An Empirical Model Of Fractionally Cointegrated Daily High And Low Stock Market Prices[1]

Jozef Baruník[a] and Sylvie Dvořáková[b]

[a)] *Institute of Economic Studies, Charles University, Opletalova 21, 110 00, Prague, Czech Republic; Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodarenskou Vezi 4, 182 00, Prague, Czech Republic. E-mail: barunik@utia.cas.cz*
[b)] *Institute of Economic Studies, Charles University, Opletalova 21, 110 00, Prague, Czech Republic*

Abstract
This work provides empirical support for the fractional cointegration relationship between daily high and low stock prices, allowing for the non-stationary volatility of stock market returns. The recently formalized fractionally cointegrated vector autoregressive (VAR) model is employed to explain both the cointegration dynamics between daily high and low stock prices and the long memory of their linear combination, i.e., the range. Daily high and low stock prices are of particular interest because they provide valuable information about range-based volatility, which is considered a highly efficient and robust estimator of volatility. We provide a comparison of the Czech PX index with other world market indices: the German Deutscher Aktienindex (DAX), U.K. Financial Times Stock Exchange (FTSE) 100, U.S. Standard and Poor's (S&P) 500 and Japanese Nihon Keizai Shimbun (NIKKEI) 225 during the 2003-2012 period, that is, before and during the financial crisis. We find that the ranges of all of the indices display long memory and are mostly in the non-stationary region, supporting recent evidence that volatility might not be a stationary process. No common pattern is detected among all of the studied indices, and different behaviors are also observed in the pre-crisis and post-crisis periods. We conclude that the fractionally cointegrated VAR approach allowing for long memory is an interesting alternative for modeling range-based volatility.

Keywords: fractional cointegration, long memory, range, volatility, daily high and low prices
JEL Classification: C32, C58, G15

1 Introduction

Daily high and low stock market prices provide valuable information about range-based volatility that is not included in the open and close prices commonly studied by researchers. More specifically, the difference between high and low prices, i.e., the range, provides an efficient estimator of volatility robust to noise (Parkinson, 1980). To date, stock prices in developed markets have generally been considered to be unpredictable and are assumed to follow a random walk. Hence, most studies consider stock prices to be integrated of order 1 (an $I(1)$ process)[2]. However, the choice between stationary $I(0)$[3] and non-stationary $I(1)$ processes can be too restrictive for the degree of integration of daily high and low prices. Because high and low prices can be modeled together as a possibly fractionally cointegrated relationship (Fiess and MacDonald, 2002; Cheung, 2007), it allows for greater flexibility. This idea is especially interesting because the error correction term from the cointegrating relationship between high and low prices is the range. Hence, a more general fractional or

[1] Jozef Barunik gratefully acknowledges support from the Czech Science Foundation Project No. P402/12/G097 DYME—"Dynamic Models in Economics."
[2] An process is a non-stationary process, where only a single differencing is sufficient to obtain stationarity.
[3] An) process is a stationary process, where no differencing is needed to achieve stationarity.

long-memory framework, where the series are assumed to be integrated of order d and cointegrated of order less than d, i.e., $CI(d-b)$, where $d, b \in \mathbb{R}$ and $0 < b \leq d$, could be more useful in capturing the empirical properties of data.

To determine the fractionally cointegrated relationship between highs and lows, we implement a fractionally cointegrated vector autoregressive model (FCVAR), as proposed by Johansen (2008) and Johansen and Nielsen (2010, 2012). The motivation for utilizing this framework is twofold. First, daily highs and lows are assumed to be cointegrated, i.e., in the short term they may diverge, but in the long term they have an embedded convergence path. Second, their specific linear combination is an efficient volatility estimate, i.e., the range, and is assumed to display a long memory.

Substantial evidence of the presence of long memory has been documented in the literature not only in the volatility of asset prices (Ding, Granger and Engel, 1993; Andersen and Bollerslev, 1997; Breidt et al., 1998; Kellard et al., 2010; or Garvey and Gallagher, 2012) but also in the interest rate differentials, inflation rates, forward premiums, and exchange rates (Baillie, 1996). Although the vast literature concludes that volatility is a long-memory process, few studies suggest that volatility is a non-stationary process with the long memory parameter d being greater or equal to 0.5 (Kellard et al., 2010). Yalama and Celik (2012) provide an excellent review of the literature studying the long memory properties of volatility and document the feature empirically as well.

This work contributes to the literature through an empirical investigation of world market indices, especially of their daily high and low prices, in the fractional cointegration framework. Their linear combination, the daily range, is found to be a non-stationary process. Whereas Caporin et al. (2013) suggest a fractionally cointegrated framework for modeling daily high and low prices in their pioneering work, we present new empirical evidence of the long memory behavior of global stock markets. Moreover, we add a long memory analysis utilizing different measures and different periods such as pre-crisis and crisis, and thus we present new empirical evidence.

The analysis is performed on four global stock market indices, the U.S. Standard and Poor's (S&P) 500, German Deutscher Aktienindex (DAX), Japanese Nihon Keizai Shimbun (NIKKEI) 225, and U.K. Financial Times Stock Exchange (FTSE) 100 index over the 10-year period 2003-2012. These results are compared to the Czech PX Index[4] over the same period. Moreover, we study the behavior of the high and low prices in two sub-periods with December 2007 as the dividing point. This analysis enables us to compare both cointegration and estimated volatility before and during the recent financial crisis. The main result is that we find significant evidence of long memory in the daily ranges falling in the non-stationary region (except for the PX Index and NIKKEI 225 in the first period). Our results also distinguish between the two sub-periods. The long memory estimates during the first period of 2003-2007 are generally lower in comparison to the second period, where primarily the years 2008 and 2009 seem to increase the long memory. Overall, the PX Index displays the lowest estimates of the order of price range integration, and its behavior is very similar to the NIKKEI 225 in this respect. The ranges of the S&P 500, FTSE 100 and DAX indices display, however, relatively higher orders of integration. Furthermore, we find that the unrestricted FCVAR performs better in detecting the stationarity of the range indicated by other applied tests than the FCVAR specification with restrictions on the cointegrating vector.

The remainder of the study is organized as follows. Section 2 describes the motivation for using daily high and low prices and descriptions of the data. In the Section 3, we conduct the preliminary analysis of daily high and low prices and the range, focusing on their long memory properties. Section 4 then suggests an empirical model of fractionally cointegrated

[4] The PX Index was introduced in March 2006 as a merging of two indices, PX-D and PX 50. The PX Index obtained the historical prices from the PX 50 and continued henceforth.

daily high and low prices and discusses the main results. Final Section 5 concludes the findings.

2 Motivation and data description

In this work, we focus on investigating daily high and low prices. By the high price, we understand the maximum price observed during the day, and by the low price, we understand the minimum price achieved during that day. These prices can be viewed as additional information about the change in direction of excess demand (Cheung, 2007). Caporin et al. (2013) nicely summarize other reasons why the daily high and low prices are of importance. First, daily high and low prices can have a role as a reference level. Stock market agents employ these reference levels to make assumptions and predictions about future developments and employ daily highs and lows as reference values. Second, daily highs and lows can function as a stop-loss indicator and may contain information about liquidity provisioning and the price discovery process. Third, high and low prices are more likely to correspond to ask and bid quotes, respectively, implying that they may be influenced by transaction costs and other market frictions (e.g., price discreteness, stale prices, and tick size). Moreover, daily high and low prices tend to react to unanticipated public announcements or other unexpected shocks.

Daily high and low stock prices are primarily valuable as a measure of dispersion, i.e., a measure of the deviation from the mean. In financial literature, dispersion measures the degree of uncertainty, and thus risk, associated with a particular asset. Parkinson (1980) was among the first to show that a variance estimator based on close-to-close returns is a far less efficient volatility estimator than the price range defined as a difference between daily high and low prices. Alizadeh et al. (2002) further demonstrate that a range-based estimator of volatility is highly statistically efficient and robust with respect to several types of microstructure frictions because it is much less contaminated by measurement error and explains not only the autocorrelation of volatility but also the volatility of volatility. Furthermore, Corwin and Schultz (2012) argue that, because daily high and low prices are mostly buy and sell trades, respectively, the price range thus represents a fundamental volatility because it reflects both the stock's variance and its bid-ask spread. Alizadeh et al. (2002) note that using the range as a volatility proxy has a "long and colorful" history in finance (e.g., Garman and Klass, 1980; Parkinson, 1980; Andersen and Bollerslev, 1998; Degiannakis and Livada, 2013). More recently, Caporin et al. (2013) find evidence of long memory in the ranges of all 30 of the components of the Dow Jones Industrial Average (DJIA) index during the 2003-2010 period.

This work focuses on analysis of the daily high and low prices of four major global indices over the 2003-2012 period covering both the calm and financial crisis periods. We consider four world indices: the U.S. S&P 500, German DAX, Japanese NIKKEI 225, and U.K. FTSE 100, available from TICK data, which are examined and compared to the Czech stock market index. Moreover, we examine the indices during the entire 10-year period from January 2003 to December 2012 and as during two sub-periods. The first sub-period covers the pre-crisis years from January 2003 until December 2007. This break point has been chosen based on the statement of the National Bureau of Economic Research (NBER), which identified December 2007 as the peak of pre-crisis economic activity. The subsequent decline in economic activity was large enough to be qualified as a recession. The second sub-period spans from January 2008 until December 2012 and covers the recent crisis period. We synchronize the data with the same time stamps but discard holidays from further analysis.

Figure 1 depicts the development of daily high and low prices of the PX index and their difference – the range. The peak of December 2007 is depicted as a vertical line in the figure.

We can see that the PX index was experiencing steep growth in the first period, but at the end of 2007, it was severely hit by the crisis. After a stable period from mid-2009 through mid-2011, another drop followed; however, it was less dramatic than the decline observed at the end of 2007.

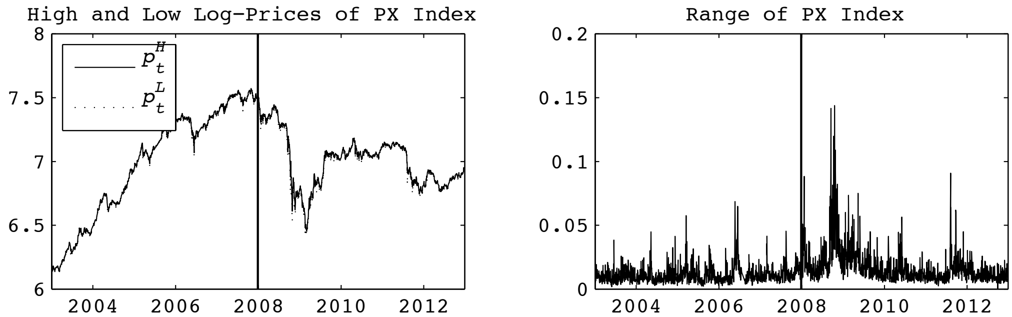

Figure 1: High and low prices of the PX index (left) and range of the PX index (right)

The range-based volatility measured as the difference between daily high and low prices is higher after the outbreak of the crisis and reaches its maximum at the end of 2008; then it gradually returns to its pre-crisis values. A similar pattern emerges for all of the other studied indices[5]. With the DAX, we document a slightly different behavior, mainly after the crisis, when it grows steadily following the drop in 2007, with the exception of a short period of decline at the end of 2011. This pattern is also the case with the S&P 500. However, the FTSE 100 displays only a very slight or no growth after the crisis, and the NIKKEI 225 actually declines.

When we compare the behavior of all of the studied markets, we find that they are influenced by similar factors; however, the reaction to the crisis is quite different for each of them. All of the indices experienced rapid growth during the 2003-2007 period followed by a steep downturn at the end of 2007. The DAX and PX indices are the two indices least affected by the crisis. The DAX is also the first index in our sample to recover from the crash and remains the best-performing one. On the other hand, the PX index is the second-least hit by the crisis, but is the second-worst performing index as of the end of 2012. Until mid-2009, it closely follows the DAX, but then it loses pace and falls at the end of 2011. The other indices also experience this fall, but in contrast to the PX Index and NIKKEI 225, they are able to resume their previous growth. Although the NIKKEI 225 outperforms the remaining indices slightly before the crisis, this is no longer true after the crisis, when its performance is the worst; it never quite recuperates from the crisis. Since the beginning of 2012, the S&P 500 begins to catch up with DAX, and the FTSE 100 follows close behind. Overall, we can see that before the crisis, all of the indices are growing together, but the reactions to the crisis vary significantly.

3 Preliminary analysis of daily high, low prices and range

Following the intuition that daily high and low prices are non-stationary and share a common trend, Cheung (2007) proposed to model them as cointegrated time series. While this is the first step for our modeling strategy, we first examine the properties of the daily high and low log-prices employing $p_t^H = \log(P_t^H)$ and $p_t^L = \log(P_t^L)$, and, as defining their difference, the

[5] Figures for all of the remaining series can be found in the Appendix (Figure 3 for the DAX, Figure 4 for the FTSE 100, Figure 5 for the NIKKEI 225, and finally, Figure 6 for the S&P 500).

range as $R_t = p_t^H - p_t^L$. We also define a vector $X_t \equiv (p_t^H, p_t^L)'$, which will be used throughout the remainder of the study.
The Augmented Dickey-Fuller (ADF) test for daily high and low prices reported in the Appendix confirm the expected result, that the series are unit root processes. Their difference, the range, is stationary; thus, daily high and low prices are cointegrated, as proposed by Cheung (2007). There are, however, two exceptions. We find that the daily high and low prices of the DAX and FTSE 100 in the first period are at the boundary of trend-stationarity. Figure 2 displays the autocorrelation function (ACF) of the DAX range for all of the examined periods. We employ this plot to motivate the need to improve the analysis of Cheung (2007). While the ADF test is designed to test for the presence of a unit root against the I(0) alternative, it has very low power against fractional processes. Despite the stationarity of the range in all periods, the ACF of the range displays a high degree of persistence. As motivated by Caporin et al. (2013), the simple cointegration analysis may not be satisfactory in explaining the relationship between high and low prices. ACF plots for all remaining tested series show similar behavior and are reported in the Appendix; all of the autocorrelations are significant even after 40 lags, except for the PX Index in the first period.

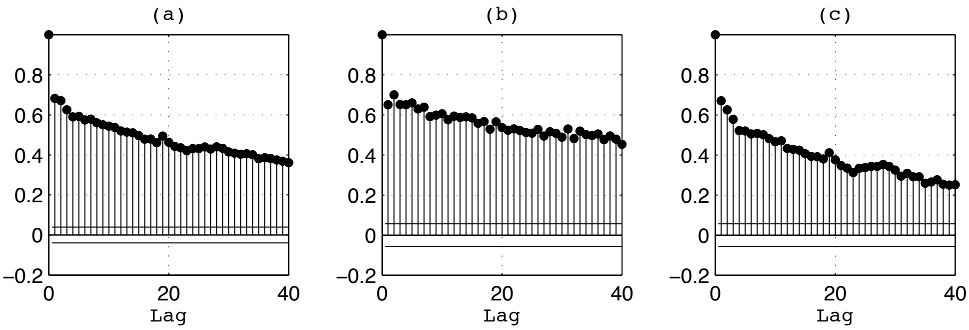

Figure 2: ACF of the DAX range in (a) 2003 – 2012, (b) 2003 – 2007 and (c) 2008 – 2012

In addition, we employ the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test of Kwiatkowski et al. (1992), which has greater power in situations when the tested series are close to being a unit root. The results reported in the Appendix confirm non-stationarity for daily high and low prices. However, the results are different for the ranges that are non-stationary, according to the KPSS test as well. While the ranges seem to be stationary according to the ADF test, we believe that this result points to the presence of long memory in the ranges. Hence, long memory needs to be utilized as a proper framework for modeling the high and low prices relationship.

3.1 Long memory properties

While Cheung (2007) first introduced the idea of modeling daily high and low prices as a cointegrated relationship, Caporin et al. (2013) first noted that the "error correction" term, the range, arising from this analysis may contain long memory. Thus, they proposed a fractionally cointegrated model to capture this feature. The preliminary analysis of our data set confirms the need to generalize to the fractional cointegration framework. Before we introduce the actual analysis, we briefly introduce the basic notion of the models here.
Long-memory models have been used by the natural sciences (specifically, hydrology and climatology) since the 1950s. They drew the attention of econometricians in approximately 1980, when Granger (1980) and Granger and Joyeux (1980) developed the autoregressive fractionally integrated moving average (ARFIMA) and Geweke and Porter-Hudak (1983) proposed a technique for estimating the long memory parameter. Later, the fractionally integrated generalized autoregressive conditional heteroskedasticity (GARCH) (p, d, q) (FIGARCH) was proposed by Baillie (1996) to capture the slowly decaying autocorrelation

functions of volatility. Substantial evidence of the presence of long memory has been demonstrated in the literature on financial data such as the volatility of asset prices, interest rate differentials, inflation rates, forward premiums, and exchange rates (Baillie, 1996; Ding, Granger and Engel, 1993; Andersen and Bollerslev, 1997; Breidt et al., 1998; Kellard et al., 2010; or Garvey and Gallagher, 2012).

We can define a long-memory process as a weakly stationary process with an autocorrelation function $\rho(\cdot)$ having hyperbolic decay (Brockwell and Davis, 1991), $\rho(h) \sim C \cdot h^{2d-1}$ as $h \to \infty$,

where $C \neq 0, d < 0.5$. In contrast, a weakly stationary process has short memory if its autocorrelation function is geometrically bounded.

Consider a vector X_t holding $I(1)$ elements as a cointegrated vector if there exists a linear combination $\beta' X_t$ that is an $I(0)$ process. Robinson and Yajima (2002) note that the possible existence of a long-run, stable relationship among non-stationary series X_t does not depend on whether the series are $I(1)$. The need for a flexible approach is solved by considering an $I(d)$ series, i.e., a series integrated of order d, with a real-valued d. Robinson and Yajima (2002) define the series X_t as an $I(d)$ process if $u_t = (1-L)^d X_t$ is $I(0)$, where L is the lag operator and $d < 0.5$. For $d \geq 0.5$, we define a non-stationary $I(d)$ series $X_t = (1-L)^{-d} u_t \mathrm{I}\{t \geq 1\}, t = 0, \pm 1, \pm 2, \ldots$, where I( ) is the indicator function (for details, see, e.g., Shimotsu and Phillips, 2005). If $d > 0$, we say that the process has long memory, and if $d < 0$, we say that the process is anti-persistent. One can easily see that if $d = 1$, then the process represents a random walk, and if $d = 0$,, then the process is stationary. The parameter d is called the fractional differencing parameter, fractional degree of persistence or fractional order of integration, and it describes the memory properties of X_t (Robinson and Yajima, 2002).

	Bandwidths			$ELW_{m=T^{0.5}}$				$ELW_{m=T^{0.6}}$			
	T	$T^{0.5}$	$T^{0.6}$	$\hat{d}_H$	$\hat{d}_L$	$\hat{d}_R$	s.e.	$\hat{d}_H$	$\hat{d}_L$	$\hat{d}_R$	s.e.
S&P 500											
2003-2012	2517	50	109	1.0972	1.0709	**0.7626**	0.0707	1.0442	1.0036	**0.6216**	0.0479
2003-2007	1258	35	72	0.9094	0.8547	**0.5870**	0.0845	0.9265	0.8932	**0.5740**	0.0589
2008-2012	1259	35	72	1.1428	1.0853	**0.6643**	0.0845	1.0512	1.0130	**0.6867**	0.0589
FTSE 100											
2003-2012	2526	50	110	0.9990	0.9763	**0.6575**	0.0707	0.9673	0.9369	**0.6324**	0.0477
2003-2007	1264	35	72	0.8518	0.8182	**0.6220**	0.0845	0.8539	0.8357	**0.6028**	0.0589
detrending				0.5765	0.5877			0.8138	0.8011		
2008-2012	1262	35	72	1.0377	1.0026	**0.6278**	0.0845	0.9799	0.9568	**0.6169**	0.0589
DAX											
2003-2012	2556	50	110	1.0221	1.0055	**0.5884**	0.0707	1.0434	1.0058	**0.5887**	0.0477
2003-2007	1274	35	72	0.9711	0.9158	**0.6142**	0.0845	0.9817	0.9461	**0.5968**	0.0589
detrending				0.9076	0.8451			0.9630	0.9289		
2008-2012	1282	35	73	1.0913	1.0766	**0.7001**	0.0845	1.0237	1.0057	**0.6229**	0.0585
NIKKEI 225											
2003-2012	2543	49	108	1.0572	1.0372	**0.4833**	0.0714	1.0564	1.0428	**0.6493**	0.0481
2003-2007	1229	35	71	0.9909	0.9630	**0.4955**	0.0845	0.9755	0.9447	**0.3916**	0.0593
2008-2012	1224	34	71	1.1082	1.1018	**0.6131**	0.0857	1.0697	1.0544	**0.6044**	0.0593
PX											
2003-2012	2494	49	109	1.0990	1.0876	**0.5158**	0.0714	1.1659	1.1367	**0.5197**	0.0479
2003-2007	1235	35	71	0.8911	0.8601	**0.2811**	0.0845	1.0174	0.9659	**0.4455**	0.0593
2008-2012	1259	35	72	1.2337	1.2189	**0.5723**	0.0845	1.1216	1.0986	**0.5165**	0.0589

Table 1: *Exact local Whittle (ELW) estimator of the fractional degree of integration parameter d based on the 2-step ELW estimator for the high prices ($\hat{d}_H$), low prices ($\hat{d}_L$), and their difference, the range ($\hat{d}_R$). We consider two bandwidths m, determining the number of periodogram ordinates employed in estimation equal to*

$m=T^{0.5}$ and $m=T^{0.6}$, similar to Nielsen and Shimotsu (2007). For the possibly trend-stationary series (DAX in the first period and FTSE 100 in the first period), we also include the results of the 2-step ELW estimator with prior detrending.

There are several methods of estimating the fractional degree of persistence. In our analysis, we employ two semi-parametric methods, a univariate exact local Whittle estimator (ELW) proposed by Shimotsu and Phillips (2005) and a procedure proposed by Geweke and Porter-Hudak (GPH) (1983). Because we are estimating the order of integration for two series (daily high and low prices), we also test for the equality of integration orders later in this study because this is a condition for further analysis.

The exact local Whittle (ELW) estimator proposed by Shimotsu and Phillips (2005) is consistent in both situations, when cointegration is present as well as absent in the tested series. This estimator is also applicable to both stationary and non-stationary cases, which removes the limitation of the original tests proposed by Robinson and Yajima (2002) working under stationary data only. It also improves the original local Whittle (LW) estimator whose asymptotic theory is discontinuous for the values of $d = 3/4$ and $d = 1$.

The results in Table 1 support our two initial hypotheses. First, the daily highs and lows are not stationary. The order of integration is generally close to 1; however, in the full period and in the second sub-period, unitary integration is substantially exceeded. Only in 25 cases out of 60 is the integration order of daily highs and lows smaller than 1. The lowest values of integration of daily highs and lows are achieved in the first period. The DAX and the FTSE 100 remain non-stationary even for the detrended case, which is in line with the KPSS test, and hence we may conclude that high and low prices are non-stationary. The difference between high and low prices (the range) is mostly non-stationary and displays long memory with parameter $d_R > 0.5$. As the robustness check, we also compute the fractional degree of integration on the data excluding the years 2008 and 2009. These data display order of integration that is lower than the second period but higher than the first period. Because the different sample lengths make it difficult to compare the results statistically, we do not present the full results here, but we can provide them upon request instead.

The results for the PX Index are different from the other indices because the range is in the stationary region in the first period. This finding is also true for the Japanese NIKKEI 225, but not for the other indices because its range is non-stationary in all of the studied periods. This finding confirms the observations from the ACF of the PX and NIKKEI 225.

	Bandwidths			$GPH_{m=T^{0.5}}$			$GPH_{m=T^{0.6}}$			
	T	$T^{0.5}$	$T^{0.6}$	$\hat{d}_H$	$\hat{d}_L$	$\hat{d}_R$	$\hat{d}_H$	$\hat{d}_L$	$\hat{d}_R$	s.e.
S&P 500										
2003-2012	2517	50	109	1.0261	1.0202	**0.7646**	1.0819	1.0593	**0.6742**	0.0256
2003-2007	1258	35	72	1.0921	1.0819	**0.5204**	1.0914	1.0945	**0.6778**	0.0362
2008-2012	1259	35	72	1.2206	1.2483	**0.7526**	1.0390	1.0398	**0.7211**	0.0361
FTSE 100										
2003-2012	2526	50	110	1.0187	1.0095	**0.7095**	1.0524	1.0306	**0.7123**	0.0255
2003-2007	1264	35	72	1.0299	1.0157	**0.5145**	1.1522	1.1606	**0.5960**	0.0361
detrending				0.6543	0.6425		0.6674	0.6732		
2008-2012	1262	35	72	1.1200	1.0774	**0.7387**	0.9274	0.9292	**0.7010**	0.0361
DAX										
2003-2012	2556	50	110	1.0756	1.0647	**0.6518**	1.1018	1.1023	**0.7059**	0.0254
2003-2007	1274	35	72	1.0695	1.0623	**0.8031**	1.0768	1.0819	**0.8264**	0.0359
detrending				0.7396	0.7131		0.8819	0.8758		
2008-2012	1282	35	73	1.1732	1.1692	**0.7099**	0.9912	0.9828	**0.7384**	0.0358
NIKKEI 225										
2003-2012	2543	49	108	1.0696	1.0841	**0.4989**	1.0753	1.0869	**0.6589**	0.0254

2003-2007	1229	35	71	1.0403	1.0310	**0.4627**	1.0703	1.0590	**0.4013**	0.0366
2008-2012	1224	34	71	1.2168	1.1852	**0.7009**	1.0606	1.0344	**0.7841**	0.0367
PX										
2003-2012	2494	49	109	1.0653	1.0673	**0.5765**	1.0871	1.0859	**0.6167**	0.0257
2003-2007	1235	35	71	1.0166	1.0159	**0.2106**	1.0158	1.0128	**0.3155**	0.0365
2008-2012	1259	35	72	0.8935	0.9182	**0.7603**	0.8478	0.8336	**0.5805**	0.0361

Table 2: *GPH estimator of the fractional degree of integration for the high prices ($\widehat{d}_H$), low prices ($\widehat{d}_L$), and their difference, the range ($\widehat{d}_R$). The bandwidth parameters are chosen in the same manner as for the ELW estimator.*

Table 2 presents complimentary results obtained with the Geweke and Porter-Hudak GPH) (1983) estimator. In the first period, the order of integration for the prices of FTSE 100 is approximately 0.65 in both specifications, and in the case of the DAX, it is even higher. This finding supports the previous results; hence, for the remainder of the study, we consider these two series non-stationary and employ them in their original form (without detrending). The rest of the GPH estimates generally support the previous findings. First, the integration orders of daily highs and lows are close to 1, and second, the range is not integrated of order 0 and displays long memory.

However, we can observe some differences between the two estimation procedures. The GPH estimates of the integration orders of daily highs and lows utilizing each bandwidth are much closer to each other than the ELW estimates. However, if we look at the first (calmer) period, the ELW estimates the integration orders of daily highs and lows below 1, but GPH estimates are mostly greater than 1 (overall, in only 8 cases out of 60 are the GPH integration orders of daily highs and lows smaller than 1). Surprisingly, from the GPH results for the $m = T^{0.6}$ bandwidth, one would believe that the prices in the second (after-crisis) period are closer to 0.5, indicating a stationarity region, than in the first (pre-crisis) period, although they remain non-stationary. If we compare this result with results using the $m = T^{0.5}$ bandwidth, we arrive at the opposite conclusion.

The estimates of integration orders of the range vary substantially between the two specifications of bandwidth. Similarly to the ELW estimator, in the case of the NIKKEI 225 for the full period, the choice of bandwidth changes the conclusion about the stationarity of the range. In 24 cases out of 30, the GPH estimator provides higher estimates of the long memory of the range than the ELW estimator. Among the indices, the integration orders of range vary substantially, from 0.21 for the PX Index in the first period up to 0.83 for the DAX in the same period. The ranges of the PX Index and of the NIKKEI 225 are definitely stationary in the first period and non-stationary in the others. The ranges of the other indices are non-stationary in all periods. One can also observe the large difference between the ELW and GPH estimates of the integration order of the range of the DAX in the first period: the GPH estimate is more than 30% higher than the ELW estimate with the same bandwidth specification. The GPH estimate is more consistent with the findings from the inspection of the DAX autocorrelation function than the ELW estimate.

Overall, we can conclude that the results are sensitive to both the selected estimator (ELW or GPH) and to the bandwidth parameter. However, it is clear that the daily high and low prices are not stationary and display long memory. The range also displays long memory, but in some cases, it is in the stationary region.

3.2 Testing the equality of integration orders

The presence or absence of cointegration is not known when we estimate the fractional integration orders. Nielsen and Shimotsu (2007) present the possibility of testing the equality of integration orders by designing two test hypotheses: the pairwise equality of the

integration orders and the equality of all of the integration orders. For each hypothesis, a test statistic is defined. However, in our bivariate case, these two possible hypotheses collapse into one, allowing us to only focus on the hypothesis of equality of all of the integration orders: $\mathscr{H}_0: d_H = d_L = d_*$.
Under the null hypothesis, Nielsen and Shimotsu (2007) prove that, if the variables are cointegrated (i.e., their cointegration rank is $r = 1$), then the test statistic $\hat{T}_0$ should converge in probability to 0. However, if they are not cointegrated ($r = 0$), then under the null hypothesis, the test statistic $\hat{T}_0$ should converge in distribution to the chi-quadrat distribution. This relationship means that, if the value of the test statistic $\hat{T}_0$ is significantly large with respect to the chi-quadrat, then we can take this as evidence that the null of the equality of integration orders is rejected.
Table 3 presents the test statistics for testing the equality of integration orders of daily high and low prices. For each index and time period, we present two test statistics; the first is estimated utilizing the $T^{0.5}$ bandwidth, and the second is estimated utilizing the $T^{0.6}$ bandwidth. Because the maximum test statistic is 1.3361 and the lowest critical value of chi-quadrat distribution with one degree of freedom is 2.71 (for the 90% confidence interval), we cannot reject the null hypothesis of equality of the integration orders for all tested series. This implies that we can perform the FCVAR estimation with the same degree of integration orders $d_H = d_L$.

	Bandwidths			$\hat{T}_0$ *statistics*	
	T	$T^{0.5}$	$T^{0.6}$	$\hat{T}_0(m = T^{0.5})$	$\hat{T}_0(m = T^{0.6})$
S&P 500					
2003-2012	2517	50	109	0.2636	1.3361
2003-2007	1258	35	72	0.7113	0.5400
2008-2012	1259	35	72	0.8033	0.7043
FTSE 100					
2003-2012	2526	50	110	0.1979	0.7689
2003-2007	1264	35	72	0.2716	0.1596
2008-2012	1262	35	72	0.3005	0.2635
DAX					
2003-2012	2556	50	110	0.1041	1.1610
2003-2007	1274	35	72	0.7085	0.6007
2008-2012	1282	35	73	0.0525	0.1604
NIKKEI 225					
2003-2012	2543	49	108	0.1493	0.1517
2003-2007	1229	35	71	0.1919	0.4660
2008-2012	1224	34	71	0.0096	0.1129
PX					
2003-2012	2494	49	109	0.0488	0.6988
2003-2007	1235	35	71	0.2329	1.2668
2008-2012	1259	35	72	0.0538	0.2520

Table 3: *Test statistics for the equality of integration orders*.

4 An empirical model of fractionally cointegrated daily high and low prices

Preliminary analysis suggests that the range displays long memory and that the integration orders of daily high and low prices are the same; hence, we can continue with the empirical fractionally cointegrated VAR model for the daily highs and lows. The fractionally cointegrated vector error correction model (FVECM) or fractionally cointegrated VAR

(FCVAR) was discussed in the work of Granger (1986) and formalized recently by Johansen (2008) and Johansen and Nielsen (2010, 2012). The main distinction from the classical cointegration analysis is that the generalized model allows X_t to be fractional of order d and cofractional of order $d-b$; that is, $\beta' X_t$ should be fractional of order $d-b \geq 0$. In other words, fractional cointegration assumes the existence of a common stochastic trend, which is integrated of order d, and the short-term departures from the long-run equilibrium being integrated of order $d-b$.

Following Johansen and Nielsen (2012) and Nielsen and Morin (2012), we describe the model in two steps. First, the usual lag operator and the difference operator are replaced by the fractional lag operator and fractional difference operator, $L_b = 1 - \Delta^b$ and $\Delta^b = (1-L)^b$, respectively. The fractional difference operator is defined by the binomial expansion $\Delta^b Z_t = \sum_{n=0}^{\infty}(-1)^n \binom{b}{n} Z_{t-n}$. Second, the resulting model is applied to $Z_t = \Delta^{d-b} X_t$. Thus, a fractionally cointegrated $\text{VAR}_{\text{d,b}}(\text{p})$ model for a vector of high and low prices $X_t \equiv (p_t^H, p_t^L)'$, is defined as

$$\Delta^d X_t = \Delta^{d-b} L_b \alpha \beta' X_t + \sum_{i=1}^{p} \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t, \qquad t = 1, \ldots T,$$

where ε_t is an $i.i.d.(0,\Omega)$, with Ω positive-definite variance matrix, and α and β are $2 \times r$ matrices, $0 \leq r \leq 2, d \geq b > 0$. Non-zero mean data, e.g., $Y_t = \mu + X_t$ can be modeled as $\Delta^a Y_t = \Delta^a(\mu + X_t) = \Delta^a X_t$ because $\Delta^a 1 = 0$ for $a > 0$. Y_t and thus satisfies the same equations as $Y_t = \mu + X_t$. This relationship means that the model with $d > b$ is invariant to the inclusion of a restricted constant term ρ. Therefore, we consider the inclusion of a constant term only in the model with $d = b$:

$$\Delta^d X_t = L_d \alpha(\beta' X_t + \rho') + \sum_{i=1}^{p} \Gamma_i \Delta^d L_d^i X_t + \varepsilon_t, t = 1, \ldots T,$$

Both models include the standard cointegrated VAR model as the special case when $d = b = 1$. The cointegration as well as adjustment towards equilibrium is more general because the model incorporates both fractional integration and cointegration. X_t is integrated of order d, and b is the strength of the cointegrating relationships (a higher b means less persistence in the cointegrating relationships; b can also be called the cointegration gap). Moreover, if $d - b < 1/2$, then $\beta' X_t$ is asymptotically a zero-mean stationary process. If we write $\Pi = \alpha\beta'$, where the $2 \times r$ matrices α and β with $r \leq 2$ are assumed to have full column rank r, the columns of β are then the r cointegrating (cofractional) relationship determining the long-run equilibrium. The rank r is called the cointegration or cofractional rank, the parameter α determines the speed of adjustment towards the equilibrium, the parameters $\Gamma = (\Gamma_1, \ldots, \Gamma_p)$ govern the short-run dynamics, and the parameter ρ is the restricted constant term (because the constant term in the model is restricted to be of the form $\mu = \alpha\rho'$) and is interpreted as the mean level of the long-run equilibrium. In the special case when $d = b$, $(\beta' X_t + \rho')$ is a zero-mean process of fractional order zero. The model parameters are estimated using the procedure[6] outlined in Johansen and Nielsen (2012).

4.1 Testing cointegration rank

The traditional tests for cointegration proposed by Johansen (1991) are not applicable in the presence of long memory. Instead, more recent tests allowing for fractional cointegration should be applied. Time series X_t is said to be fractionally cointegrated $CI(d,b)$ if X_t has $I(d)$ elements and for some $b > 0$, there exists β such that $\beta' X_t$ is integrated of order $(d-b)$. In our work, we use two cointegration rank tests proposed by Nielsen and Shimotsu (2007) and

[6] For the estimation, we employ the software available from the authors Nielsen and Morin (2012).

Johansen and Nielsen (2012). The cointegration rank determination procedure proposed by Nielsen and Shimotsu (2007) extends the procedure found in Robinson and Yajima (2002) because it allows for both stationary and non-stationary fractionally integrated processes. The ability to consider any value of the fractional differencing parameter d follows from the application of the exact local Whittle analysis of Shimotsu and Phillips (2005). The exact local Whittle estimate of d is then employed to examine the rank of the spectral density matrix of the d'th differenced process around the origin to provide a consistent estimate of the cointegration rank. This semi-parametric method only requires information about the behavior of the spectral density matrix around the origin, but it also relies on the choice of the bandwidth and threshold parameters. However, this approach does not require the cointegrating vectors to be estimated to determine the cointegration rank.

Table 4 summarizes the results indicating that there is one cointegrating relationship. $L(0)$ and $L(1)$ are the values of the loss function evaluated with regard to a cointegration rank of 0 or 1. The cointegration rank $\hat{r}$ is then determined by the $arg\,min$ of $L(u)$. In all cases, the loss function $L(1)$ is smaller than $L(0)$, implying that there is exactly one fractionally cointegrating relationship between the daily high and low prices.

In addition, we also utilize the testing procedures of Johansen and Nielsen (2012) and Nielsen and Morin (2012). In the fractionally cointegrated VAR model, we test the hypothesis $\mathcal{H}_r : rank(\Pi) = r$ against $\mathcal{H}_n : rank(\Pi) = n$. Let $L(d, b, r)$ be the profile likelihood function given a rank r, where (α, β, Γ) have been concentrated out by regression and reduced rank regression (see Johansen and Nielsen, 2012, p. 23). In the case of the model with a constant, we test $\mathcal{H}_r : rank(\Pi, \mu) = r$ against $\mathcal{H}_n : rank(\Pi, \mu) = n$, and the profile likelihood function given rank r is then $L(d, r)$, where again the parameters $(\alpha, \beta, \rho, \Gamma)$ have been concentrated out.[7]

	Eigenvalues			Rank estimates					
				$\upsilon(T) = m_1^{-0.45}$			$\upsilon(T) = m_1^{-0.05}$		
	$\bar{d}_*$	δ_1	δ_2	$L(0)$	$L(1)$	$\hat{r}$	$L(0)$	$L(1)$	$\hat{r}$
S&P 500									
2003-2012	1.0239	0.3169	0.0010	-1.7117	-1.8491	1	-0.3872	-1.1868	1
2003-2007	0.9099	0.2151	0.0009	-1.6561	-1.8197	1	-0.3553	-1.1694	1
2008-2012	1.0321	0.5076	0.0022	-1.6561	-1.8189	1	-0.3553	-1.1686	1
FTSE 100									
2003-2012	0.9521	0.4360	0.0010	-1.7117	-1.8511	1	-0.3872	-1.1889	1
2003-2007	0.8448	0.3286	0.0014	-1.6561	-1.8194	1	-0.3553	-1.1691	1
2008-2012	0.9684	0.6313	0.0018	-1.6561	-1.8222	1	-0.3553	-1.1718	1
DAX									
2003-2012	1.0246	0.5136	0.0017	-1.7117	-1.8491	1	-0.3872	-1.1869	1
2003-2007	0.9639	0.4063	0.0025	-1.6591	-1.8172	1	-0.3569	-1.1662	1
2008-2012	1.0147	0.7520	0.0029	-1.6591	-1.8217	1	-0.3569	-1.1706	1
NIKKEI 225									
2003-2012	1.0496	0.4847	0.0011	-1.7099	-1.8504	1	-0.3861	-1.1885	1
2003-2007	0.9601	0.4551	0.0010	-1.6561	-1.8235	1	-0.3553	-1.1731	1
2008-2012	1.0621	0.6491	0.0021	-1.6529	-1.8199	1	-0.3537	-1.1703	1
PX									
2003-2012	1.1513	0.3563	0.0009	-1.7099	-1.8496	1	-0.3861	-1.1877	1
2003-2007	0.9917	0.4132	0.0016	-1.6561	-1.8200	1	-0.3553	-1.1696	1
2008-2012	1.1101	0.5811	0.0035	-1.6561	-1.8159	1	-0.3553	-1.1656	1

Table 4: *Fractional cointegration rank test by Nielsen and Shimotsu (2007). We report the common integration order $\bar{d}_*$, which is used in the fractional cointegration analysis and is simply computed as an average of the estimated integration orders of daily high and low prices from the ELW estimator based on a given bandwidth. Eigenvalues of the estimated statistics are reported as well.*

[7] The model with an inclusion of a constant is considered only when and that is why the profile likelihood function depends only on the given rank and the parameter .

We maximize the profile likelihood function under both hypotheses $\mathcal{H}_r$ and $\mathcal{H}_n$; the likelihood ratio (LR) statistic is $LR_T(q) = 2\log(L(\hat{d}_n, \hat{b}_n, n)/L(\hat{d}_r, \hat{b}_r, r))$, where $q = n - r$ and $L(\hat{d}_n, \hat{b}_n, n) = \max_{d,b} L(d, b, n)$, $L(\hat{d}_r, \hat{b}_r, r) = \max_{d,b} L(d, b, r)$, similarly for the model with a constant. The asymptotic distribution of $LR_T(q)$ depends both qualitatively and quantitatively on the parameter b and on $q = n - r$. This dependence on the unknown parameter b makes the empirical analysis more complicated; however, MacKinnon and Nielsen (2012) provide asymptotic critical values for the LR rank test. In the case of "weak cointegration," when $0 < b < 1/2$, $LR_T(q)$ has a standard asymptotic distribution, $LR_T(q) \xrightarrow{D} \chi^2(q^2)$. The situation is, however, different when $1/2 < b \leq d$. Then, the asymptotic theory is nonstandard and

$$LR_T(q) \xrightarrow{D} \mathrm{Tr}\left\{\int_0^1 dW(s)F'(s)\left(\int_0^1 F(s)F'(s)ds\right)^{-1}\int_0^1 F(s)dW'(s)\right\},$$

where the vector process dW is the increment of the ordinary vector Brownian motion of dimension $q = n - r$. The vector F depends on deterministics in a similar way as the CVAR model in Johansen (1996). If we do not include any deterministic terms in the model, then $F(u) = W_b(u)$. If the restricted constant term is included in the model, then $F(u) = (W_b'(u), 1)'$, where $W_b(u) = \Gamma(b)^{-1}\int_0^u (u - s)^{b-1}dW(s)$ is vector fractional type-II Brownian motion.

	$r = 0$					$r = 1$				$r = 2$	
	IV	$\hat{d}$	$\hat{b}$	LR	$CV_{5\%}$	$\hat{d}$	$\hat{b}$	LR	$CV_{5\%}$	$\hat{d}$	$\hat{b}$
S&P 500											
2003-2012	70	0.681	0.372	271.232	9.490	**0.998**	**0.302**	0.058	3.840	1.002	0.295
2003-2007	80	0.517	0.517	117.535	9.362	**0.999**	**0.467**	1.338	3.840	0.984	0.481
2008-2012	60	0.719	0.329	136.629	9.490	**0.981**	**0.334**	1.699	3.840	1.033	0.233
FTSE 100											
2003-2012	80	0.623	0.416	254.914	9.490	**0.970**	**0.357**	2.368	3.840	0.954	0.387
2003-2007	70	0.497	0.497	118.642	9.490	**0.984**	**0.380**	2.510	3.840	0.968	0.424
2008-2012	30	0.667	0.361	120.148	9.490	**0.987**	**0.463**	2.035	3.840	1.009	0.416
DAX											
2003-2012	60	0.609	0.439	233.689	9.490	**0.987**	**0.388**	3.808	3.840	0.966	0.408
2003-2007	50	0.561	0.524	38.776	9.359	**1.042**	**0.476**	0.323	3.840	1.055	0.480
2008-2012	80	0.638	0.405	133.836	9.490	**0.967**	**0.336**	0.302	3.840	0.955	0.361
NIKKEI 225											
2003-2012	80	0.591	0.476	138.368	9.490	**1.004**	**0.517**	0.254	3.636	0.997	0.515
2003-2007	60	0.513	0.513	80.636	9.365	**1.019**	**0.635**	0.003	3.587	1.018	0.635
2008-2012	60	0.708	0.370	79.292	9.490	**0.996**	**0.010**	28.181	3.840	0.978	0.547
PX											
2003-2012	80	0.520	0.520	109.031	9.360	**1.007**	**0.475**	1.289	3.840	0.988	0.472
2003-2007	60	0.532	0.532	65.024	9.360	**1.021**	**0.680**	5.991	6.373*	0.977	0.666
2008-2012	60	0.511	0.511	53.246	9.367	**0.982**	**0.454**	1.092	3.840	1.010	0.451

Table 5: *Cointegration rank test by Johansen and Nielsen (2012). The first column, IV, stands for the number of initial values used in the estimation. For each rank* $r = 0, 1, 2,,$ *we present the estimates of the parameter of the fractional order of integration* ($\hat{d}$)*, the parameter of the cointegration gap* ($\hat{b}$)*, and the corresponding likelihood ratio statistic* (LR) *and its critical value at a 5% level of significance. When b is smaller than 0.5, it follows the* $\chi^2(q^2)$ *distribution; this means that for cointegration rank* $r = 0$*,* $q = 2$ *and* $\chi^2_{0.95}(4) = 9.49$*. When* $r = 1$*, then* $q = 1$ *and* $\chi^2_{0.95}(1) = 3.84$*. If b is greater than 0.51, we use MacKinnon and Nielsen (2012) for critical values. Asterisk (*) denotes the 1% critical value rather than the 5% critical value. The cointegrating relationship is not significant at the 5% level of significance, but it is significant at the 1% level of significance.*

The results of the cointegration rank test by Johansen and Nielsen (2012) are presented in Table 5. We find one significant cointegrating relationship, except for the NIKKEI 225 in the second period, where no cointegrating vector is found. In the case of the PX Index in the first period, the LR statistic for one cointegrating vector ($r = 1$) is significant at the 1% level. When $r = 0$, the likelihood ratio (LR) statistic is significantly larger than the corresponding critical value, meaning that we reject the null hypothesis of zero cointegrating relationships When $r = 1,,$, the LR statistic is significantly smaller than the corresponding critical value, and thus, we do not reject the null of one cointegrating relationship. This finding is not true for the NIKKEI 225 in the second period, where we do not find a significant cointegrating vector, even though its presence was hinted at by the results of the cointegration rank test proposed by Nielsen and Shimotsu (2007).

4.2 FCVAR model of daily high and low prices

Having in mind the result that in each specification there is one significant cointegrating vector (except for the NIKKEI 225 in the second period), we build an FCVAR for the daily high and low prices. We employ one lag for the short-term deviations $p = 1$ because it sufficiently captures the autocorrelation of residuals. MacKinnon and Nielsen (2012) state that a single lag is usually sufficient in the fractional model, which is in contrast with the standard cointegrated VAR, where several more lags are needed to capture the serial correlation in residuals. The number of initial values utilized for the estimation of the FCVAR model is the same number as the number of initial values utilized for the estimation of the cointegration rank in the previous section. Next, we consider restricting the order of integration of the range. Because the range is defined as the difference between the maximum and minimum daily prices, i.e., $(p_t^H - p_t^L)$, we would like the cointegrating vector to be $(1, -1)$. If the cointegrating vector is different from $(1, -1)$, we cannot interpret the difference $(d - b)$ as the order of integration of the range. That is why we first estimate the model without any restrictions imposed to see whether the model yields a significant cointegrating vector and significant estimates of $\hat{d}$ and $\hat{b}$, and then we impose the $(1, -1)$ restriction on the cointegrating vector. However, when a restriction is imposed, the standard errors are not provided; thus, we cannot make any inference regarding the significance of the estimated parameters. We estimate the model for the case when $d \neq b$; however, the procedure is capable of detecting whether d and b are close to equality; if they were, we would re-estimate the model with the restriction $d = b$. This situation did, however, not occur, which is consistent with our previous findings because the equality of the d and b parameters would imply that the order of integration of the range is 0, which we previously rejected based on both the ELW and GPH estimator results.

Table 6 presents the results from the FCVAR estimation without any restrictions imposed. In all of the specifications, the parameters of interest ($\hat{d}$ and $\hat{b}$) are significantly different from zero and different from each other (thus $d \neq b$). Additionally, the estimates of the cointegrating vector $\hat{\beta}$ are very close to the desired vector of $(1, -1)$. The results suggest that a linear combination of the daily high and low prices (the range) is integrated of a non-zero order. However, because the estimate of the cointegrating vector is not exactly $(1, -1)$, we cannot interpret the difference $(d - b)$ as the order of integration of the range.

		$\hat{d}$	$\hat{b}$	$\hat{\beta}$	α_H	α_L	γ_{11}	γ_{12}	γ_{21}	γ_{22}
S&P 500										
u	2003-2012	0.998	0.302	(1,-1.007)	-0.195	3.986	-0.773	1.050	-1.614	2.619
		(0.014)	(0.028)		(0.302)	(0.945)	(0.310)	(0.316)	(0.727)	(0.776)
	2003-2007	0.999	0.467	(1,-1.002)	-1.072	1.680	0.153	0.076	-0.282	0.924
		(0.009)	(0.038)		(0.254)	(0.363)	(0.178)	(0.193)	(0.258)	(0.283)
u	2008-2012	0.981	0.334	(1,-1.008)	-0.013	3.081	-0.755	1.049	-0.935	1.890
		(0.015)	(0.035)		(0.348)	(0.922)	(0.363)	(0.370)	(0.696)	(0.752)
FTSE 100										
	2003-2012	0.970	0.357	(1,-1.005)	-0.657	2.308	-0.263	0.569	-0.620	1.505
		(0.011)	(0.024)		(0.218)	(0.466)	(0.190)	(0.204)	(0.358)	(0.386)
u	2003-2007	0.984	0.380	(1,-1.003)	-1.344	2.339	0.245	-0.009	-0.994	1.727
		(0.009)	(0.027)		(0.331)	(0.513)	(0.252)	(0.280)	(0.426)	(0.442)
	2008-2012	0.987	0.463	(1,-1.004)	-0.429	1.344	-0.309	0.524	-0.049	0.669
		(0.009)	(0.031)		(0.189)	(0.305)	(0.162)	(0.171)	(0.231)	(0.247)
DAX										

2003-2012	0.987	0.388	(1,-1.004)	-0.436	1.891	-0.332	0.571	-0.433	1.101
	(0.011)	(0.033)		(0.209)	(0.440)	(0.195)	(0.202)	(0.327)	(0.354)
2003-2007	1.042	0.476	(1,-1.001)	-0.551	1.520	-0.145	0.319	-0.169	0.526
	(0.013)	(0.064)		(0.259)	(0.429)	(0.194)	(0.203)	(0.279)	(0.305)
2008-2012	0.967	0.336	(1,-1.007)	-0.016	2.651	-0.695	0.980	-0.910	1.718
	(0.016)	(0.035)		(0.393)	(0.846)	(0.413)	(0.415)	(0.675)	(0.711)
NIKKEI 225									
2003-2012	1.004	0.517	(1,-1.002)	-0.175	1.318	-0.295	0.535	-0.066	0.543
	(0.012)	(0.043)		(0.134)	(0.271)	(0.124)	(0.133)	(0.183)	(0.205)
2003-2007	1.019	0.635	(1,-1.001)	-0.361	1.144	-0.154	0.309	-0.171	0.522
	(0.010)	(0.053)		(0.151)	(0.242)	(0.120)	(0.125)	(0.168)	(0.186)
2008-2012	-	-	-	-	-	-	-	-	-
PX									
2003-2012	1.007	0.475	(1,-1.003)	-0.682	1.281	-0.172	0.340	-0.447	0.861
	(0.017)	(0.045)		(0.168)	(0.310)	(0.127)	(0.136)	(0.236)	(0.258)
2003-2007	1.021	0.680	(1,-1.002)	-0.410	0.612	0.021	0.204	0.123	0.243
	(0.015)	(0.065)		(0.119)	(0.168)	(0.095)	(0.096)	(0.120)	(0.130)
2008-2012	0.982	0.454	(1,-1.003)	-0.732	1.346	-0.272	0.453	-0.600	1.031
	(0.017)	(0.056)		(0.253)	(0.445)	(0.194)	(0.211)	(0.355)	(0.389)

Table 6: *FCVAR estimation results (no restrictions). Note: Standard errors are given in the brackets, and "u" denotes that the model is unstable (some of the roots of the characteristic polynomial lie outside the unit root circle).*

Daily high and low prices are integrated of an order close to 1. Surprisingly, in the first period, which we consider the calmer period, daily prices are further from stationarity than in the second period or in the full period. Additionally, the orders of integration of daily prices are smaller than unity in 9 out of 14 cases.

The adjustment coefficients α_H and α_L capturing the speed of adjustment of p_t^H and p_t^L toward equilibrium are significantly different from zero with the expected signs; α_H is negative, and α_L is positive, implying that they move in opposite directions to restore equilibrium after a shock to the system occurs. We can note that the absolute values of the estimates of α_H are much smaller than α_L, which suggests that the correction in the equation for daily lows overshoots the long-run equilibrium. Caporin et al. (2013) obtained similar results when analyzing DJIA stocks. When interpreting the short-run dynamics parameters (γ_{11}, …, γ_{22}), we may notice that the coefficients of the lagged daily highs are mostly negative, whereas the coefficients of the lagged daily lows are mostly positive. Cheung (2007) states that negative coefficients imply a regressive behavior, whereas positive coefficients are an indication of spill-over effects. He argues that higher daily highs tend to fall to a lower level, lower daily highs tend to drift up to a higher level, and higher daily lows lead to higher daily highs.

	$\hat{d}$	$\hat{b}$	$\hat{\beta}$	α_H	α_L	γ_{11}	γ_{12}	γ_{21}	γ_{22}
S&P 500									
u 2003-2012	1.035	0.339	(1,-1)	-0.258	3.483	-0.692	0.846	-1.362	2.139
d 2003-2007	1.000	0.426	(1,-1)	-1.114	1.834	0.145	0.129	-0.396	0.997
u 2008-2012	1.024	0.366	(1,-1)	-0.154	2.788	-0.646	0.819	-0.827	1.583
FTSE 100									
2003-2012	1.013	0.390	(1,-1)	-0.623	2.135	-0.303	0.487	-0.595	1.286
d 2003-2007	1.000	0.419	(1,-1)	-1.055	1.890	0.078	0.140	-0.743	1.295
2008-2012	1.018	0.468	(1,-1)	-0.593	1.215	-0.224	0.369	0.021	0.515
DAX									
2003-2012	1.018	0.409	(1,-1)	-0.398	1.876	-0.392	0.543	-0.495	1.039
2003-2007	1.054	0.446	(1,-1)	-0.685	1.670	-0.103	0.253	-0.252	0.591
2008-2012	1.013	0.373	(1,-1)	-0.098	2.343	-0.647	0.795	-0.790	1.385
NIKKEI 225									
2003-2012	1.028	0.489	(1,-1)	-0.210	1.511	-0.342	0.539	-0.202	0.645
2003-2007	1.046	0.507	(1,-1)	-0.590	1.631	-0.138	0.261	-0.468	0.828
2008-2012	1.001	0.516	(1,-1)	-0.079	1.123	-0.295	0.542	0.188	0.297
PX									
2003-2012	1.039	0.455	(1,-1)	-0.753	1.510	-0.201	0.306	-0.642	0.993

	2003-2007	1.053	0.520	(1,-1)	-0.624	1.049	0.041	0.160	-0.121	0.493
d	2008-2012	1.000	0.427	(1,-1)	-0.863	1.493	-0.252	0.399	-0.727	1.122

Table 7: *FCVAR estimation results (with restrictions). Note: "u" denotes that the model is unstable (some of the roots of the characteristic polynomial lie outside the unit root circle), and "d" denotes that the restriction* $d = 1$ *had to be imposed to achieve model convergence.*

In the model without any restrictions imposed, in one case, namely the second period of the NIKKEI 225, we were unable to estimate the FCVAR model without restrictions because the model did not converge for any of the specified number of initial values. Additionally, we must note that the model was not stable in three cases (the S&P 500 in the full period and in the second sub-period and the FTSE 100 in the first sub-period). We state that a model is stable when the roots of the characteristic polynomial are smaller than unity. In these three situations, the roots exceeding unity are 1.365, 1.374, and 1.101, respectively, and thus, the model should be interpreted with caution in these cases. In the Tables, we mark this situation with the letter *"u"* set before the period specification of the affected periods. We also test the residuals for the remaining autocorrelation and heteroskedasticity. Based on the Ljung-Box Q-test, we reject in most cases the null of no autocorrelation; however, the value of the statistic is rather marginal. Based on the visualization of the autocorrelation functions, the dependency is weak, and it disappears after the second lag. Additionally, based on the visualization of the autocorrelation function of squared residuals, we can detect some heteroskedasticity; however, it is again very weak. Neither of these findings impacts the quality of our estimates.[8]

Since the estimates of the cointegrating vector from the unrestricted model are very close to the desired vector of $(1,-1)$, we impose the restriction on the vector to be exactly $(1,-1)$ to be able to interpret the cointegrating relation as the range. Table 7 contains the results of this estimation. In three specifications moreover, we had to impose the restriction $d = 1$ to achieve convergence of the model (we mark this restriction with the letter *"d"* set before the period specification of the affected periods). Imposition of this restriction solved the instability of the model for the FTSE 100 in the first period. However, this restriction was of no use in the case of the S&P 500. In the case of the S&P 500, the models in the full period and in the second sub-period remain unstable (which we mark with *"u"*). The roots of the characteristic polynomial exceeding unity are 1.205 and 1.237, respectively.

We can see that the order of integration of daily prices is in all cases greater than 1 (which is different from the estimation without the restriction, where the order of integration was mostly below 1). The estimates of the cointegration gap $\hat{b}$ are quite similar to the unrestricted specification, apart from the case of the NIKKEI 225 and the PX Index in the first period. In these two situations, the difference $(d-b)$ changed from 0.384 to 0.539 and from 0.341 to 0.533, respectively. The estimates of the order of integration of the range have thus changed from the stationary region into the non-stationary region. We have already discussed these two indices when analyzing their ACFs, where we noted that the presence of long memory in the first period is arguable; this finding was further supported by the results from the GPH and ELW estimators, where the range was found to be stationary but not integrated of order 0. In all of the other cases, the maximum change in the estimate of the difference $(d-b)$ was 0.05, and the implication for stationarity or non-stationarity remained unchanged. In the cases of the NIKKEI 225 and PX Index in the first period, we can also note the highest differences among the α and γ parameters in the estimation with and without the restriction. In the other specifications, these six parameters vary slightly, but nowhere near as much as in these two cases. The adjustment coefficients α_H are again negative, and the adjustment coefficients α_L remain positive. When interpreting these signs, we can make use of the fact that the

[8] For the sake of brevity, this residual diagnostic is not presented here but is available upon request.

cointegrating vector is now the range. An increase in the daily range is reduced the next day by decreasing the high price and boosting the low price for that day. The short-run dynamics parameters (γ_{11}, …, γ_{22}) are again mostly negative for the lagged daily highs and mostly positive for the lagged daily lows; the interpretation remains the same as in the previous specification. We can conclude that the imposition of the $(1, -1)$ restriction on the cointegrating vector may have impaired the results, even though the original estimates of the cointegrating vector are fairly close to the desired $(1, -1)$. The residual diagnostics are very similar to the previous case.

Finally, having all the estimates, we compare the level of integration of the range. Table 8 summarizes the GPH, ELW and final FCVAR estimates. The estimates of long memory in the range are quite sensitive to both the chosen methodology and the chosen bandwidth parameter. In the case of the S&P 500, even though there were some problems with the stability of the FCVAR model, we can see that the FCVAR estimates of long memory in the range are consistent with both the GPH and ELW estimates (with the smaller bandwidth parameter specification). For the FTSE 100 index in the second period, the FCVAR estimates of long memory appear to be quite underestimated because the values would imply that the second period is less non-stationary than the first one, which contrasts with all of the other results (the ELW and GPH estimates and inspection of ACF). In the case of the DAX, we can note that each estimator chooses different periods to be the most and the least non-stationary, despite the fact that the estimates are of magnitudes that are close to each other.

	GPH		ELW		FCVAR	
	$m = T^{0.5}$	$m = T^{0.6}$	$m = T^{0.5}$	$m = T^{0.6}$	R	NR
S&P 500						
2003-2012	0.765	0.674	0.763	0.622	0.696	0.696
2003-2007	0.520	0.678	0.587	0.574	0.574	0.532
2008-2012	0.753	0.721	0.664	0.687	0.658	0.647
FTSE 100						
2003-2012	0.710	0.712	0.658	0.632	0.623	0.613
2003-2007	0.515	0.596	0.622	0.603	0.581	0.604
2008-2012	0.739	0.701	0.628	0.617	0.550	0.524
DAX						
2003-2012	0.652	0.706	0.588	0.589	0.609	0.599
2003-2007	0.803	0.826	0.614	0.597	0.608	0.566
2008-2012	0.710	0.738	0.700	0.623	0.640	0.631
NIKKEI 225						
2003-2012	0.499	0.659	0.483	0.649	0.539	0.487
2003-2007	0.463	0.401	0.496	0.392	0.539	0.384
2008-2012	0.701	0.784	0.613	0.604	0.485	-
PX						
2003-2012	0.577	0.617	0.516	0.520	0.584	0.532
2003-2007	0.211	0.316	0.281	0.446	0.533	0.341
2008-2012	0.760	0.581	0.572	0.517	0.573	0.528

Table 8: *Comparison of integration orders of range. Note: "R" denotes a model with restrictions on the cointegrating vector, and "NR" denotes a model without restrictions.*

Contrary to the ELW and GPH estimates, the FCVAR results for the NIKKEI 225 with restrictions do not confirm the stationarity of the range in the first period, and the model overestimates the order of integration. Moreover, in the second period, the dependence is significantly underestimated. We can observe that the FCVAR model also fails to confirm the stationarity of the range of the PX Index in the first period and overestimates the dependence as well. The FCVAR results in the two remaining periods are quite similar to both the ELW and GPH estimates (with the larger bandwidth parameter).

To summarize the results, the FCVAR model with restrictions fails to detect the lower orders of integration of the range and suggests that the range is in the non-stationary region when it should be stationary according to the results of other applied tests. However, we should note that the FCVAR model does detect the stationarity of the "range" when the restriction on the cointegrating vector is not imposed. However, without the restriction, interpreting the error correction term in the model as the range is incorrect, even though the cointegrating vector is fairly close to the value required for the interpretation to be valid even in the unrestricted model.

The most unanimous conclusion is that, except for the ranges of the PX Index and the NIKKEI 225 in the first period, which are in the stationary region, the remaining ranges are non-stationary and display long memory. The best results can be observed in the case of the S&P 500, where all four different methods for examining long memory yield results closest to each other, despite the instability of the FCVAR model for this index.

5 Conclusion

This work provides empirical support for the fractional cointegration of daily high and low stock prices in several markets. The main motivation for examining these maximum and minimum daily prices is that they provide valuable information about range-based volatility. The range, defined as the difference between daily high and low prices, is considered a highly efficient and robust estimator of volatility. An empirical model based on the fractionally cointegrated VAR framework is able to capture both the cointegration between daily high and low prices and the long memory of their linear combination, i.e., the range. In this concept, the range is the error correction term in the FCVAR model and is allowed to fall into a non-stationary region.

The analysis is performed on four major global indices, namely, the U.S. S&P 500, German DAX, Japanese NIKKEI 225 and U.K. FTSE 100, and the results are compared with the Czech PX Index. We consider three periods, the base period being 2003-2012, and its division into two sub-periods, with the year 2007 as the dividing point. The first sub-period captures the relatively calm behavior before the crisis, whereas the second period covers the outbreak of the crisis and the post-crisis turbulence. We find significant evidence that the range-based volatility estimated as an error correction from the FCVAR of daily high and low prices displays long memory. Moreover, the range is in the non-stationary region in most of the cases, with the exception of ranges of the PX Index and NIKKEI 225 in the first period. In general, the estimates of the long memory parameters are mostly lower in the first period before the crisis. We also demonstrate that the results in the FCVAR framework with restrictions imposed are slightly inferior to the original unrestricted FCVAR model, primarily in the situation when the range should be in the stationary region based on other applied tests. Furthermore, the results for the PX Index are very similar to the results for the NIKKEI 225 because their ranges display the lowest estimates of integration orders. Integration orders of the ranges of the S&P 500, FTSE 100 and DAX indices are, however, relatively higher.

These results can be useful for the predictability of asset prices. The fact that we find a long memory in the range allows the predictability of the variance to be embedded in a model for the mean dynamics of high and low prices and to obtain better forecasts of future extreme prices based on past values. Our results thus provide compelling evidence that daily high and low prices are predictable and can be modeled. This evidence can be materialized in future research in several areas. More precise estimates of daily ranges can be used to enhance trading strategies because many trading strategies are based on daily ranges. Further, it would be interesting to investigate not only how the results can improve risk analysis and management but also other broad areas employing precise volatility estimates as derivative pricing.

References

Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002): Range–Based Estimation of Stochastic Volatility Models, The Journal of Finance, 57(3), 1047–1091.

Andersen, T. G., and T. Bollerslev (1997): Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long–Run in High Frequency Returns, The Journal of Finance, 52(3), 975–1005.

Andersen, T., and T. Bollerslev (1998): Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, International Economic Review, 39(4), 885–905.

Baillie, R. T. (1996): Long Memory Processes and Fractional Integration in Econometrics, Journal of Econometrics, 73(1), 5–59.

Breidt, F. J., N. Crato, and P. de Lima (1998): The detection and estimation of long memory in stochastic volatility, Journal of Econometrics, 83 (1–2), 325–348.

Brockwell, P.J., and R.A. Davis (1991): Time Series: Theory and Methods, 2nd edition, Springer–Verlag, New York.

Caporin, M., Ranaldo, A., and P. Santucci de Magistris (2013): On the predictability of stock prices: A case for high and low prices, Journal of Banking & Finance 37(12), 5132–5146.

Cheung, Y. (2007): An empirical model of daily highs and lows, International Journal of Finance and Economics, 12(1), 1–20.

Cipra, T. (2008): Finanční ekonometrie, (Financial Econometrics), first edition, Ekopress 2008.

Corwin, S. A., and P. H. Schultz (2012): A Simple Way to Estimate Bid–Ask Spreads from Daily High and Low Prices, The Journal of Finance, 67(2), 719–760.

Degiannakis, S. and Livada, A. (2013): Realized volatility or price range: Evidence from a discrete simulation of the continuous time diffusion process, Economic Modelling, 30, 212-216.

Dickey, D. A., and W. A. Fuller (1981): Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica: Journal of the Econometric Society, 1057-1072.

Ding, Z., Granger, C.W.J., and R.T. Engle (1993): A long memory property of stock market returns and a new model, Journal Empirical Finance 1(1), 83–106.

Engle, R., and C.W.J. Granger (1987): Cointegration and error correction: representation estimation, and testing, Econometrica, 55(2), 251–276.

Fiess, N.M. and MacDonald, R. (2002): Towards the fundamentals of technical analysis: analysing the information content of High, Low and Close prices, Economic Modelling, 19(3), 353-374.

Garman, M. B., and M. J. Klass (1980): On the Estimation of Security Price Volatilities from Historical Data, The Journal of Business, 53(1), 67–78.

Garvey, J. F., and L.A. Gallagher (2012): The Realised–Implied Volatility Relationship: Recent Empirical Evidence from FTSE–100 Stocks. J. Forecast., 31(7), 639–660.

Geweke, J., and S. Porter–Hudak (1983): The estimation and application of long memory time series models, Journal of Time Series Analysis 4(4), 221–238.

Granger, C.W.J. (1980): Long memory relationships and the aggregation of dynamic models, Journal of Econometrics, 14(2), 227–238.

Granger, C.W. J. (1986): Developments in the study of cointegrated economic variables, Oxford Bulletin of Economics and Statistics, 48(3), 213-228.

Granger, C.W.J., and N. Hyung (2004): Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns, Journal of Empirical Finance, 11(3), 399–421.

Granger, C. W. J., and R. Joyeux, (1980): An introduction to long memory time series models and fractional differencing, Journal of Time Series Analysis, 1(1), 15–29.

Greene, W. H. (2008): Econometric Analysis, Sixth Edition, Prentice Hall: New Jersey

Johansen, S. (1991): Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Econometrica 59(6), 1551–1581.

Johansen, S. (1996): Likelihood–Based Inference in Cointegrated Vector Autoregressive Models, 2nd edition, Oxford University Press, Oxford.

Johansen, S. (2008): A representation theory for a class of vector autoregressive models for fractional processes, Econometric Theory, 24(3), 651–676.

Johansen, S., and M. O. Nielsen (2010): Likelihood inference for a nonstationary fractional autoregressive model, Journal of Econometrics 158(1), 51–66.

Johansen, S., and M. O. Nielsen (2012) : Likelihood inference for a fractionally cointegrated vector autoregressive model, Econometrica, 80(6), 2667–2732.

Kanzler, L. (1998): GPH: MATLAB module to calculate Geweke–Porter–Hudak long memory statistic, Statistical Software Components.

Kellard, N., Dunis, C., and N. Sarantis (2010): Foreign exchange, fractional cointegration and the implied–realized volatility relation, Journal of Banking & Finance, 34(4), 882–891.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Y. Shin (1992): Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root, Journal of Econometrics 54(1), 159–178.

Lobato, I.N., and N.E. Savin (1997): Real and spurious long–memory properties of stock–market data, Journal of Business and Economic Statistics, 16(3), 261–268.

MacKinnon, J. G., and M. Ø. Nielsen (2012): Numerical distribution functions of fractional unit root and cointegration tests, Journal of Applied Econometrics.

NBER: US Business Cycle Expansions and Contractions, available at: http://www.nber.org/cycles/, last accessed on 15. 5. 2012.

Nielsen, M., and L. Morin (2012): FCVARmodel.m: A MATLAB software package for estimation and testing in the fractionally cointegrated VAR model, QED working paper 1273, Queen's University.

Nielsen, M. Ø., and K. Shimotsu (2007): Determining the cointegration rank in nonstationary fractional system by the exact local Whittle approach, Journal of Econometrics, 141(2), 574–596.

Parkinson, M. (1980): The Extreme Value Method for Estimating the Variance of the Rate of Return, The Journal of Business, 53(1), 61–65.

Robinson, P. M., and Y. Yajima (2002): Determination of cointegrating rank in fractional systems, Journal of Econometrics, 106(2), 217–241.

Shimotsu, K., and P. Phillips (2005): Exact local Whittle estimation of fractional integration, Annals of Statistics, 33(4), 1890–1933.

Yalama, A., and Celik, S.(2013): Real or spurious long memory characteristics of volatility: Empirical evidence from an emerging market, Economic Modelling, 30, 67-72.

Appendix: Figures

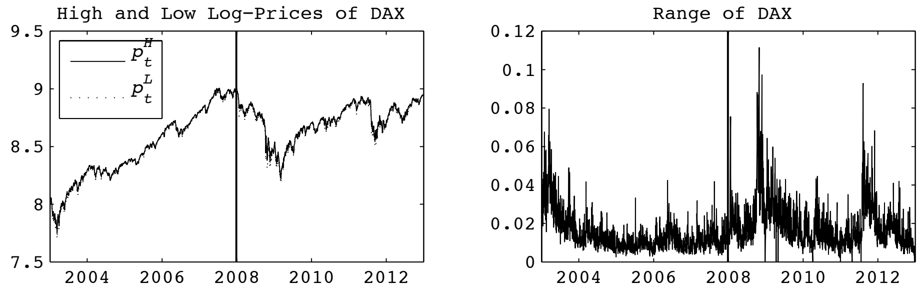

Figure 3: High and low prices of the DAX index (left) and range of the DAX index (right)

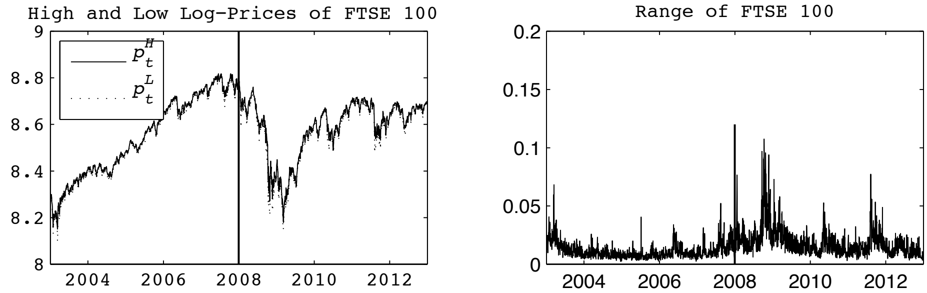

Figure 4: High and low prices of the FTSE 100 index (left) and range of the FTSE 100 index (right)

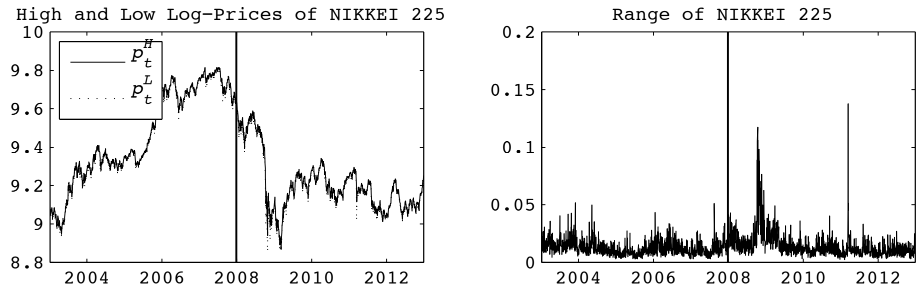

Figure 5: High and low prices of the NIKKEI 225 index (left) and range of the NIKKEI 225 index (right)

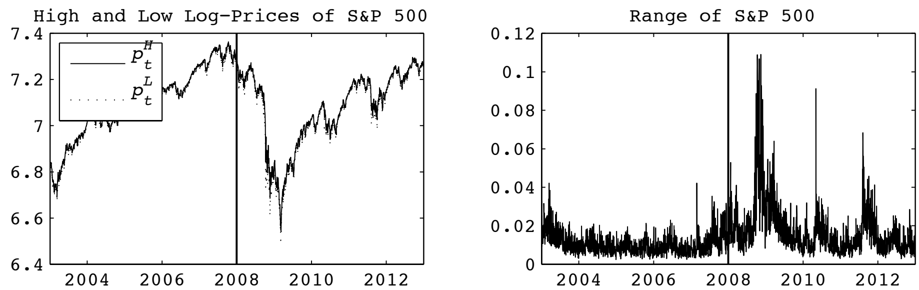

Figure 6: High and low prices of the S&P 500 index (left) and range of the S&P 500 index (right)

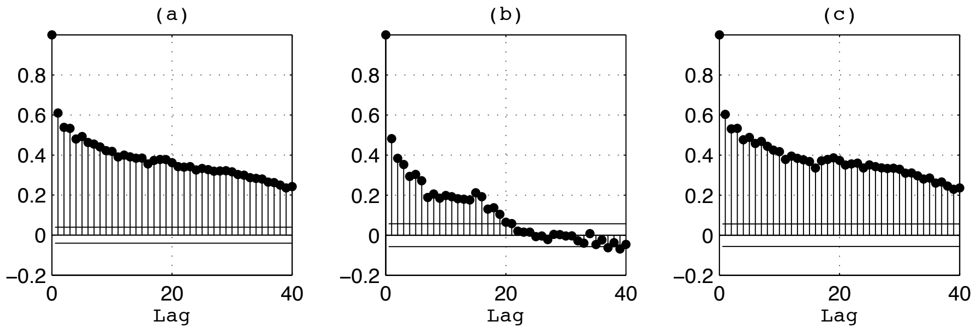

Figure 7: ACF of the PX range in (a) 2003 – 2012, (b) 2003 – 2007 and (c) 2008 – 2012

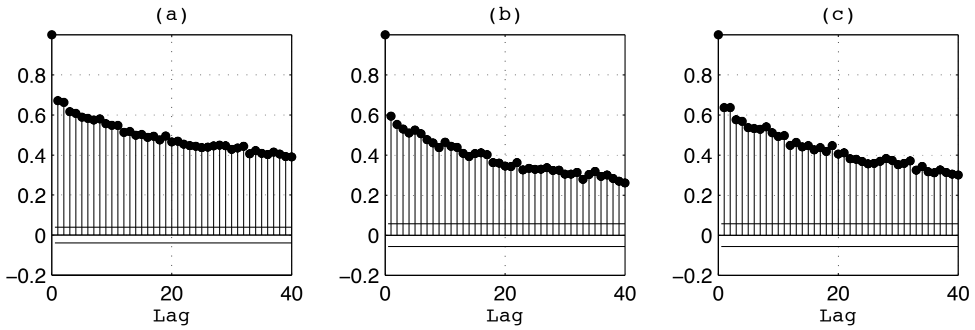

Figure 8: ACF of the FTSE 100 range in (a) 2003 – 2012, (b) 2003 – 2007 and (c) 2008 – 2012

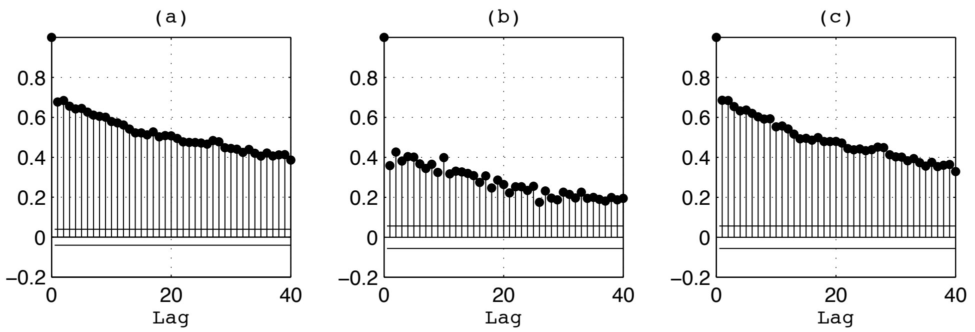

Figure 9: ACF of the S&P 500 range in (a) 2003 – 2012, (b) 2003 – 2007 and (c) 2008 – 2012

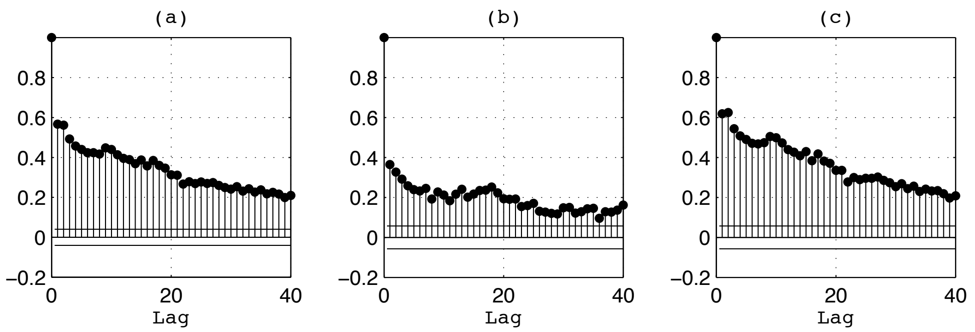

Figure 10: ACF of the NIKKEI 225 range in (a) 2003 – 2012, (b) 2003 – 2007 and (c) 2008 – 2012

Appendix: Tables

		ADF_H		ADF_L		ADF_R
		Level	First-differences	Level	First-differences	Level
S&P 500						
c	2003-2012	0.2947	0.001	0.2697	0.001	0.001
c, t	2003-2007	0.6369	0.001	0.2830	0.001	0.001
c	2008-2012	0.5718	0.001	0.5333	0.001	0.001
FTSE 100						
c	2003-2012	0.2683	0.001	0.1467	0.001	0.001
c, t	2003-2007	**0.0225**	-	0.0180	-	0.001
c	2008-2012	0.2863	0.001	0.1917	0.001	0.001
DAX						
c	2003-2012	0.3666	0.001	0.3380	0.001	0.001
c, t	2003-2007	**0.0419**	-	0.0329	-	0.001
c	2008-2012	0.3246	0.001	0.1799	0.001	0.001
NIKKEI 225						
c	2003-2012	0.4489	0.001	0.3779	0.001	0.001
c	2003-2007	0.4578	0.001	0.4349	0.001	0.001
c	2008-2012	0.1224	0.001	0.0979	0.001	0.001
PX						
c	2003-2012	0.1107	0.001	0.1162	0.001	0.001
c	2003-2007	0.2578	0.001	0.3114	0.001	0.001
c	2008-2012	0.2488	0.001	0.1570	0.001	0.001

Table 9: *P-values for an ADF test for variables based on levels and first-differences. Note: the letter "c" denotes the inclusion of a constant only, and the letter "t" denotes the additional inclusion of a trend for daily high and low prices in levels only. The reported p-value of 0.001 is the smallest reported p-value.*

		$KPSS_H$		$KPSS_L$		$KPSS_R$	
		Short lag	Long lag	Short lag	Long lag	Short lag	Long lag
S&P 500							
c	2003-2012	0.01	0.01	0.01	0.01	0.01	0.0131
c, t	2003-2007	0.01	0.01	0.01	0.01	0.01	0.01
c	2008-2012	0.01	0.01	0.01	0.01	0.01	0.01
FTSE 100							
c	2003-2012	0.01	0.01	0.01	0.01	0.01	0.0143
c, t	2003-2007	0.01	0.01	0.01	0.01	0.01	0.001
c	2008-2012	0.01	0.01	0.01	0.01	0.01	0.001
DAX							
c	2003-2012	0.01	0.01	0.01	0.01	0.01	**0.0784**
c, t	2003-2007	0.01	0.01	0.01	0.0188	0.01	0.001
c	2008-2012	0.01	0.01	0.01	0.01	0.01	0.001
NIKKEI 225							
c	2003-2012	0.01	0.01	0.01	0.01	0.01	**0.0943**
c	2003-2007	0.01	0.01	0.01	0.01	0.01	0.01
c	2008-2012	0.01	0.01	0.01	0.01	0.01	0.01
PX							
c	2003-2012	0.01	0.01	0.01	0.01	0.01	0.01
c	2003-2007	0.01	0.01	0.01	0.01	0.01	**0.0772**
c	2008-2012	0.01	0.01	0.01	0.01	0.01	0.01

Table 10: *P-values for a KPSS test for variables based on levels with two lag specifications Note: the letter "c" denotes the inclusion of a constant only, and the letter "t" denotes the additional inclusion of a trend for daily high and low prices in levels only. The p-value of 0.01 is the minimum reported p-value.*