^{p+1} \gamma_i \mathbf{X}_{t-i} + \varepsilon_t \tag{2}$$

where μ is the intercept term, γ_i's are coefficient matrices, and ε_t is the innovation vector. To test whether the elements in $\mathbf{X}_t$ are cointegrated, the Johansen procedure tests for significant canonical correlations between $\Delta\mathbf{X}_t$ and $\mathbf{X}_{t-p-1}$ after adjusting for all intervening lags. Johansen (1991) and Johansen and Juselius (1990), for example, give a detailed description of the test.

The cointegration test results are reported in Table 3. Again, the Bayesian information criterion is used to select the lag parameter p and diagnostic tests are conducted to ensure the selected lag structure adequately describes data dynamics. According to both maximum eigenvalue and trace statistics, the null hypothesis of no cointegration is rejected. Further, there is no evidence that there exists more than one cointegrating vector. These results suggest that, for a given stock index, its daily high and daily low series are cointegrated. That is, the high and low series are driven by the same stochastic trend and individually wander randomly over time. However, an appropriate linear combination of highs and lows can eliminate the effects of the stochastic trend and form a stationary mean reverting series.

Table 2. Sample correlations

	ΔO	ΔH	ΔL	ΔC	V
(*A*) *The Dow Jones Industrial index*					
ΔH	0.573				
ΔL	0.597	0.800			
ΔC	0.076	0.667	0.639		
V	0.013	0.065	−0.034	0.016	
R	−0.134	−0.015	−0.217	−0.038	0.542
(*B*) *The NASDAQ index*					
ΔH	0.762				
ΔL	0.761	0.782			
ΔC	0.260	0.646	0.640		
V	0.002	0.030	−0.019	0.017	
R	−0.232	−0.105	−0.337	−0.121	0.474
(*C*) *The S&P 500 index*					
ΔH	0.634	1.000			
ΔL	0.651	0.641	1.000		
ΔC	0.075	0.585	0.570	1.000	
V	−0.063	−0.024	−0.094	−0.026	
R	−0.200	0.034	−0.366	−0.062	0.498

Note: The sample correlations between the stationary variables ΔO_t, ΔH_t, ΔL_t, ΔC_t, and V_t of the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index are reported.

DOI: 10.1002/ijfe

Table 3. Cointegration test results

	EIGENV	TRACE	*H*	*L*	LAG
(*A*) *The Dow Jones Industrial index*					
$r = 1$	1.218	1.218			
$r = 0$	75.968*	77.186*			
					7
Q5			0.089	0.182	
Q10			2.599	6.519	
C. Vector			1.000	−1.007	
(*B*) *The NASDAQ index*					
$r = 1$	1.652	1.652			
$r = 0$	112.184*	113.836*			
					8
Q5			0.149	0.364	
Q10			1.910	5.719	
C. Vector			1.000	−1.010	
(*C*) *The S&P 500 index*					
$r = 1$	1.223	1.223			
$r = 0$	102.923*	104.145*			
					8
Q5			0.173	0.151	
Q10			2.497	3.673	
C. Vector			1.000	−1.005	

Note: The results of testing for cointegration between highs and lows of the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index are reported in Panels A–C. Eigenvalue and trace statistics are given under the columns 'EIGENV' and 'TRACE.' '$r = 0$' corresponds to the null hypothesis of no cointegration and '$r = 1$' corresponds to the hypothesis of one cointegration vector. The no-cointegration null is rejected and the hypothesis of one-cointegration vector is not rejected. '*H*' and '*L*' identify the *Q*-statistics associated with the daily high and daily low equations. All the *Q*-statistics are insignificant. The rows labelled 'C. Vector' give cointegrating vectors with the coefficient of the high normalized to one. 'LAG', the lag parameters used to conduct the test.
*Indicates significance at the 5% level.

The estimated cointegrating vectors with the coefficient of the daily high series H_t normalized to one are also reported in Table 3. According to the estimated cointegrating vectors, there is approximately a one-to-one correspondence between movements in daily high and daily low over time. Recall that the range is defined by $R_t = H_t - L_t$. The stationarity result of the range R_t reported in Table 1 is supportive of the $(1, -1)$ specification of the cointegrating vector. Thus, in the subsequent analyses, we impose the $(1, -1)$ cointegrating restriction in estimating the vector error correction model.[7] The diagnostic *Q*-statistics are all insignificant; indicating the selected lag structures are appropriate.

3.2. Vector error correction model

Given that the daily high and daily low series are cointegrated, a vector error correction model (VECM) is used to examine their long-run and short-run interactions. Imposing the $(1,-1)$ cointegrating vector restriction, the VECM can be written as

$$\Delta \mathbf{X}_t = \mu + \sum_{i=1}^{p} \Gamma_i \Delta \mathbf{X}_{t-i} + \alpha R_{t-1} + \xi D_t + \varepsilon_t \tag{3}$$

The variable $D_t \equiv (d_{2t}, d_{3t}, d_{4t}, d_{5t})'$ containing dummies for Tuesday, Wednesday, Thursday, and Friday is included to allow for the possible day-of-the-week effect. The VECM results are presented in Table 4. The

DOI: 10.1002/ijfe

Table 4. Estimates of the basic vector error correction models

	ΔH		ΔL	
	COEFF	T-STAT	COEFF	T-STAT
(*A*) *The Dow Jones Industrial index*				
ΔH (−1)	−0.201	−6.492	0.550	15.691
ΔH (−2)	−0.256	−6.951	0.279	6.701
ΔH (−3)	−0.151	−3.877	0.259	5.863
ΔH (−4)	−0.099	−2.507	0.219	4.902
ΔH (−5)	−0.046	−1.202	0.215	4.950
ΔH (−6)	0.028	0.804	0.216	5.553
ΔL (−1)	0.388	14.254	−0.232	−7.509
ΔL (−2)	0.143	4.449	−0.354	−9.707
ΔL (−3)	0.156	4.549	−0.236	−6.045
ΔL (−4)	0.126	3.605	−0.183	−4.625
ΔL (−5)	0.013	0.376	−0.251	−6.501
ΔL (−6)	−0.024	−0.789	−0.206	−5.908
μ	0.001	1.896	−0.001	−1.226
R(.)	−0.011	−0.605	0.070	3.474
d_{2t}	0.000	0.243	0.000	−0.612
d_{3t}	0.000	−1.083	−0.001	−2.421
d_{4t}	−0.001	−1.539	−0.001	−1.772
d_{5t}	−0.001	−2.063	−0.001	−1.506
Adjusted R^2	0.082		0.096	
Q5	0.062		0.278	
Q10	2.564		7.367	
(*B*) *The NASDAQ index*				
ΔH (−1)	−0.274	−8.790	0.564	15.019
ΔH (−2)	−0.297	−7.775	0.341	7.437
ΔH (−3)	−0.213	−5.186	0.351	7.112
ΔH (−4)	−0.091	−2.153	0.370	7.238
ΔH (−5)	−0.077	−1.829	0.253	4.975
ΔH (−6)	0.002	0.045	0.229	4.719
ΔH (−7)	−0.006	−0.170	0.196	4.594
ΔL (−1)	0.428	16.530	−0.282	−9.044
ΔL (−2)	0.209	6.437	−0.407	−10.438
ΔL (−3)	0.215	6.057	−0.302	−7.075
ΔL (−4)	0.117	3.155	−0.319	−7.159
ΔL (−5)	0.075	2.002	−0.290	−6.470
ΔL (−6)	0.003	0.095	−0.269	−6.261
ΔL (−7)	0.033	1.021	−0.175	−4.515
μ	0.001	1.637	−0.001	−1.138
R(.)	−0.053	−2.514	0.047	1.851
d_{2t}	0.001	0.928	0.000	0.156
d_{3t}	0.000	0.078	0.000	−0.504
d_{4t}	0.001	1.541	0.001	1.713
d_{5t}	0.000	−0.690	0.001	0.925
Adjusted R^2	0.094		0.079	
Q5	0.174		0.537	
Q10	1.857		9.257	
(*C*) *The S&P 500 index*				
ΔH (−1)	−0.302	−11.980	0.730	24.937
ΔH (−2)	−0.414	−12.471	0.374	9.687
ΔH (−3)	−0.254	−6.906	0.403	9.429

DOI: 10.1002/ijfe

Table 4. (*continued*)

	ΔH		ΔL	
	COEFF	T-STAT	COEFF	T-STAT
ΔH (−4)	−0.208	−5.433	0.365	8.218
ΔH (−5)	−0.122	−3.198	0.334	7.519
ΔH (−6)	−0.074	−2.013	0.324	7.562
ΔH (−7)	−0.067	−2.063	0.211	5.584
ΔL (−1)	0.533	24.699	−0.294	−11.706
ΔL (−2)	0.230	8.141	−0.503	−15.311
ΔL (−3)	0.283	8.912	−0.336	−9.108
ΔL (−4)	0.180	5.371	−0.363	−9.322
ΔL (−5)	0.120	3.543	−0.342	−8.664
ΔL (−6)	0.050	1.535	−0.330	−8.663
ΔL (−7)	0.056	1.924	−0.238	−6.985
μ	0.001	2.621	−0.001	−1.834
R(.)	−0.034	−1.520	0.120	4.589
d_{2t}	0.000	−0.245	0.000	−0.543
d_{3t}	−0.001	−1.322	−0.001	−1.262
d_{4t}	0.000	−0.863	0.000	−0.873
d_{5t}	−0.001	−1.434	−0.001	−1.789
Adjusted R^2	0.165		0.176	
Q5	0.165		0.207	
Q10	2.410		4.229	

Note: The estimates of the vector error correction model (3) are reported. Panels A–C give the results for the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index. Results pertaining to the high and the low equations are reported under the headings 'ΔH' and 'ΔL.' t-statistics are given in parentheses next to the parameter estimates. 'μ' is the constant term. R(.) is the lagged range; which is the error correction term with the (1, −1) coefficient restriction. d_{2t}, d_{3t}, d_{4t}, are d_{5t} are the Tuesday, Wednesday, Thursday, and Friday dummy variables capturing the day-of-the-week effects. The adjusted R^2 statistics are reported in the row labelled 'Adjusted R^2.' Q5 and Q10 give the Q-statistics calculated from the first 5 and 10 sample autocorrelations, respectively. All the Q-statistics are insignificant.

Q-statistics affirm that the selected VECM models adequately capture the data dynamics and the resulting disturbance terms display no statistically significant serial correlation.

Since we do not have a theoretical model underpinning the VECM (3), we do not want to over-interpret the estimation results. Nonetheless, there are a few interesting observations. First, in all three cases, the range variable has a negative coefficient in the daily high equation and a positive coefficient in the daily low equation. An increase in the daily range tends to bring down the next daily high and boost the next daily low and, hence, reduces the next daily range. Thus, the estimated dynamics implies the range variable is regressive and is in accordance with its stationary property. The result is consistent with the cointegration result and indicates the range variable is not an unreasonable proxy for the error correction term. While the range variable is statistically significant in all three daily low equations, it is significant in only one daily high equation—the NASDAQ daily high equation. We do not have a good reason to explain the range is mostly significant in daily low but not daily high equations.

Second, for all the three stock indexes, the significant coefficient estimates of lagged dependent variables are all negative and those of the other lagged variables are positive. For instance, consider the Dow Jones Industrial index daily high equation in Panel A, the coefficient estimates of the lagged daily high differences are negative whereas those of the lagged daily low differences are positive. The negative coefficients suggest regressive behaviour. Higher daily highs tend to drift down to a lower level, and lower daily highs tend to move up to a higher level. On the other hand, the positive coefficients of the lagged daily low differences are indicative of spillover effects. Higher (lower) daily lows lead to higher (lower) daily highs.

DOI: 10.1002/ijfe

Third, the explanatory power of these error correction equations is quite decent for stock price changes. The two S&P 500 equations presented in Panel C have the highest adjusted R^2 statistics of 16.5% and 17.6%. For the other two stock indexes, the adjusted R^2 is between 7.9% and 9.6%.

The estimation results indicate day-of-the-week effects in daily high and daily low data are quite weak. Most of the day-of-the-week dummy variables are not statistically significant. For the few that are significant, the (absolute) size of the estimates is quite small. When the variable D_t is omitted from (3), the other estimates are essentially the same (in terms of both magnitude and statistical significance) and the adjusted R^2 is reduced by less than 0.1% in most cases. Indeed, for all the three cases, the Bayesian information criterion selects the specification without the day-of-the-week dummy variables, which passes diagnostic tests with essentially the same Q-statistics. Thus, for brevity, the day-of-the-week effect is not considered in the subsequent analyses.

A remark on modelling the range is in order. The VECM (3) implies that the use of the historical range data to model the range dynamics may not be most efficient. Multiply both sides of (3) by the vector $(1, -1)'$, we have

$$\Delta R_t = c + \sum_{i=1}^{p} (a_i \Delta H_{t-i} - b_i \Delta L_{t-i}) + w R_{t-1} + u_t \tag{4}$$

where c, a_i, b_i, w and u_t are functions of the coefficients in (3). Only when the difference of the rows in each Γ_i is a constant vector, we have $a_i = b_i$ and ΔR_{t-i} under the summation sign on the right-hand side of (4). Thus, under the VECM specification, a proper specification of the range R_t requires information on the high and the low, and beyond the history of R_t itself.

4. AUGMENTED MODELS

4.1. Additional price variables

Equation (3) uses only histories of highs and lows as explanatory variables. Since the open and close are realizations from the same price series, they contain useful information about the evolution of the high and the low. Consider, say, changes in the daily closing price and the daily high, ΔC_{t-1} and ΔH_t. Because the close and the high are recorded at different times of the day, the information arrived between H_{t-1} and C_{t-1} is contained in ΔC_{t-1} but not available in ΔH_{t-1}. ΔC_{t-1} does not contain extra information when $H_{t-1} = C_{t-1}$. Thus, adding data on opens and closes would enhance the performance of (3). The role of other price variables in explaining ΔH_t and ΔL_t is examined using the augmented model:

$$\Delta \mathbf{X}_t = \mu + \sum_{i=1}^{p} \Gamma_i \Delta \mathbf{X}_{t-i} + \alpha R_{t-1} + \sum_{i=1}^{q} \Lambda_i \Delta Y_{t-i} + \sum_{i=1}^{r} \theta_i CO_{t-i} + \varepsilon_t \tag{5}$$

where ΔY_{t-i} is a vector containing ΔO_{t-i} and ΔC_{t-i}, CO_{t-i} is given by $C_{t-i} - O_{t-i}$, and Λ_i and θ_i are the corresponding coefficient matrix and vector. The results of fitting (5) to the data are presented in Table 5. The lag parameters q and r are chosen based on the significance of Λ_i and θ_i. The significant coefficients of ΔY_{t-i} and CO_{t-i} in these equations are all positive; indicating that increases in inter-day movements in opens and closes and in intraday open-to-close spreads imply gains in the high and the low. The local price momentum (information) captured by these additional price variables helps explain variations in both highs and lows.

The inclusion of these additional price variables has some systematic impacts on the original VECM coefficient estimates. The coefficient estimates of the lagged dependent variables become more negative, and those of the other variables shrink and turn negative in some cases. For instance, in the case of the Dow Jones Industrial daily low equation presented in Panel A, the coefficient estimates of the first few lagged changes in lows display a larger negative impact than those in Table 4. The effect of the lagged changes in highs is smaller; the coefficient estimates of the first two lags are, in fact, significantly negative.

DOI: 10.1002/ijfe

Table 5. Estimates of the vector error correction models with additional price variables

	ΔH		ΔL	
	COEFF	T-STAT	COEFF	T-STAT
(*A*) *The Dow Jones Industrial index*				
ΔH (−1)	−0.786	−24.727	−0.168	−4.743
ΔH (−2)	−0.560	−14.247	−0.143	−3.256
ΔH (−3)	−0.415	−9.996	−0.070	−1.510
ΔH (−4)	−0.212	−5.318	0.091	2.053
ΔH (−5)	−0.089	−2.706	0.167	4.526
ΔH (−6)	0.000	−0.015	0.179	5.671
ΔL (−1)	−0.022	−0.786	−0.742	−24.214
ΔL (−2)	−0.008	−0.226	−0.582	−15.628
ΔL (−3)	0.023	0.647	−0.406	−10.441
ΔL (−4)	0.070	2.111	−0.244	−6.588
ΔL (−5)	0.048	1.702	−0.205	−6.450
ΔL (−6)	0.003	0.136	−0.172	−6.075
μ	0.001	2.157	−0.001	−2.782
ΔO (−1)	0.397	4.400	0.363	3.605
ΔO (−2)	0.114	1.539	0.165	2.000
ΔO (−3)	0.107	1.923	0.133	2.147
ΔO (−4)	0.078	2.793	0.080	2.563
ΔC (−1)	0.396	3.867	0.619	5.408
ΔC (−2)	0.219	2.447	0.457	4.584
ΔC (−3)	0.318	4.442	0.406	5.091
ΔC (−4)	0.141	2.790	0.165	2.913
CO (−1)	−0.457	−4.464	−0.406	−3.553
R(.)	−0.021	−1.442	0.057	3.489
Adjusted R^2	0.376		0.404	
Q5	1.736		3.205	
Q10	5.385		5.356	
(*B*) *The NASDAQ index*				
ΔH (−1)	−0.764	−22.075	−0.066	−1.637
ΔH (−2)	−0.487	−11.280	0.056	1.106
ΔH (−3)	−0.367	−8.873	0.136	2.803
ΔH (−4)	−0.149	−4.477	0.291	7.450
ΔH (−5)	−0.042	−1.299	0.291	7.600
ΔH (−6)	−0.009	−0.283	0.212	5.776
ΔH (−7)	−0.017	−0.610	0.183	5.662
ΔL (−1)	0.065	2.262	−0.751	−22.285
ΔL (−2)	0.152	4.318	−0.519	−12.550
ΔL (−3)	0.156	4.532	−0.402	−9.987
ΔL (−4)	0.152	5.285	−0.279	−8.249
ΔL (−5)	0.063	2.173	−0.298	−8.817
ΔL (−6)	0.031	1.135	−0.229	−7.081
ΔL (−7)	0.018	0.711	−0.194	−6.658
μ	0.001	3.916	0.000	−1.224
ΔO (−1)	0.308	5.935	0.437	7.201
ΔO (−2)	0.142	3.129	0.242	4.542
ΔO (−3)	0.098	3.107	0.159	4.280
ΔC (−1)	0.292	5.456	0.309	4.926
ΔC (−2)	0.148	3.214	0.210	3.886
ΔC (−3)	0.101	3.093	0.127	3.344
CO (−1)	−0.594	−11.072	−0.816	−12.970
R(.)	−0.037	−2.258	0.069	3.628

DOI: 10.1002/ijfe

Table 5. (*continued*)

	ΔH		ΔL	
	COEFF	T-STAT	COEFF	T-STAT
Adjusted R^2	0.457		0.477	
Q5	1.994		0.918	
Q10	9.216		4.089	
(*C*) *The S&P 500 index*				
ΔH (−1)	−0.880	−31.033	0.035	1.049
ΔH (−2)	−0.666	−18.452	0.031	0.740
ΔH (−3)	−0.500	−13.582	0.119	2.781
ΔH (−4)	−0.259	−8.328	0.321	8.874
ΔH (−5)	−0.135	−4.471	0.321	9.154
ΔH (−6)	−0.121	−4.167	0.273	8.076
ΔH (−7)	−0.087	−3.390	0.187	6.291
ΔL (−1)	0.079	3.314	−0.842	−30.423
ΔL (−2)	0.102	3.439	−0.689	−20.020
ΔL (−3)	0.135	4.415	−0.507	−14.317
ΔL (−4)	0.192	7.166	−0.342	−10.990
ΔL (−5)	0.144	5.363	−0.320	−10.251
ΔL (−6)	0.098	3.801	−0.276	−9.204
ΔL (−7)	0.088	3.786	−0.204	−7.597
μ	0.001	3.412	−0.001	−4.130
ΔO (−1)	0.369	4.953	0.257	2.967
ΔO (−2)	0.305	5.347	0.276	4.161
ΔO (−3)	0.220	7.612	0.224	6.680
ΔC (−1)	0.355	4.128	0.597	5.984
ΔC (−2)	0.232	3.303	0.508	6.243
ΔC (−3)	0.107	2.187	0.247	4.356
CO (−1)	−0.501	−5.759	−0.415	−4.106
R(.)	−0.048	−2.729	0.104	5.033
Adjusted R^2	0.481		0.489	
Q5	3.416	[0.636]	5.123	[0.401]
Q10	5.723	[0.838]	9.025	[0.530]

Note: The estimates of the augmented vector error correction model (5) for the high and the low equations are reported. Panels A–C give the results for the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index. ΔO(.), ΔC(.), and CO(.) are the extra price variables added to the basic VECM (3). See also the Note to Table 4.

However, the effects of these additional price variables on the range variable's coefficient estimates are not similar across the three US stock indexes. For instance, compared with Table 4, the estimated range effect in Table 5 is smaller for the Dow Jones Industrial and the S&P 500 daily low equations but is larger in the case of the NASDAQ daily low equation. For daily high equations, the range effect is mitigated in the case of the NASDAQ index but is stronger and becomes significant for the S&P 500 index.

The most noticeable change is the adjusted R^2 statistic. The NASDAQ daily low equation experiences the largest improvement. The adjusted R^2 statistic of the augmented equation (47.67%) is six times of the original error correction equation (7.87%). The smallest increase is given by the S&P 500 daily low equation; the adjusted R^2 improves from 17.59% to 48.91%. The additional price variables do not qualitatively deteriorate the diagnostic Q-statistics, which still indicate the estimated residuals are well behaved. Thus, the explanatory power is enhanced without scarifying the modelling quality.

DOI: 10.1002/ijfe

4.2. Trading volume

Trading volume is an exogenous variable quite commonly considered by studies of financial price dynamics. Intuitively, trading volume is a relevant variable since prices are determined by the interplay of demand and supply. Indeed, there is a rich literature that covers the theory and empirics of interactions between returns and trading volume.[8] We investigate the effect trading volume has on highs and lows using the regression:

$$\Delta \mathbf{X}_t = \mu + \sum_{i=1}^{p} \Gamma_i \Delta \mathbf{X}_{t-i} + \alpha R_{t-1} + \sum_{i=0}^{s} \delta_i V_{t-i} + \varepsilon_t \tag{6}$$

Following a common practice in extant literature, we include the contemporaneous trading volume and set the lag parameter s to 1. The estimation results are given in Table 6.

Table 6. Estimates of the vector error correction models with trading volume

	ΔH		ΔL	
	COEFF	T-STAT	COEFF	T-STAT
(*A*) *The Dow Jones Industrial index*				
ΔH (−1)	−0.185	−5.670	0.584	15.718
ΔH (−2)	−0.252	−6.501	0.329	7.456
ΔH (−3)	−0.145	−3.540	0.315	6.789
ΔH (−4)	−0.098	−2.393	0.284	6.107
ΔH (−5)	−0.057	−1.428	0.283	6.244
ΔH (−6)	0.011	0.295	0.283	6.851
ΔL (−1)	0.377	13.101	−0.263	−8.055
ΔL (−2)	0.143	4.217	−0.403	−10.421
ΔL (−3)	0.150	4.138	−0.290	−7.011
ΔL (−4)	0.128	3.490	−0.246	−5.908
ΔL (−5)	0.026	0.736	−0.318	−7.788
ΔL (−6)	−0.010	−0.310	−0.272	−7.250
μ	0.001	1.904	−0.003	−4.829
V	0.005	7.415	−0.004	−5.685
V (−1)	−0.004	−4.861	0.000	−0.159
R(.)	−0.027	−1.280	0.131	5.525
Adjusted R^2	0.095		0.104	
Q5	0.230		0.196	
Q10	2.890		7.351	
(*B*) *The NASDAQ index*				
ΔH (−1)	−0.264	−8.233	0.581	14.986
ΔH (−2)	−0.281	−7.072	0.369	7.679
ΔH (−3)	−0.200	−4.685	0.382	7.376
ΔH (−4)	−0.091	−2.058	0.402	7.516
ΔH (−5)	−0.088	−2.001	0.289	5.438
ΔH (−6)	−0.020	−0.461	0.276	5.391
ΔH (−7)	−0.025	−0.667	0.246	5.374
ΔL (−1)	0.422	15.901	−0.298	−9.273
ΔL (−2)	0.196	5.813	−0.431	−10.549
ΔL (−3)	0.202	5.421	−0.329	−7.317
ΔL (−4)	0.115	2.980	−0.348	−7.438
ΔL (−5)	0.083	2.145	−0.322	−6.843
ΔL (−6)	0.019	0.505	−0.311	−6.833
ΔL (−7)	0.050	1.443	−0.220	−5.287
μ	0.001	3.420	−0.001	−2.091

DOI: 10.1002/ijfe

Table 6. (*continued*)

	ΔH		ΔL	
	COEFF	T-STAT	COEFF	T-STAT
V	0.009	7.923	−0.006	−4.035
V (−1)	−0.007	−6.049	0.002	1.592
R(.)	−0.071	−2.957	0.089	3.048
Adjusted R^2	0.108		0.083	
Q5	0.198		0.467	
Q10	2.670		6.804	
(*C*) *The S&P 500 index*				
ΔH (−1)	−0.299	−10.163	0.731	21.456
ΔH (−2)	−0.403	−10.258	0.386	8.476
ΔH (−3)	−0.244	−5.608	0.429	8.520
ΔH (−4)	−0.195	−4.342	0.420	8.069
ΔH (−5)	−0.117	−2.606	0.398	7.630
ΔH (−6)	−0.067	−1.538	0.405	8.032
ΔH (−7)	−0.054	−1.383	0.288	6.408
ΔL (−1)	0.517	20.461	−0.311	−10.610
ΔL (−2)	0.220	6.589	−0.526	−13.597
ΔL (−3)	0.270	7.182	−0.360	−8.256
ΔL (−4)	0.162	4.101	−0.426	−9.332
ΔL (−5)	0.108	2.696	−0.415	−8.909
ΔL (−6)	0.039	0.999	−0.410	−9.056
ΔL (−7)	0.052	1.468	−0.307	−7.464
μ	0.001	1.314	−0.003	−5.470
V	0.002	3.327	−0.008	−9.947
V (−1)	−0.003	−3.883	0.004	4.191
R(.)	−0.014	−0.497	0.212	6.467
Adjusted R^2	0.167		0.199	
Q5	0.264		0.280	
Q10	3.037		3.837	

Note: The estimates of the augmented vector error correction model (6) for the high and the low equations are reported. Panels A–C give the results for the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index. V and V (−1) are the contemporaneous and lagged trading volume variables added to the basic VECM (3). See also the Note to Table 4.

The contemporaneous trading volume is positively correlated with the change in the daily high. The lagged trading volume, on the other hand, has a negative impact. The results are quite different for the daily low equation. The contemporaneous trading volume is found to be negatively correlated with changes in daily lows. The lagged trading volume, on the other hand, has a significant positive effect for the NASDAQ and S&P 500 indexes and an insignificant effect for the Dow Jones Industrial index.

When we combine the effects on the daily high and low equations, a high level of contemporaneous trading volume implies a large range value (because of an increase in the high and a reduction in the low). Since the range is a proxy of volatility, the result is in accordance with the assertion that a high level of trading volume is associated with a high level of volatility. The lagged trading volume, on the other hand, is negatively related to the range—a result that is comparable to its negative effect on volatility reported in the literature. Thus, the estimated trading volume effect is broadly consistent with the notion of joint dependence of returns and volume on a common latent variable and with empirical findings on the interaction between returns and volatility.

The presence of trading volume does not materially change the estimates of the original VECM model. The coefficient estimates of the lagged changes in Table 6 have signs and magnitudes that are quite

DOI: 10.1002/ijfe

comparable to those in Table 4. Similar to the additional price variables considered in Table 5, the trading volume does not have a systematic effect on the range coefficient estimates. Specifically, the Dow Jones Industrial high and low equations exhibit range effects that are larger than those in Table 4, Panel A. On other hand, the presence of trading volume reduces the range effects for the S&P 500 equations and yields mixed impacts for the NASDAQ equations. The diagnostic Q-statistics reported in Table 6 are all insignificant. The incremental explanatory power of trading volume is small relative to the price variables considered in the previous subsection. The inclusion of trading volume, in general, strenghtens the value of the adjusted R^2 statistic by 1–2%. The additional price variables in the previous subsection, on the other hand, boost the statistic by over 30%.

4.3. The combined model

The combined effects of the additional price variables and trading volume are examined using

$$\Delta \mathbf{X}_t = \mu + \sum_{i=1}^{p} \Gamma_i \Delta \mathbf{X}_{t-i} + \alpha R_{t-1} + \sum_{i=1}^{q} \Lambda_i \Delta Y_{t-i} + \sum_{i=1}^{r} \theta_i CO_{t-i} + \sum_{i=0}^{s} \delta_i V_{t-i} + \varepsilon_t \tag{7}$$

The results are presented in Table 7. In a nutshell, the coefficient estimates of the price variables are quite similar to those in Table 5, the trading volume effects are comparable to those reported in Table 6, the adjusted R^2 statistics are marginally higher than those in Table 5, and the Q-statistics are good.

The explanatory power of V_t relative to ΔY_t and CO_t is in accordance with the notion that trading volume is secondary in importance while price is the most important piece of information. In technical analysis, trading volume patterns are usually used to confirm price patterns but not used as the primary indicator. Overall, (7) offers a promising specification of the high and low dynamics. It explains close to 50% of the variations in changes in highs and lows, as indicated by the adjusted R^2 statistics.

Table 7. Estimates of the vector error correction models with additional price variables and trading volume

	ΔH		ΔL	
	COEFF	T-STAT	COEFF	T-STAT
(*A*) *The Dow Jones Industrial index*				
$\Delta H\ (-1)$	−0.780	−23.692	−0.137	−3.721
$\Delta H\ (-2)$	−0.565	−13.936	−0.098	−2.152
$\Delta H\ (-3)$	−0.418	−9.818	−0.027	−0.566
$\Delta H\ (-4)$	−0.226	−5.554	0.138	3.026
$\Delta H\ (-5)$	−0.116	−3.390	0.221	5.776
$\Delta H\ (-6)$	−0.028	−0.944	0.236	7.039
$\Delta L\ (-1)$	−0.031	−1.103	−0.763	−24.077
$\Delta L\ (-2)$	−0.005	−0.154	−0.617	−16.055
$\Delta L\ (-3)$	0.024	0.670	−0.444	−11.018
$\Delta L\ (-4)$	0.084	2.446	−0.289	−7.529
$\Delta L\ (-5)$	0.074	2.463	−0.259	−7.720
$\Delta L\ (-6)$	0.029	1.050	−0.227	−7.449
μ	0.001	3.228	−0.002	−4.908
V	0.005	9.275	−0.004	−6.720
$V\ (-1)$	−0.003	−5.155	0.000	0.684
$\Delta O\ (-1)$	0.392	4.392	0.372	3.712
$\Delta O\ (-2)$	0.105	1.437	0.171	2.095
$\Delta O\ (-3)$	0.105	1.914	0.137	2.221
$\Delta O\ (-4)$	0.081	2.938	0.078	2.524
$\Delta C\ (-1)$	0.418	4.124	0.595	5.228
$\Delta C\ (-2)$	0.232	2.621	0.439	4.431
$\Delta C\ (-3)$	0.331	4.678	0.394	4.961
$\Delta C\ (-4)$	0.143	2.860	0.160	2.843
$CO\ (-1)$	−0.437	−4.321	−0.426	−3.751
$R(.)$	−0.045	−2.653	0.107	5.569

DOI: 10.1002/ijfe

Table 7. (*continued*)

	ΔH		ΔL	
	COEFF	T-STAT	COEFF	T-STAT
Adjusted R^2	0.390		0.412	
Q5	1.800		2.148	
Q10	4.402		5.065	
(*B*) *The NASDAQ index*				
ΔH (−1)	−0.776	−22.387	−0.054	−1.310
ΔH (−2)	−0.493	−11.283	0.070	1.345
ΔH (−3)	−0.364	−8.616	0.147	2.929
ΔH (−4)	−0.160	−4.647	0.307	7.496
ΔH (−5)	−0.069	−2.037	0.313	7.793
ΔH (−6)	−0.049	−1.495	0.239	6.174
ΔH (−7)	−0.054	−1.846	0.209	6.042
ΔL (−1)	0.056	1.944	−0.753	−22.025
ΔL (−2)	0.143	4.038	−0.523	−12.405
ΔL (−3)	0.154	4.417	−0.410	−9.918
ΔL (−4)	0.162	5.434	−0.293	−8.259
ΔL (−5)	0.085	2.846	−0.317	−8.925
ΔL (−6)	0.065	2.241	−0.253	−7.366
ΔL (−7)	0.050	1.911	−0.218	−6.946
μ	0.002	5.106	−0.001	−2.019
V	0.011	11.731	−0.004	−3.886
V (−1)	−0.008	−8.067	0.002	2.007
ΔO (−1)	0.325	6.367	0.426	7.004
ΔO (−2)	0.149	3.336	0.236	4.423
ΔO (−3)	0.094	3.008	0.159	4.285
ΔC (−1)	0.296	5.601	0.313	4.978
ΔC (−2)	0.155	3.416	0.210	3.887
ΔC (−3)	0.101	3.163	0.128	3.366
CO (−1)	−0.605	−11.405	−0.805	−12.762
R(.)	−0.069	−3.727	0.092	4.181
Adjusted R^2	0.476		0.479	
Q5	1.676		0.805	
Q10	7.662		3.166	
(*C*) *The S&P 500 index*				
ΔH (−1)	−0.891	−27.577	0.033	0.895
ΔH (−2)	−0.686	−16.449	0.023	0.487
ΔH (−3)	−0.493	−11.546	0.147	2.971
ΔH (−4)	−0.258	−7.075	0.358	8.507
ΔH (−5)	−0.133	−3.739	0.383	9.332
ΔH (−6)	−0.120	−3.514	0.344	8.668
ΔH (−7)	−0.086	−2.803	0.251	7.086
ΔL (−1)	0.063	2.315	−0.847	−26.885
ΔL (−2)	0.081	2.379	−0.711	−18.040
ΔL (−3)	0.121	3.426	−0.527	−12.959
ΔL (−4)	0.187	5.969	−0.393	−10.843
ΔL (−5)	0.139	4.390	−0.382	−10.418
ΔL (−6)	0.096	3.112	−0.345	−9.714
ΔL (−7)	0.088	3.147	−0.267	−8.260
μ	0.001	2.499	−0.002	−6.120
V	0.004	6.808	−0.007	−9.927
V (−1)	−0.004	−6.413	0.003	3.769
ΔO (−1)	0.374	4.741	0.273	2.997

DOI: 10.1002/ijfe

Table 7. (*continued*)

	ΔH		ΔL	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta O\ (-2)$	0.320	5.229	0.310	4.384
$\Delta O\ (-3)$	0.222	6.923	0.231	6.222
$\Delta C\ (-1)$	0.384	4.269	0.570	5.480
$\Delta C\ (-2)$	0.261	3.561	0.503	5.924
$\Delta C\ (-3)$	0.121	2.372	0.221	3.741
$CO\ (-1)$	−0.477	−5.220	−0.437	−4.141
$R(.)$	−0.041	−1.853	0.180	7.006
Adjusted R^2	0.485		0.507	
Q5	3.383		2.073	
Q10	5.484		8.160	

Note: The estimates of the augmented vector error correction model (7) for the high and the low are reported. Panels A–C give the results for the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index. $\Delta O(.)$, $\Delta C(.)$, $CO(.)$, and $V(.)$ are the extra explanatory variables added to the basic VECM (3). See also the Note to Table 4.

4.4. Impulse responses

In this subsection, we employ the generalized impulse response technique (Pesaran and Shin, 1998) to examine the effects of shocks to the daily high and daily low under different model specifications. Unlike the usual approach based on Cholesky decomposition and orthogonalized shocks, the Pesaran–Shin approach incorporates correlation between shocks and yields unique impulse response functions that are invariant to the ordering of variables. Only in the limiting case of a diagonal variance matrix of the error vector do the traditional and the generalized approaches coincide.

Let the error vector ε_t has a zero mean and a variance $\Sigma = (\sigma_{ij})$. The generalized impulse response of $\mathbf{X}_{t+h}$ with respect to a unit shock to the jth variable ($j = 1$ for a shock to the high and $j = 2$ for a shock to the low) at time t is given by

$$B_h \sum_j e/\sigma_{jj}, \quad h = 0, 1, 2, \ldots \tag{8}$$

where $B_h = \gamma_1 B_{h-1} + \gamma_2 B_{h-2} + \cdots + \gamma_p B_{h-p} + \gamma_{p+1} B_{h-p-1}$, $h = 1, 2, \ldots, B_0 = I$, and $B_h = 0$ for $h < 0$. Note that the matrices $\{B_h, t = 1, 2, \ldots\}$ constitute the coefficient matrices of the (infinite order) moving-average representation of $\mathbf{X}_t$. The term e_j is a selection vector with unity as its jth element and zeros elsewhere. It is shown that (8) is valid for a system of cointegrated variables. See Pesaran and Shin (1998) for a more detailed discussion.

The generalized impulse responses of $\Delta\mathbf{X}_t$ to normalized unit shocks calculated from models (3), (5)–(7) are summarized in Figure 1. The impulse response patterns are different across these models but these patterns are quite similar among the three stock indexes. In general, the effects of the shocks on changes in highs and lows are short-lived; a typical result reported for financial price returns. For the basic VECM model (3) and one-day lagged responses, innovations in daily highs have a larger impact on daily lows than they do on daily highs.[9] On the other hand, innovations in daily lows have a larger impact on daily highs than on daily lows. All the one-day lagged responses are positive, and the effect of the shock dies off pretty quickly after the first day.

The responses to these shocks change quite substantially in the presence of data on openings and closings. In contrast to the basic VECM model, one-day lagged responses to shocks are negative for all the three stock indexes under specification (5). The magnitude of the first day responses is larger than the one

DOI: 10.1002/ijfe

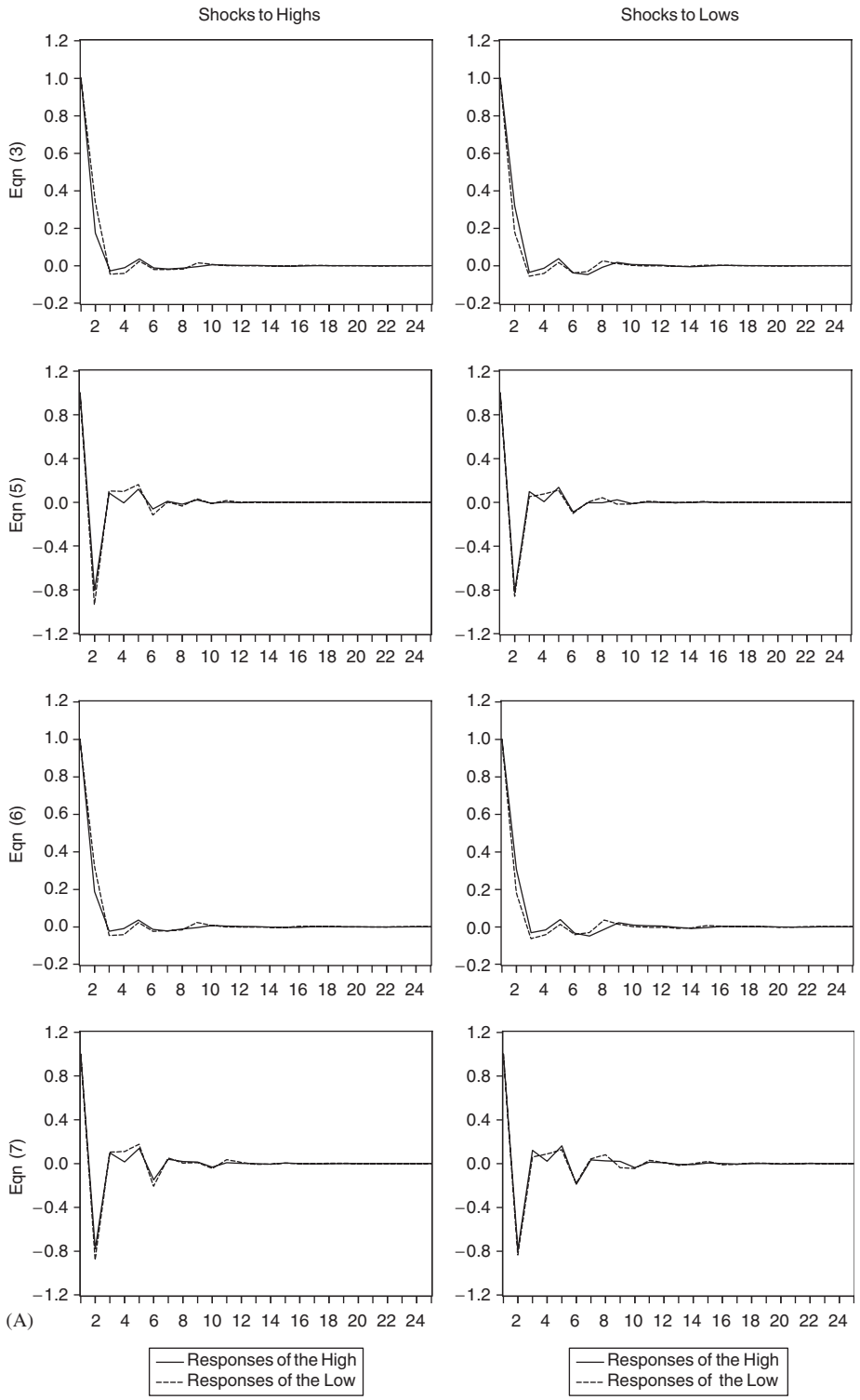

Figure 1. Generalized impulse responses. (A) The Dow Jones Industrial index, (B) the NASDAQ index, (C) the S&P 500 index.

DOI: 10.1002/ijfe

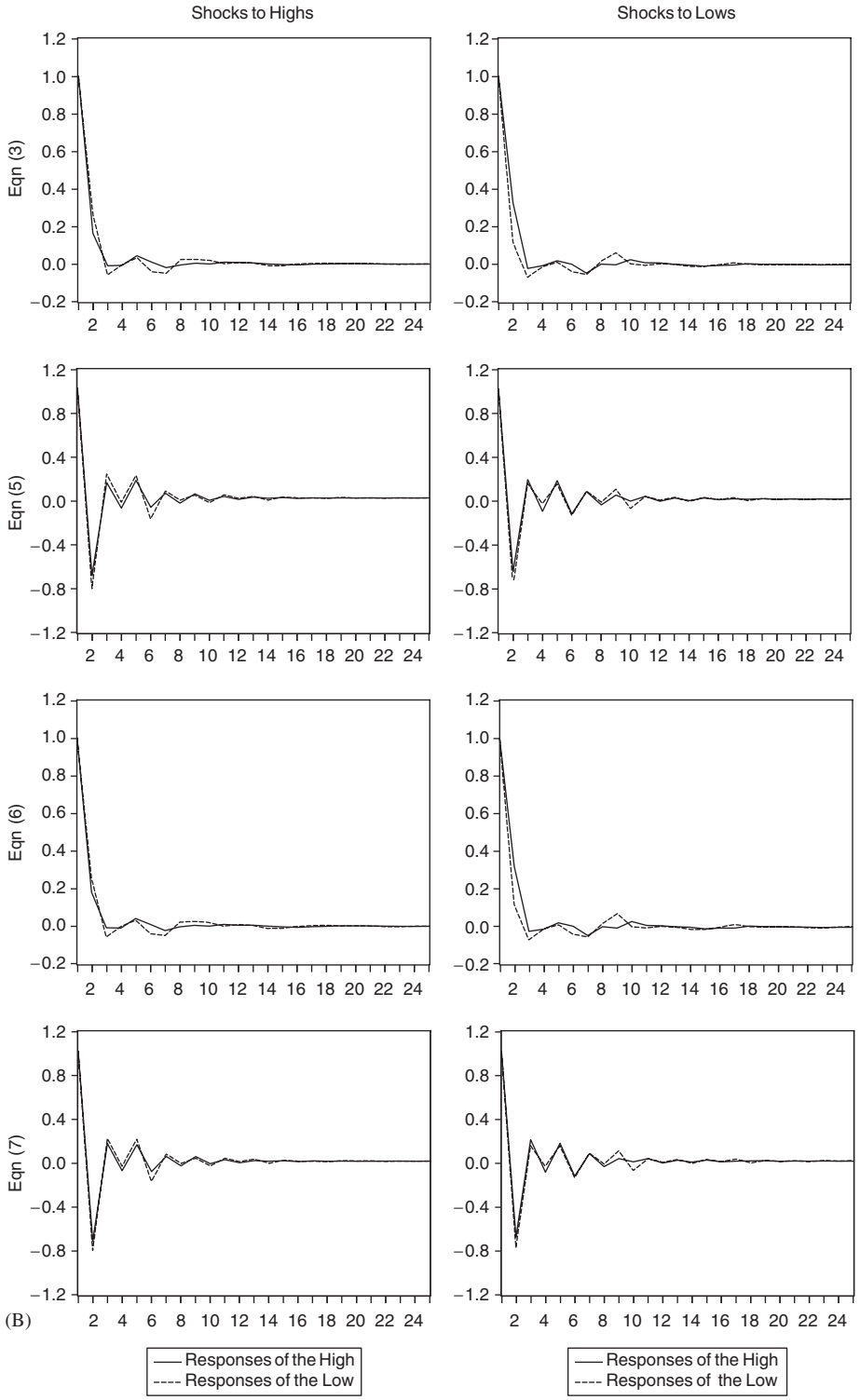

Figure 1—Continued.

DOI: 10.1002/ijfe

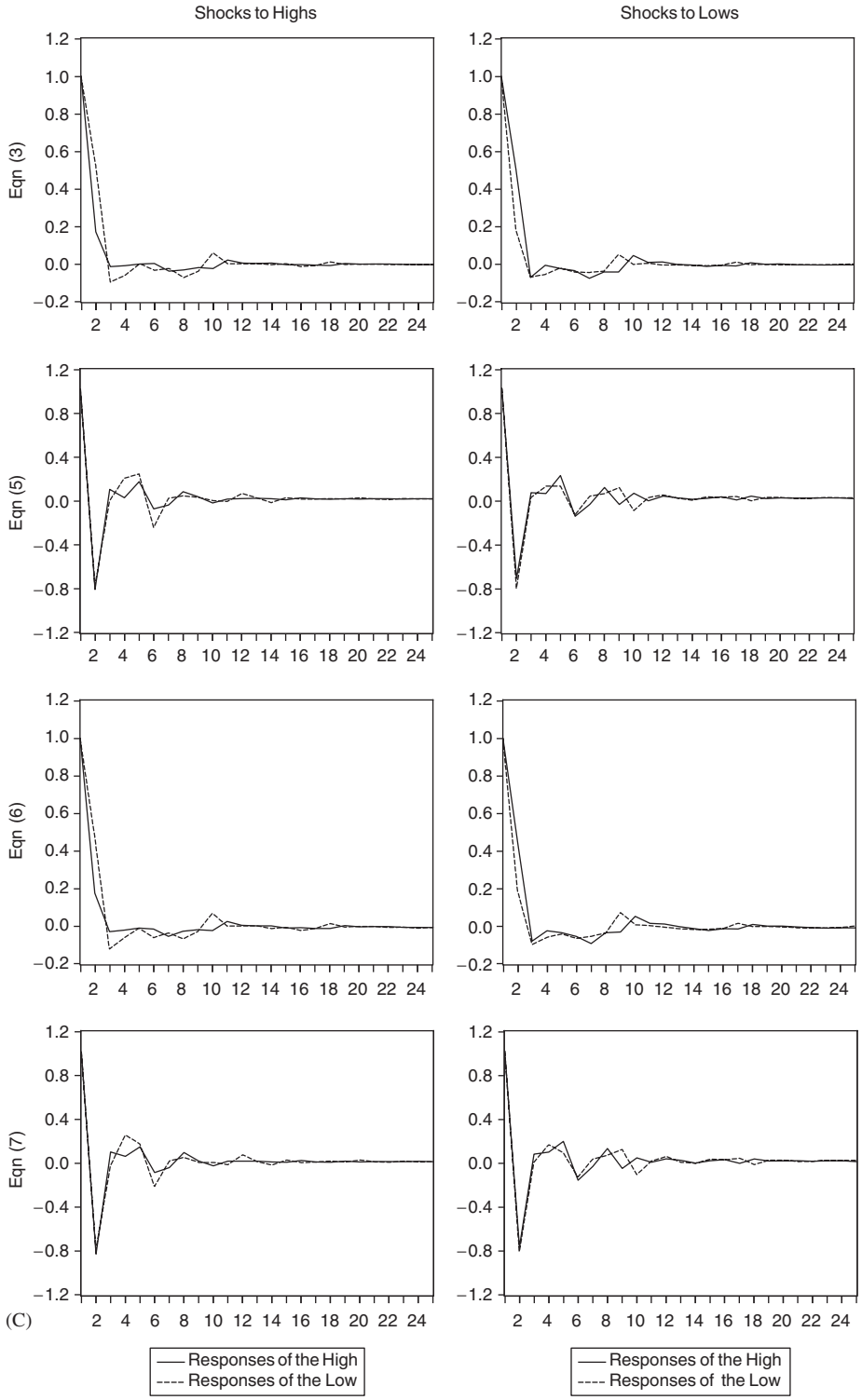

Figure 1—Continued.

DOI: 10.1002/ijfe

from the basic VECM model. While the impulse responses drop off quite fast, their absolute magnitudes are usually larger than the ones from (3).

The trading volume does not appear to have a substantial impact on the impulse response patterns for the three stock indexes. The impulse responses computed from (6) are very similar to those from (3). The combined model (7), as expected, generates impulse responses comparable to those obtained from (5). To summarize, the general impulse response analysis corroborates the analyses conducted in the previous subsections—the open and the close have significant information about the dynamics of the high and the low and their information is richer than that contained in trading volume data.

5. CONCLUDING REMARKS

Motivated by the intuition that daily highs and lows of stock indexes in the US do not drift apart over time, we constructed an empirical model of these two variables based on the cointegration concept. Our empirical results show that daily highs and daily lows of three main US stock price indexes are cointegrated. The difference of the high and the low, which is the price range examined in the literature, is stationary and can be interpreted as the error correction term of the cointegration system comprising highs and lows.

Data on openings, closings, and trading volume are found to offer incremental explanatory power for highs and lows in the VECM framework. The incremental explanatory power of openings and closings is considerably higher than that of trading volume. With all these variables, the augmented VECM models explain 40–50% of the variations in daily highs and lows. The generalized impulse response analysis reveals that the responses of daily highs and daily lows to their shocks depend on whether data on openings, closing, and trading volume are included in the analysis.

The perspective of the current exercise is different from some recent studies that focus on price range dynamics and the ability of price ranges to capture volatility. The current exercise is on modelling the high and the low, which are the constituting elements of the price range. The cointegration result implies that using only the history of the range to model range dynamics does not constitute a complete strategy. A proper specification of the range should also include information on highs and lows. Also, while price ranges can be constructed from highs and lows, it is rather difficult, if not impossible, to recover highs and lows from data on price ranges. Thus, a model of highs and lows is complementary to extant studies on modelling ranges.

The exploratory analysis conducted here indicates that the proposed model has good explanatory power. While we are not claiming the superiority of the empirical high–low model, the results do bear some implications for studying stock price dynamics. For instance, in specifying a GARCH-type specification of stock return behaviour, the range variable derived from the empirical high–low model can be used to model conditional volatility. The use of range may improve the performance of GARCH-type models. Further, range is an efficient estimator of volatility. The empirical model offers a reasonable alternative to generate volatility forecasts that are crucial inputs for options pricing and risk management.[10] In general, the empirical high–low model should complement studies in which (conditional) volatility plays a significant role. Further research, which is beyond the scope of the current exercise, on the implications of the proposed model for pricing exotic options and for evaluating technical trading methods that involve high and low variables is warranted.

ACKNOWLEDGEMENTS

We thank editor Mark Taylor, an anonymous referee, and Giorgio Valente for their helpful comments and suggestions. The financial support of Faculty Research Funds of the University of California, Santa Cruz is gratefully acknowledged.

NOTES

1. Modifications and variations of the Parkinson result are provided by, for example, Beckers (1983), Garman and Klass (1980), Kunitomo (1992), Rogers and Satchell (1991), and Yang and Zhang (2000).

DOI: 10.1002/ijfe

2. See, for example, Alizadeh *et al*. (2002), Brandt and Diebold (2003), Brunetti and Lildholdt (2005), Chou (2005), Engle and Gallo (2003), Fernandes *et al*. (2005), and Gallant *et al*. (1999).
3. In a related literature, the range is used to determine the persistence (strength of memory) of data. See, for example, Hurst (1951), Lo (1991), and Cheung (1993).
4. For example, a knock-out option will expire and become worthless when the price reaches a pre-specified level. A lookback option, on the other hand, offers the retrospective right to exercise the contract at the lowest price (for a call, or the highest price for a put) during the period stipulated in the contract before its expiration.
5. See, for example, Edwards and Magee (1997) for some popular technical trading techniques. The popularity of trading rules in financial markets is documented in, for example, Cheung and Wong (2000), Cheung and Chinn (2001), and Taylor and Allen (1992). Lo *et al*. (2000) provides an extensive analysis of technical trading.
6. See Elliott *et al*. (1996) and Cheung and Lai (1995) for a detailed description of the testing procedure and the related finite sample critical values.
7. The results pertaining to models without the $(1, -1)$ restriction are very similar to those reported in the text. These results are available upon request.
8. See Karpoff (1987) for a detailed review of early studies on the topic. A recent and extensive study is provided by Lo and Wang (in press).
9. Number '2' on the horizontal axis corresponds to the one-day lagged response to the initial unit shock.
10. Indeed, in a companion study (Cheung *et al*., 2005), it is showed that the range forecasts generated from the VECM specification are better than those from, say, ARMA specifications of the range variable.

REFERENCES

Alizadeh S, Brandt MW, Diebold FX. 2002. Range-based estimation of stochastic volatility models. *Journal of Finance* **57**: 1047–1092.

Beckers S. 1983. Variance of security price returns based on high, low, and closing prices. *Journal of Business* **56**: 97–112.

Brandt MW, Diebold FX. 2003. A no-arbitrage approach to range-based estimation of return covariances and correlations. PIER Working Paper 03-013.

Brunetti C, Lildholdt PM. 2005. Relative efficiency of return- and range-based volatility estimators. *Manuscript*, Johns Hopkins University.

Cheung Y-W. 1993. Long memory in foreign exchange rates. *Journal of Business & Economic Statistics* **11**: 93–102.

Cheung Y-W, Cheung YL, Wan A. 2005. A high-low model for forecasting daily stock price ranges. *Manuscript*, UCSC.

Cheung Y-W, Chinn M. 2001. Currency traders and exchange rate dynamics: a survey of the US market. *Journal of International Money and Finance* **20**: 439–471.

Cheung Y-W, Lai KS. 1995. Lag order and critical values of a modified Dickey–Fuller test. *Oxford Bulletin of Economics and Statistics* **57**: 411–419.

Cheung Y-W, Wong CYP. 2000. A survey of market practitioners' views on exchange rate dynamics. *Journal of International Economics* **51**: 401–419.

Chou R. 2005. Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. *Journal of Money, Credit, and Banking* **37**: 561–582.

Edwards RD, Magee J. 1997. *Technical Analysis of Stock Trends* (7th edn). Amacom: New York.

Elliott G, Rothenberg TJ, Stock JH. 1996. Efficient tests for an autoregressive unit root. *Econometrica* **64**: 813–836.

Engle RF, Gallo GM. 2003. A multiple indicators model for volatility using intra-daily data. NBER Working Paper 10117.

Fernandes M, de Sá Mota B, Rocha G. 2005. A multivariate conditional autoregressive range model. *Economics Letters* **86**: 435–440.

Gallant AR, Hsu CT, Tauchen GE. 1999. Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Review of Economics and Statistics* **81**: 617–631.

Garman MB, Klass MJ. 1980. On the estimation of price volatility from historical data. *Journal of Business* **53**: 67–78.

Hurst HE. 1951. Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers* **116**: 770–799.

Johansen S. 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* **59**: 1551–1581.

Johansen S, Juselius K. 1990. Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxford Bulletin of Economics and Statistics* **52**: 169–210.

Karpoff J. 1987. The relation between price changes and trading volume: a survey. *Journal of Financial and Quantitative Analysis* **22**: 109–126.

Kunitomo N. 1992. Improving the Parkinson method of estimating security price volatilities. *Journal of Business* **65**: 295–302.

Lo AW. 1991. Long-term memory in stock market prices. *Econometrica* **59**: 1279–1314.

Lo AW, Mamaysky H, Wang J. 2000. Foundations of technical analysis: computational algorithms, statistical inference, and empirical implementation. *Journal of Finance* **55**: 1075–1765.

Lo AW, Wang J. Stock market trading volume. In *Handbook of Financial Econometrics*, Aït-Sahalia Y, Hansen LP (eds). Elsevier Science Publishing Company, Inc.: New York, in press.

Mok DMY, Lam K, Li WK. 2000. Using daily high/low time to test for intraday random walk in two index futures markets. *Review of Quantitative Finance and Accounting* **14**: 381–397.

Parkinson M. 1980. The extreme value method for estimating the variance of the rate of return. *Journal of Business* **53**: 61–65.

Pesaran MH, Shin Y. 1998. Generalized impulse response analysis in linear multivariate models. *Economics Letters* **58**: 17–29.

Rogers LCG, Satchell SE. 1991. Estimating variance from high, low and closing prices. *Annals of Applied Probability* **1**: 504–512.

Taylor MP, Allen H. 1992. The use of technical analysis in the foreign exchange market. *Journal of International Money and Finance* **11**: 304–314.

Yang D, Zhang Q. 2000. Drift-independent volatility estimation based on high, low, open and close prices. *Journal of Business* **73**: 477–491.

DOI: 10.1002/ijfe