

# AN EMPIRICAL MODEL OF DAILY HIGHS AND LOWS

YIN-WONG CHEUNG

CESIFO WORKING PAPER NO. 1695  
CATEGORY 6: MONETARY POLICY AND INTERNATIONAL FINANCE  
MARCH 2006

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.de](http://www.CESifo-group.de)

# AN EMPIRICAL MODEL OF DAILY HIGHS AND LOWS

## 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% to 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.

JEL Code: C32, G10.

Keywords: high, low open, close, trading volume, VECM model.

*Yin-Wong Cheung*  
*Department of Economics, E2*  
*University of California, Santa Cruz*  
*401 Engineering 2*  
*1156 High Street*  
*Santa Cruz, California 95064*  
*USA*  
*cheung@ucsc.edu*

This Version: February, 2006

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.

## 1. Introduction

Stock price behavior 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.<sup>1</sup> Thus, there is no apparent reason to ignore information on the other three price variables in studying stock price behavior.

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 behavior and, thus, incorporate it in various GARCH and stochastic volatility models to construct conditional or local variance estimators.<sup>2</sup> 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.<sup>3</sup>

The current exercise offers an exploratory analysis of the empirical properties of highs and lows. There are several reasons for analyzing 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 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

---

<sup>1</sup> 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).

<sup>2</sup> 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).

<sup>3</sup> 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).

interest such as the price range. In our exercise, it is shown that modeling the range using only its own history may be inferior to a model that jointly describes the behavior 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.<sup>4</sup>

Third, the high and the low are key components of some technical trading techniques.<sup>5</sup> 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 analyzed in the presence of different groups of explanatory variables.

---

<sup>4</sup> 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.

<sup>5</sup> 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.

## 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 January 2, 1990 to December 31, 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, Rothenberg, Stock, 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:

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

where  $L$  is the lag operator,  $Y_t^\tau$  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  $(\tilde{Y}_1, \tilde{Y}_2, \dots, \tilde{Y}_T) = (Y_1, (1 - \bar{\alpha} L)Y_2, \dots, (1 - \bar{\alpha} L)Y_T)$ ,  $(\tilde{z}_1, \tilde{z}_2, \dots, \tilde{z}_T) = (z_1, (1 - \bar{\alpha} L)z_2, \dots, (1 - \bar{\alpha} L)z_T)$ , and  $\varepsilon_t$  is the error term. The local alternative  $\bar{\alpha}$  is defined by  $\bar{\alpha} = 1 + \bar{c}/T$  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  $\alpha_0 = 0$  against the alternative of  $\alpha_0 < 0$ , is significant.<sup>6</sup>

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

<sup>6</sup> See Elliott, Rothenberg, Stock (1996) and Cheung and Lai (1995) for a detailed description of the testing procedure and the related finite sample critical values.

Table 1. Unit Root Tests

	ADF-GLS				ADF			
	STAT	LAG	Q5	Q10	STAT	LAG	Q5	Q10
A. The Dow Jones Industrial index								
<i>O</i>	-1.352	8	0.007	1.422	-1.249	8	0.007	1.383
<i>H</i>	-1.332	3	5.663	7.909	-1.225	3	5.657	7.910
<i>L</i>	-1.513	6	0.069	5.640	-1.432	6	0.070	5.648
<i>C</i>	-1.53	1	6.455	14.612	-1.447	1	6.483	14.642
<i>V</i>	-7.106*	11	0.899	12.937	-9.588*	10	0.646	11.395
<i>R</i>	-5.043*	10	0.324	6.950	-6.253*	10	0.170	3.858
B. The NASDAQ index								
<i>O</i>	-1.235	6	0.043	11.818	-1.210	6	0.044	11.857
<i>H</i>	-1.235	5	0.733	3.359	-1.202	5	0.749	3.425
<i>L</i>	-1.245	7	0.024	5.698	-1.223	7	0.024	5.672
<i>C</i>	-1.264	3	2.572	7.610	-1.232	3	2.572	7.633
<i>V</i>	-4.556*	13	1.257	11.019	-5.281*	10	1.916	15.072
<i>R</i>	-6.413*	10	1.127	12.329	-7.122*	10	0.847	9.226
C. The S&P 500 index								
<i>O</i>	-1.071	8	0.025	1.149	-0.977	8	0.025	1.13
<i>H</i>	-1.184	2	2.611	9.948	-1.099	2	2.657	10.035
<i>L</i>	-1.156	10	0.006	0.229	-1.065	10	0.005	0.235
<i>C</i>	-1.085	8	0.014	0.660	-0.98	8	0.014	0.631
<i>V</i>	-5.081*	10	1.368	13.592	-5.548*	10	1.168	11.301
<i>R</i>	-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” gives the ADF-GLS and ADF test results. “STAT” gives the test statistics, “LAG” gives the lag parameters used in the test procedures, “\*” indicates the rejection of the unit root null hypothesis at the 5% level and “Q5” and “Q10” gives the Box-Ljung Q-statistics calculated from the first 5 and 10 estimated residual autocorrelations. None of the Q-statistics is significant.

$\Delta Y_t = \delta + \beta t + \gamma Y_{t-1} + \sum_{j=1}^p \beta_j \Delta Y_{t-j} + \varepsilon_t$ . 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.

Table 2. Sample Correlations

A. The Dow Jones Industrial index						
	$\Delta O$	$\Delta H$	$\Delta L$	$\Delta C$	$V$	
$\Delta H$	0.573					
$\Delta L$	0.597	0.800				
$\Delta C$	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						
	$\Delta O$	$\Delta H$	$\Delta L$	$\Delta C$	$V$	
$\Delta H$	0.762					
$\Delta L$	0.761	0.782				
$\Delta C$	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						
	$\Delta O$	$\Delta H$	$\Delta L$	$\Delta C$	$V$	
$\Delta H$	0.634	1.000				
$\Delta L$	0.651	0.641	1.000			
$\Delta C$	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  $\Delta O_t$ ,  $\Delta H_t$ ,  $\Delta L_t$ ,  $\Delta C_t$ , and  $V_t$  of the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index are reported.

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  $\Delta O_t$ ,  $\Delta H_t$ ,  $\Delta L_t$ ,  $\Delta C_t$ ,  $V_t$ , 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.  $\Delta C_t$  tends to have a low correlation with  $\Delta O_t$  but a high correlation with  $\Delta H_t$  and  $\Delta L_t$ . 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.

### 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 2x1 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:

$$\mathbf{X}_t = \boldsymbol{\mu} + \sum_{i=1}^{p+1} \boldsymbol{\gamma}_i \mathbf{X}_{t-i} + \boldsymbol{\varepsilon}_t, \quad (2)$$

where  $\boldsymbol{\mu}$  is the intercept term,  $\boldsymbol{\gamma}_i$ 's are coefficient matrices, and  $\boldsymbol{\varepsilon}_t$  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



Table 3. Cointegration Test Results

	EIGENV	TRACE	<i>H</i>	<i>L</i>	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, B, and 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 labeled “C. Vector” give cointegrating vectors with the coefficient of the high normalized to one. “LAG” gives the lag parameters used to conduct the test.

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.

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 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.<sup>7</sup> 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:

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

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

---

<sup>7</sup> 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.

Table 4. Estimates of the Basic Vector Error Correction Models

A. The Dow Jones Industrial index

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.201	-6.492	0.550	15.691
$\Delta H(-2)$	-0.256	-6.951	0.279	6.701
$\Delta H(-3)$	-0.151	-3.877	0.259	5.863
$\Delta H(-4)$	-0.099	-2.507	0.219	4.902
$\Delta H(-5)$	-0.046	-1.202	0.215	4.950
$\Delta H(-6)$	0.028	0.804	0.216	5.553
$\Delta L(-1)$	0.388	14.254	-0.232	-7.509
$\Delta L(-2)$	0.143	4.449	-0.354	-9.707
$\Delta L(-3)$	0.156	4.549	-0.236	-6.045
$\Delta L(-4)$	0.126	3.605	-0.183	-4.625
$\Delta L(-5)$	0.013	0.376	-0.251	-6.501
$\Delta L(-6)$	-0.024	-0.789	-0.206	-5.908
$\mu$	0.001	1.896	-0.001	-1.226
$R(\cdot)$	-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

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.274	-8.790	0.564	15.019
$\Delta H(-2)$	-0.297	-7.775	0.341	7.437
$\Delta H(-3)$	-0.213	-5.186	0.351	7.112
$\Delta H(-4)$	-0.091	-2.153	0.370	7.238
$\Delta H(-5)$	-0.077	-1.829	0.253	4.975
$\Delta H(-6)$	0.002	0.045	0.229	4.719
$\Delta H(-7)$	-0.006	-0.170	0.196	4.594
$\Delta L(-1)$	0.428	16.530	-0.282	-9.044
$\Delta L(-2)$	0.209	6.437	-0.407	-10.438
$\Delta L(-3)$	0.215	6.057	-0.302	-7.075
$\Delta L(-4)$	0.117	3.155	-0.319	-7.159
$\Delta L(-5)$	0.075	2.002	-0.290	-6.470
$\Delta L(-6)$	0.003	0.095	-0.269	-6.261
$\Delta L(-7)$	0.033	1.021	-0.175	-4.515
$\mu$	0.001	1.637	-0.001	-1.138
$R(\cdot)$	-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 <sup>2</sup>	0.094		0.079	
Q5	0.174		0.537	
Q10	1.857		9.257	

C. The S&P 500 index

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.302	-11.980	0.730	24.937
$\Delta H(-2)$	-0.414	-12.471	0.374	9.687
$\Delta H(-3)$	-0.254	-6.906	0.403	9.429
$\Delta H(-4)$	-0.208	-5.433	0.365	8.218
$\Delta H(-5)$	-0.122	-3.198	0.334	7.519
$\Delta H(-6)$	-0.074	-2.013	0.324	7.562
$\Delta H(-7)$	-0.067	-2.063	0.211	5.584
$\Delta L(-1)$	0.533	24.699	-0.294	-11.706
$\Delta L(-2)$	0.230	8.141	-0.503	-15.311
$\Delta L(-3)$	0.283	8.912	-0.336	-9.108
$\Delta L(-4)$	0.180	5.371	-0.363	-9.322
$\Delta L(-5)$	0.120	3.543	-0.342	-8.664
$\Delta L(-6)$	0.050	1.535	-0.330	-8.663
$\Delta L(-7)$	0.056	1.924	-0.238	-6.985
$\mu$	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 <sup>2</sup>	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, B, and 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 “ $\Delta H$ ” and “ $\Delta L$ .” t-statistics are given in parentheses next to the parameter estimates. “ $\mu$ ” 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-squared statistics are reported in the row labeled “Adjusted R<sup>2</sup>.” Q5 and Q10 give the Q-statistics calculated from the first 5 and 10 sample autocorrelations, respectively. All the Q-statistics are insignificant.

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 behavior. 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.

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-squared statistics of 16.5% and 17.6%. For the other two stock indexes, the adjusted R-squared is between 7.9% to 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-squared 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 modeling 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

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

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  $\Gamma_i$  is a constant vector, we have  $a_i = b_i$  and  $\Delta R_{t-i}$  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,  $\Delta C_{t-1}$  and  $\Delta H_t$ . 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  $\Delta C_{t-1}$  but not available in  $\Delta H_{t-1}$ .  $\Delta C_{t-1}$  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  $\Delta H_t$  and  $\Delta L_t$  is examined using the augmented model:

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

where  $\Delta Y_{t-i}$  is a vector containing  $\Delta O_{t-i}$  and  $\Delta C_{t-i}$ ,  $CO_{t-i}$  is given by  $C_{t-i} - O_{t-i}$ , and  $\Lambda_i$  and  $\theta_i$  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  $\Lambda_i$  and  $\theta_i$ . The significant coefficients of  $\Delta Y_{t-i}$  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.

Table 5. Estimates of the Vector Error Correction Models with Additional Price Variables

A. The Dow Jones Industrial index

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.786	-24.727	-0.168	-4.743
$\Delta H(-2)$	-0.560	-14.247	-0.143	-3.256
$\Delta H(-3)$	-0.415	-9.996	-0.070	-1.510
$\Delta H(-4)$	-0.212	-5.318	0.091	2.053
$\Delta H(-5)$	-0.089	-2.706	0.167	4.526
$\Delta H(-6)$	0.000	-0.015	0.179	5.671
$\Delta L(-1)$	-0.022	-0.786	-0.742	-24.214
$\Delta L(-2)$	-0.008	-0.226	-0.582	-15.628
$\Delta L(-3)$	0.023	0.647	-0.406	-10.441
$\Delta L(-4)$	0.070	2.111	-0.244	-6.588
$\Delta L(-5)$	0.048	1.702	-0.205	-6.450
$\Delta L(-6)$	0.003	0.136	-0.172	-6.075
$\mu$	0.001	2.157	-0.001	-2.782
$\Delta O(-1)$	0.397	4.400	0.363	3.605
$\Delta O(-2)$	0.114	1.539	0.165	2.000
$\Delta O(-3)$	0.107	1.923	0.133	2.147
$\Delta O(-4)$	0.078	2.793	0.080	2.563
$\Delta C(-1)$	0.396	3.867	0.619	5.408
$\Delta C(-2)$	0.219	2.447	0.457	4.584
$\Delta C(-3)$	0.318	4.442	0.406	5.091
$\Delta C(-4)$	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 <sup>2</sup>	0.376		0.404	
Q5	1.736		3.205	
Q10	5.385		5.356	



B. The NASDAQ index

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.764	-22.075	-0.066	-1.637
$\Delta H(-2)$	-0.487	-11.280	0.056	1.106
$\Delta H(-3)$	-0.367	-8.873	0.136	2.803
$\Delta H(-4)$	-0.149	-4.477	0.291	7.450
$\Delta H(-5)$	-0.042	-1.299	0.291	7.600
$\Delta H(-6)$	-0.009	-0.283	0.212	5.776
$\Delta H(-7)$	-0.017	-0.610	0.183	5.662
$\Delta L(-1)$	0.065	2.262	-0.751	-22.285
$\Delta L(-2)$	0.152	4.318	-0.519	-12.550
$\Delta L(-3)$	0.156	4.532	-0.402	-9.987
$\Delta L(-4)$	0.152	5.285	-0.279	-8.249
$\Delta L(-5)$	0.063	2.173	-0.298	-8.817
$\Delta L(-6)$	0.031	1.135	-0.229	-7.081
$\Delta L(-7)$	0.018	0.711	-0.194	-6.658
$\mu$	0.001	3.916	0.000	-1.224
$\Delta O(-1)$	0.308	5.935	0.437	7.201
$\Delta O(-2)$	0.142	3.129	0.242	4.542
$\Delta O(-3)$	0.098	3.107	0.159	4.280
$\Delta C(-1)$	0.292	5.456	0.309	4.926
$\Delta C(-2)$	0.148	3.214	0.210	3.886
$\Delta C(-3)$	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
Adjusted R <sup>2</sup>	0.457		0.477	
Q5	1.994		0.918	
Q10	9.216		4.089	

C. The S&P 500 index

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.880	-31.033	0.035	1.049
$\Delta H(-2)$	-0.666	-18.452	0.031	0.740
$\Delta H(-3)$	-0.500	-13.582	0.119	2.781
$\Delta H(-4)$	-0.259	-8.328	0.321	8.874
$\Delta H(-5)$	-0.135	-4.471	0.321	9.154
$\Delta H(-6)$	-0.121	-4.167	0.273	8.076
$\Delta H(-7)$	-0.087	-3.390	0.187	6.291
$\Delta L(-1)$	0.079	3.314	-0.842	-30.423
$\Delta L(-2)$	0.102	3.439	-0.689	-20.020
$\Delta L(-3)$	0.135	4.415	-0.507	-14.317
$\Delta L(-4)$	0.192	7.166	-0.342	-10.990
$\Delta L(-5)$	0.144	5.363	-0.320	-10.251
$\Delta L(-6)$	0.098	3.801	-0.276	-9.204
$\Delta L(-7)$	0.088	3.786	-0.204	-7.597
$\mu$	0.001	3.412	-0.001	-4.130
$\Delta O(-1)$	0.369	4.953	0.257	2.967
$\Delta O(-2)$	0.305	5.347	0.276	4.161
$\Delta O(-3)$	0.220	7.612	0.224	6.680
$\Delta C(-1)$	0.355	4.128	0.597	5.984
$\Delta C(-2)$	0.232	3.303	0.508	6.243
$\Delta C(-3)$	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 <sup>2</sup>	0.481		0.489	
Q5	3.416	[.636]	5.123	[.401]
Q10	5.723	[.838]	9.025	[.530]

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

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.

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-squared statistic. The NASDAQ daily low equation experiences the largest improvement. The adjusted R-squared 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-squared 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 modeling quality.

#### 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.<sup>8</sup> We investigate the effect trading volume has on highs and lows using the regression:

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

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.

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,

---

<sup>8</sup> See Karpoff (1987) for a detailed review of early studies on the topic. A recent and extensive study is provided by Lo and Wang (2001)

Table 6. Estimates of the Vector Error Correction Models with Trading Volume

A. The Dow Jones Industrial index

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.185	-5.670	0.584	15.718
$\Delta H(-2)$	-0.252	-6.501	0.329	7.456
$\Delta H(-3)$	-0.145	-3.540	0.315	6.789
$\Delta H(-4)$	-0.098	-2.393	0.284	6.107
$\Delta H(-5)$	-0.057	-1.428	0.283	6.244
$\Delta H(-6)$	0.011	0.295	0.283	6.851
$\Delta L(-1)$	0.377	13.101	-0.263	-8.055
$\Delta L(-2)$	0.143	4.217	-0.403	-10.421
$\Delta L(-3)$	0.150	4.138	-0.290	-7.011
$\Delta L(-4)$	0.128	3.490	-0.246	-5.908
$\Delta L(-5)$	0.026	0.736	-0.318	-7.788
$\Delta L(-6)$	-0.010	-0.310	-0.272	-7.250
$\mu$	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
$R(\cdot)$	-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

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.264	-8.233	0.581	14.986
$\Delta H(-2)$	-0.281	-7.072	0.369	7.679
$\Delta H(-3)$	-0.200	-4.685	0.382	7.376
$\Delta H(-4)$	-0.091	-2.058	0.402	7.516
$\Delta H(-5)$	-0.088	-2.001	0.289	5.438
$\Delta H(-6)$	-0.020	-0.461	0.276	5.391
$\Delta H(-7)$	-0.025	-0.667	0.246	5.374
$\Delta L(-1)$	0.422	15.901	-0.298	-9.273
$\Delta L(-2)$	0.196	5.813	-0.431	-10.549
$\Delta L(-3)$	0.202	5.421	-0.329	-7.317
$\Delta L(-4)$	0.115	2.980	-0.348	-7.438
$\Delta L(-5)$	0.083	2.145	-0.322	-6.843
$\Delta L(-6)$	0.019	0.505	-0.311	-6.833
$\Delta L(-7)$	0.050	1.443	-0.220	-5.287
$\mu$	0.001	3.420	-0.001	-2.091
$V$	0.009	7.923	-0.006	-4.035
$V(-1)$	-0.007	-6.049	0.002	1.592
$R(\cdot)$	-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

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.299	-10.163	0.731	21.456
$\Delta H(-2)$	-0.403	-10.258	0.386	8.476
$\Delta H(-3)$	-0.244	-5.608	0.429	8.520
$\Delta H(-4)$	-0.195	-4.342	0.420	8.069
$\Delta H(-5)$	-0.117	-2.606	0.398	7.630
$\Delta H(-6)$	-0.067	-1.538	0.405	8.032
$\Delta H(-7)$	-0.054	-1.383	0.288	6.408
$\Delta L(-1)$	0.517	20.461	-0.311	-10.610
$\Delta L(-2)$	0.220	6.589	-0.526	-13.597
$\Delta L(-3)$	0.270	7.182	-0.360	-8.256
$\Delta L(-4)$	0.162	4.101	-0.426	-9.332
$\Delta L(-5)$	0.108	2.696	-0.415	-8.909
$\Delta L(-6)$	0.039	0.999	-0.410	-9.056
$\Delta L(-7)$	0.052	1.468	-0.307	-7.464
$\mu$	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
$R(\cdot)$	-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, B, and 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.

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 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-squared statistic by 1% to 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

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

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-squared statistics are marginally higher than those in Table 5, and the Q-statistics are good.

The explanatory power of  $V_t$  relative to  $\Delta Y_t$  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-squared statistics.

Table 7. Estimates of the Vector Error Correction Models with Additional Price Variables and Trading Volume

A. The Dow Jones Industrial index

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-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
$\mu$	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
Adjusted R <sup>2</sup>	0.390		0.412	
Q5	1.800		2.148	
Q10	4.402		5.065	



B. The NASDAQ index

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.776	-22.387	-0.054	-1.310
$\Delta H(-2)$	-0.493	-11.283	0.070	1.345
$\Delta H(-3)$	-0.364	-8.616	0.147	2.929
$\Delta H(-4)$	-0.160	-4.647	0.307	7.496
$\Delta H(-5)$	-0.069	-2.037	0.313	7.793
$\Delta H(-6)$	-0.049	-1.495	0.239	6.174
$\Delta H(-7)$	-0.054	-1.846	0.209	6.042
$\Delta L(-1)$	0.056	1.944	-0.753	-22.025
$\Delta L(-2)$	0.143	4.038	-0.523	-12.405
$\Delta L(-3)$	0.154	4.417	-0.410	-9.918
$\Delta L(-4)$	0.162	5.434	-0.293	-8.259
$\Delta L(-5)$	0.085	2.846	-0.317	-8.925
$\Delta L(-6)$	0.065	2.241	-0.253	-7.366
$\Delta L(-7)$	0.050	1.911	-0.218	-6.946
$\mu$	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
$\Delta O(-1)$	0.325	6.367	0.426	7.004
$\Delta O(-2)$	0.149	3.336	0.236	4.423
$\Delta O(-3)$	0.094	3.008	0.159	4.285
$\Delta C(-1)$	0.296	5.601	0.313	4.978
$\Delta C(-2)$	0.155	3.416	0.210	3.887
$\Delta C(-3)$	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

	$\Delta H$		$\Delta L$	
	COEFF	T-STAT	COEFF	T-STAT
$\Delta H(-1)$	-0.891	-27.577	0.033	0.895
$\Delta H(-2)$	-0.686	-16.449	0.023	0.487
$\Delta H(-3)$	-0.493	-11.546	0.147	2.971
$\Delta H(-4)$	-0.258	-7.075	0.358	8.507
$\Delta H(-5)$	-0.133	-3.739	0.383	9.332
$\Delta H(-6)$	-0.120	-3.514	0.344	8.668
$\Delta H(-7)$	-0.086	-2.803	0.251	7.086
$\Delta L(-1)$	0.063	2.315	-0.847	-26.885
$\Delta L(-2)$	0.081	2.379	-0.711	-18.040
$\Delta L(-3)$	0.121	3.426	-0.527	-12.959
$\Delta L(-4)$	0.187	5.969	-0.393	-10.843
$\Delta L(-5)$	0.139	4.390	-0.382	-10.418
$\Delta L(-6)$	0.096	3.112	-0.345	-9.714
$\Delta L(-7)$	0.088	3.147	-0.267	-8.260
$\mu$	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
$\Delta O(-1)$	0.374	4.741	0.273	2.997
$\Delta O(-2)$	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
$R(\cdot)$	-0.041	-1.853	0.180	7.006
Adjusted R <sup>2</sup>	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, B, and C give the results for the Dow Jones Industrial index, the NASDAQ index, and the S&P 500 index.  $\Delta O(\cdot)$ ,  $\Delta C(\cdot)$ ,  $CO(\cdot)$ , and  $V(\cdot)$  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  $\varepsilon_t$  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  $j$ -th variable ( $j = 1$  for a shock to the high and  $j = 2$  for a shock to the low) at time  $t$  is given by

$$B_h \Sigma e_j / \sigma_{jj}, \quad h = 0, 1, 2, \dots \quad (8)$$

where  $B_h = \gamma_1 B_{h-1} + \gamma_2 B_{h-2} + \dots + \gamma_p B_{h-p} + \gamma_{p+1} B_{h-p-1}$ ,  $h = 1, 2, \dots$ ,  $B_0 = I$ , and  $B_h = 0$  for  $h < 0$ . Note that the matrices  $\{B_h, t = 1, 2, \dots\}$  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  $j$ -th 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), (6), and (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.<sup>9</sup> 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.

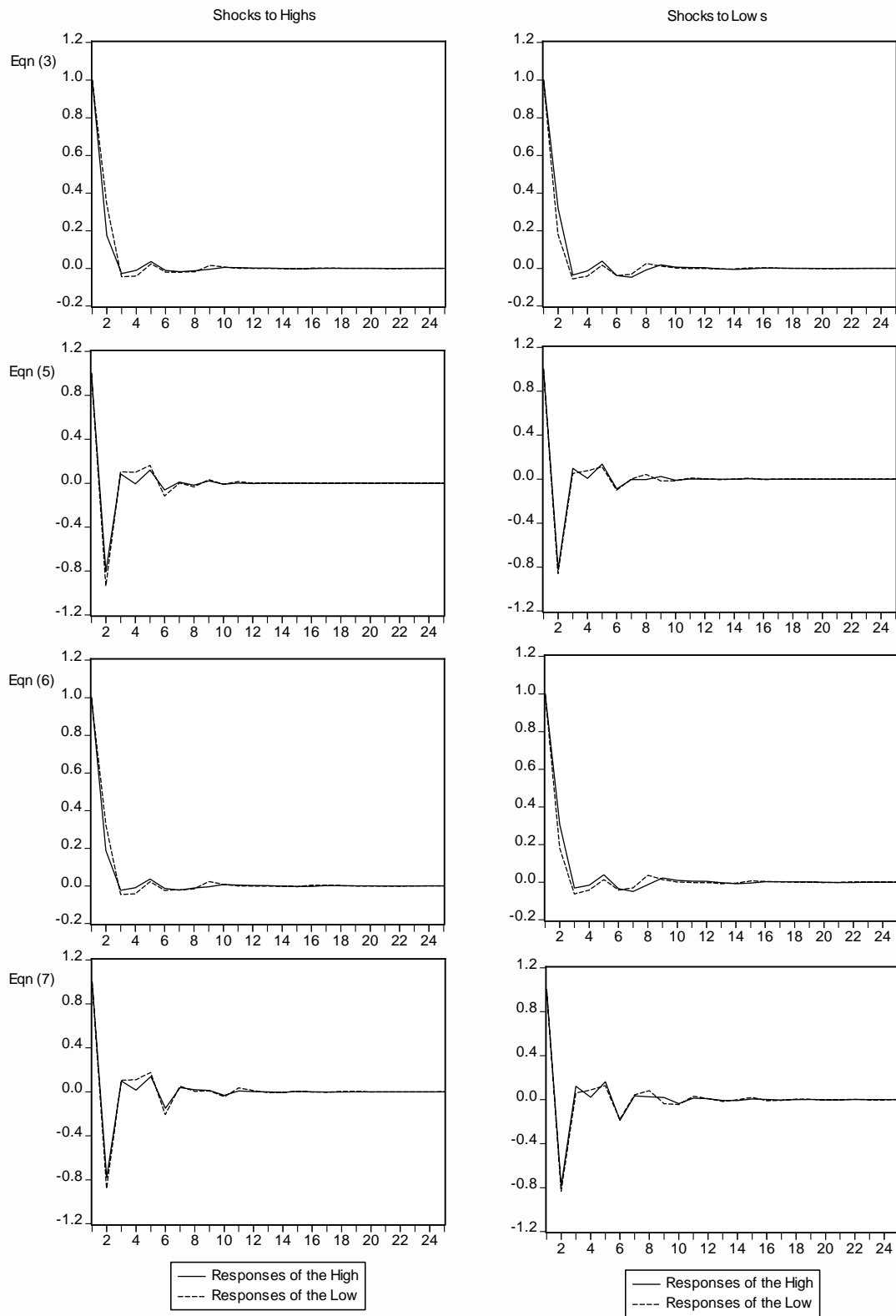
The responses to these shocks change quite substantially in the presence of data on openings and closings. In contrast to the basic VECM model, one-day lagged responses to

---

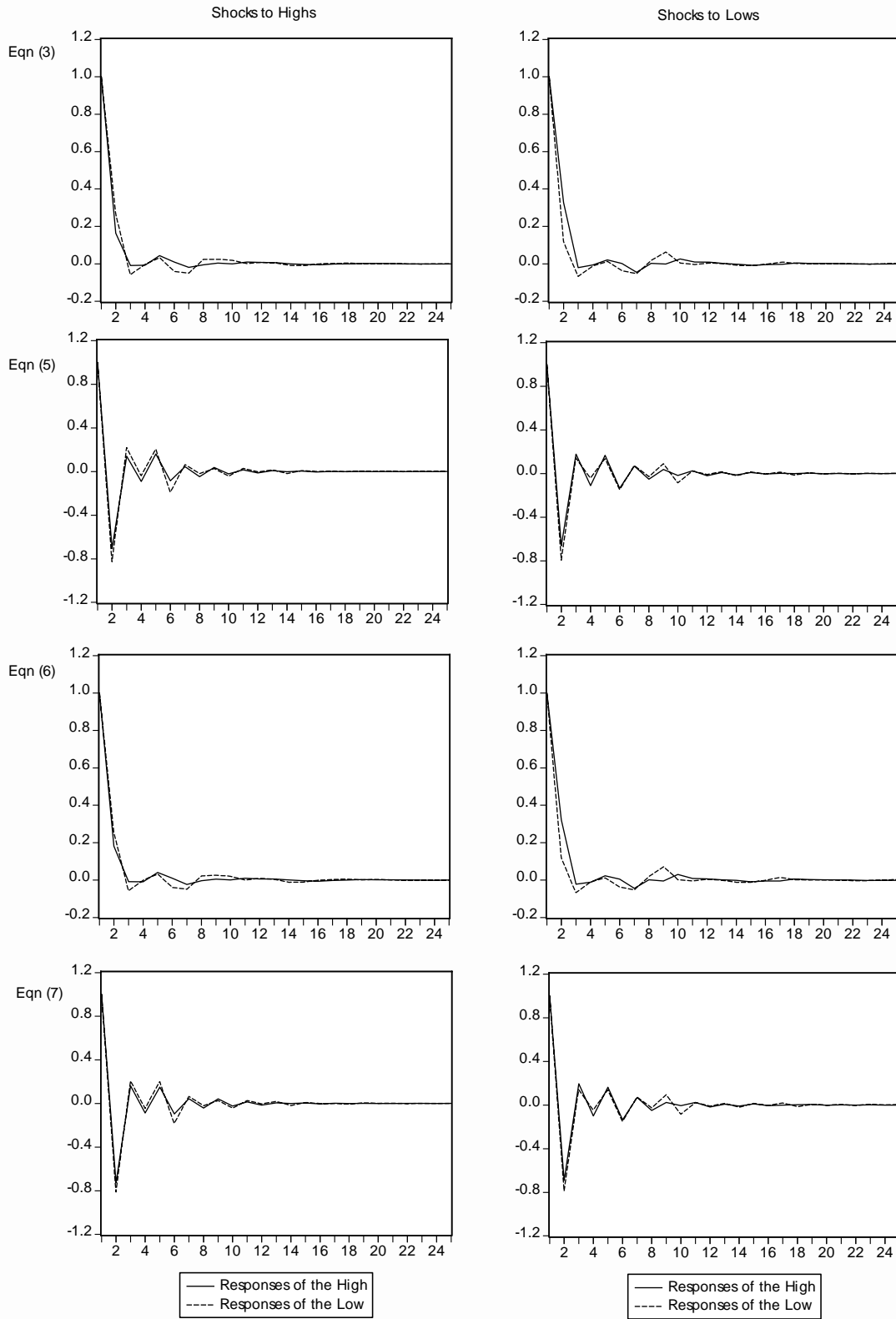
<sup>9</sup> Number “2” on the horizontal axis corresponds to the one-day lagged response to the initial unit shock.

Figure 1. Generalized Impulse Responses

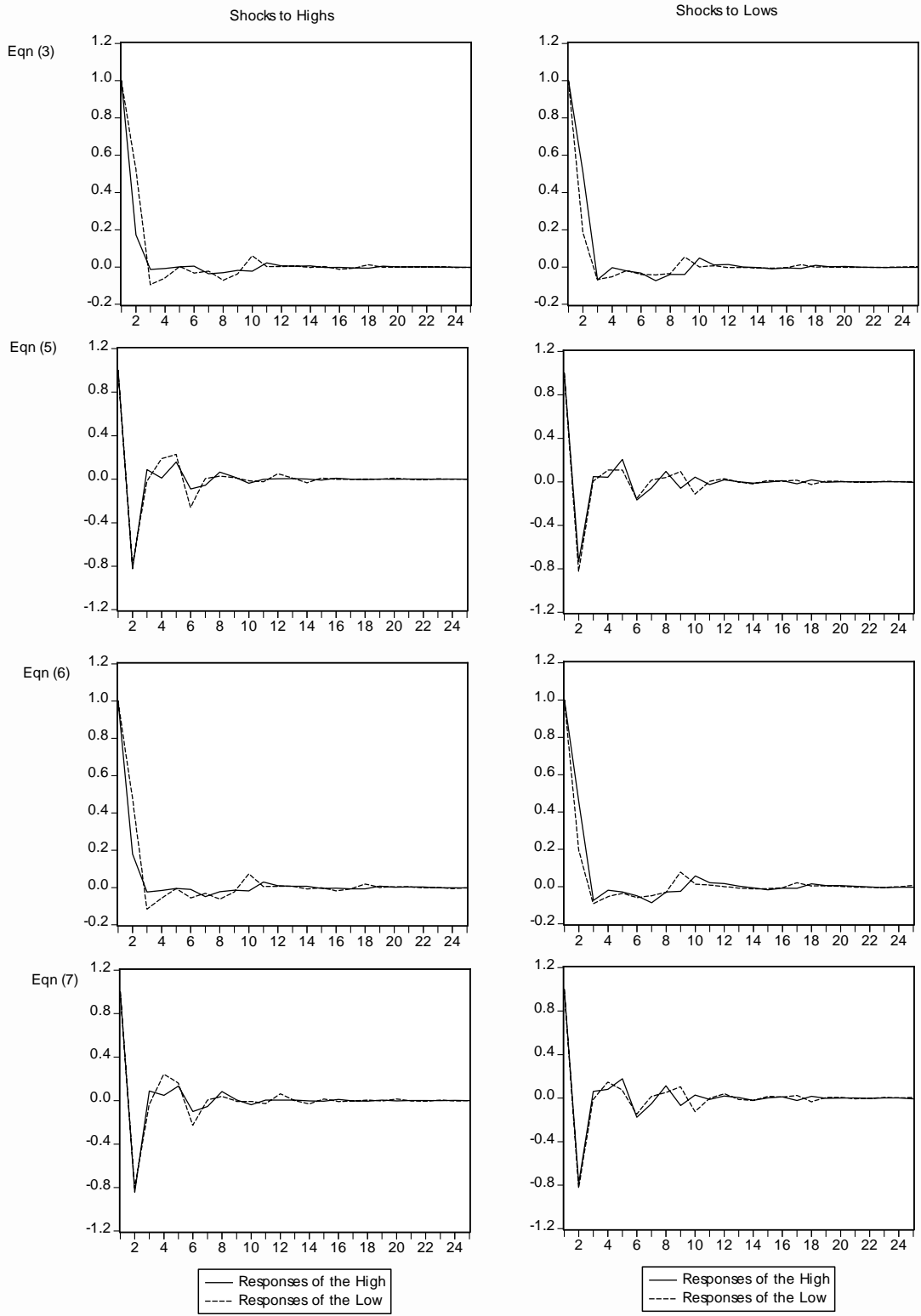
A. The Dow Jones Industrial Index



B. The NASDAQ Index



C. The S&P 500 Index



shocks are negative for all the three stock indexes under specification (5). The magnitude of the first day responses is larger than the one 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.

## **6. 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% to 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 modeling 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 modeling 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 behavior, 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.<sup>10</sup> 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.

---

<sup>10</sup> 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., M. W. Brandt, and F. X. Diebold, 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.
- Cheung, Y.-W. (1993), Long Memory in Foreign Exchange Rates, *Journal of Business & Economic Statistics* 11, 93-102.
- Cheung, Y.-W. and C.Y.P. Wong, 2000, A Survey of Market Practitioners' Views on Exchange Rate Dynamics, *Journal of International Economics* 51, 401-419.
- Cheung, Y.-W., Y.L. Cheung, and A. Wan, 2005, A High-Low Model for Forecasting Daily Stock Price Ranges, manuscript, UCSC.
- Cheung, Y.-W. and M. Chinn, 2001, Currency Traders and Exchange Rate Dynamics: A Survey of the U.S. Market, *Journal of International Money and Finance* 20, 439-471.
- Cheung, Y.-W. & K. S. Lai, 1995, Lag order and critical values of a modified Dickey-Fuller test, *Oxford Bulletin of Economics and Statistics* 57, 411-419.
- 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, R.D. and Magee, J., 1997, *Technical analysis of stock trends*, 7<sup>th</sup> ed. New York : Amacom.
- Elliott, G., T. J. Rothenberg, & J. H. Stock, 1996, Efficient tests for an autoregressive unit root, *Econometrica* 64, 813-836.
- Engle, R.F. and G.M. Gallo, 2003, A Multiple Indicators Model for Volatility Using Intra-Daily Data, NBER working paper 10117.
- Gallant, A.R., C.T. Hsu, and G.E. Tauchen, 1999, Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance, *Review of Economics and Statistics* 81, 617-631.
- Garman, M.B., and M. J. Klass, 1980, On the Estimation of Price Volatility from Historical Data, *Journal of Business* 53, 67-78.
- Hurst, H. E., 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. and K. Juselius, 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, A.W., 1991, Long-Term Memory in Stock Market Prices, *Econometrica* 59, 1279-1314.
- Lo, A.W., H. Mamaysky, and J. Wang, 2000, Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation, *Journal of Finance* 55, 1075-1765.
- Lo, A.W. and J.Wang, 2001, Stock Market Trading Volume, manuscript. (to appear in the *Handbook of Financial Econometrics*.)
- Mok, D. M.Y., K. Lam, and W.K. Li, 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, M.H. and Y. Shin, 1998, Generalized Impulse Response Analysis in Linear Multivariate Models, *Economics Letters* 58, 17-29.
- Rogers, L.C.G., and S.E. Satchell, 1991, Estimating Variance from High, Low and Closing Prices, *Annals of Applied Probability* 1, 504-512.
- Taylor, M.P., Allen, H., 1992, The Use of Technical Analysis in The Foreign Exchange Market, *Journal of International Money and Finance* 11, 304–314.
- Yang, D., and Qiang Z., 2000, Drift-Independent Volatility Estimation Based on High, Low, Open and Close Prices, *Journal of Business* 73, 477-491.

# CESifo Working Paper Series

(for full list see [www.cesifo-group.de](http://www.cesifo-group.de))

---

- 1633 Michiel Evers, Ruud A. de Mooij and Daniel J. van Vuuren, What Explains the Variation in Estimates of Labour Supply Elasticities?, December 2005
- 1634 Matthias Wrede, Health Values, Preference Inconsistency, and Insurance Demand, December 2005
- 1635 Hans Jarle Kind, Marko Koethenbueger and Guttorm Schjelderup, Do Consumers Buy Less of a Taxed Good?, December 2005
- 1636 Michael McBride and Stergios Skaperdas, Explaining Conflict in Low-Income Countries: Incomplete Contracting in the Shadow of the Future, December 2005
- 1637 Alfons J. Weichenrieder and Oliver Busch, Artificial Time Inconsistency as a Remedy for the Race to the Bottom, December 2005
- 1638 Aleksander Berentsen and Christopher Waller, Optimal Stabilization Policy with Flexible Prices, December 2005
- 1639 Panu Poutvaara and Mikael Priks, Violent Groups and Police Tactics: Should Tear Gas Make Crime Preventers Cry?, December 2005
- 1640 Yin-Wong Cheung and Kon S. Lai, A Reappraisal of the Border Effect on Relative Price Volatility, January 2006
- 1641 Stefan Bach, Giacomo Corneo and Viktor Steiner, Top Incomes and Top Taxes in Germany, January 2006
- 1642 Johann K. Brunner and Susanne Pech, Optimum Taxation of Life Annuities, January 2006
- 1643 Naércio Aquino Menezes Filho, Marc-Andreas Muendler and Garey Ramey, The Structure of Worker Compensation in Brazil, with a Comparison to France and the United States, January 2006
- 1644 Konstantinos Angelopoulos, Apostolis Philippopoulos and Vangelis Vassilatos, Rent-Seeking Competition from State Cooffers: A Calibrated DSGE Model of the Euro Area, January 2006
- 1645 Burkhard Heer and Bernd Suessmuth, The Savings-Inflation Puzzle, January 2006
- 1646 J. Stephen Ferris, Soo-Bin Park and Stanley L. Winer, Political Competition and Convergence to Fundamentals: With Application to the Political Business Cycle and the Size of Government, January 2006

- 1647 Yu-Fu Chen, Michael Funke and Kadri Männasoo, Extracting Leading Indicators of Bank Fragility from Market Prices – Estonia Focus, January 2006
- 1648 Panu Poutvaara, On Human Capital Formation with Exit Options: Comment and New Results, January 2006
- 1649 Anders Forslund, Nils Gottfries and Andreas Westermark, Real and Nominal Wage Adjustment in Open Economies, January 2006
- 1650 M. Hashem Pesaran, Davide Pettenuzzo and Allan G. Timmermann, Learning, Structural Instability and Present Value Calculations, January 2006
- 1651 Markku Lanne and Helmut Luetkepohl, Structural Vector Autoregressions with Nonnormal Residuals, January 2006
- 1652 Helge Berger, Jakob de Haan and Jan-Egbert Sturm, Does Money Matter in the ECB Strategy? New Evidence Based on ECB Communication, January 2006
- 1653 Axel Dreher and Friedrich Schneider, Corruption and the Shadow Economy: An Empirical Analysis, January 2006
- 1654 Stefan Brandauer and Florian Englmaier, A Model of Strategic Delegation in Contests between Groups, January 2006
- 1655 Jan Zápál and Ondřej Schneider, What are their Words Worth? Political Plans and Economic Pains of Fiscal Consolidations in New EU Member States, January 2006
- 1656 Thiess Buettner, Sebastian Hauptmeier and Robert Schwager, Efficient Revenue Sharing and Upper Level Governments: Theory and Application to Germany, January 2006
- 1657 Daniel Haile, Abdolkarim Sadrieh and Harrie A. A. Verbon, Cross-Racial Envy and Underinvestment in South Africa, February 2006
- 1658 Frode Meland and Odd Rune Straume, Outsourcing in Contests, February 2006
- 1659 M. Hashem Pesaran and Ron Smith, Macroeconometric Modelling with a Global Perspective, February 2006
- 1660 Alexander F. Wagner and Friedrich Schneider, Satisfaction with Democracy and the Environment in Western Europe – a Panel Analysis, February 2006
- 1661 Ben J. Heijdra and Jenny E. Ligthart, Fiscal Policy, Monopolistic Competition, and Finite Lives, February 2006
- 1662 Ludger Woessmann, Public-Private Partnership and Schooling Outcomes across Countries, February 2006
- 1663 Topi Miettinen and Panu Poutvaara, Political Parties and Network Formation, February 2006

- 1664 Alessandro Cigno and Annalisa Luporini, Optimal Policy Towards Families with Different Amounts of Social Capital, in the Presence of Asymmetric Information and Stochastic Fertility, February 2006
- 1665 Samuel Muehlemann and Stefan C. Wolter, Regional Effects on Employer Provided Training: Evidence from Apprenticeship Training in Switzerland, February 2006
- 1666 Laszlo Goerke, Bureaucratic Corruption and Profit Tax Evasion, February 2006
- 1667 Ivo J. M. Arnold and Jan J. G. Lemmen, Inflation Expectations and Inflation Uncertainty in the Eurozone: Evidence from Survey Data, February 2006
- 1668 Hans Gersbach and Hans Haller, Voice and Bargaining Power, February 2006
- 1669 Françoise Forges and Frédéric Koessler, Long Persuasion Games, February 2006
- 1670 Florian Englmaier and Markus Reisinger, Information, Coordination, and the Industrialization of Countries, February 2006
- 1671 Hendrik Hakenes and Andreas Irmen, Something out of Nothing? Neoclassical Growth and the 'Trivial' Steady State, February 2006
- 1672 Torsten Persson and Guido Tabellini, Democracy and Development: The Devil in the Details, February 2006
- 1673 Michael Rauber and Heinrich W. Ursprung, Evaluation of Researchers: A Life Cycle Analysis of German Academic Economists, February 2006
- 1674 Ernesto Reuben and Frans van Winden, Reciprocity and Emotions when Reciprocators Know each other, February 2006
- 1675 Assar Lindbeck and Mats Persson, A Model of Income Insurance and Social Norms, February 2006
- 1676 Horst Raff, Michael Ryan and Frank Staehler, Asset Ownership and Foreign-Market Entry, February 2006
- 1677 Miguel Portela, Rob Alessie and Coen Teulings, Measurement Error in Education and Growth Regressions, February 2006
- 1678 Andreas Haufler, Alexander Klemm and Guttorm Schjelderup, Globalisation and the Mix of Wage and Profit Taxes, February 2006
- 1679 Kurt R. Brekke and Lars Sjørgard, Public versus Private Health Care in a National Health Service, March 2006
- 1680 Dominik Grafenhofer, Christian Jaag, Christian Keuschnigg and Mirela Keuschnigg, Probabilistic Aging, March 2006

- 1681 Wladimir Raymond, Pierre Mohnen, Franz Palm and Sybrand Schim van der Loeff, Persistence of Innovation in Dutch Manufacturing: Is it Spurious?, March 2006
- 1682 Andrea Colciago, V. Anton Muscatelli, Tiziano Ropele and Patrizio Tirelli, The Role of Fiscal Policy in a Monetary Union: Are National Automatic Stabilizers Effective?, March 2006
- 1683 Mario Jametti and Thomas von Ungern-Sternberg, Risk Selection in Natural Disaster Insurance – the Case of France, March 2006
- 1684 Ken Sennewald and Klaus Waelde, “Itô’s Lemma“ and the Bellman Equation for Poisson Processes: An Applied View, March 2006
- 1685 Ernesto Reuben and Frans van Winden, Negative Reciprocity and the Interaction of Emotions and Fairness Norms, March 2006
- 1686 Françoise Forges, The Ex Ante Incentive Compatible Core in Exchange Economies with and without Indivisibilities, March 2006
- 1687 Assar Lindbeck, Mårten Palme and Mats Persson, Job Security and Work Absence: Evidence from a Natural Experiment, March 2006
- 1688 Sebastian Buhai and Coen Teulings, Tenure Profiles and Efficient Separation in a Stochastic Productivity Model, March 2006
- 1689 Gebhard Kirchgaessner and Silika Prohl, Sustainability of Swiss Fiscal Policy, March 2006
- 1690 A. Lans Bovenberg and Peter Birch Sørensen, Optimal Taxation and Social Insurance in a Lifetime Perspective, March 2006
- 1691 Moritz Schularick and Thomas M. Steger, Does Financial Integration Spur Economic Growth? New Evidence from the First Era of Financial Globalization, March 2006
- 1692 Burkhard Heer and Alfred Maussner, Business Cycle Dynamics of a New Keynesian Overlapping Generations Model with Progressive Income Taxation, March 2006
- 1693 Jarko Fidrmuc and Iikka Korhonen, Meta-Analysis of the Business Cycle Correlation between the Euro Area and the CEECs, March 2006
- 1694 Steffen Henzel and Timo Wollmershaeuser, The New Keynesian Phillips Curve and the Role of Expectations: Evidence from the Ifo World Economic Survey, March 2006
- 1695 Yin-Wong Cheung, An Empirical Model of Daily Highs and Lows, March 2006