

Volatility of stock market indices - an analysis based on SEMIFAR models

Jan Beran^{†,1} and Dirk Ocker[‡]

[†]Department of Mathematics and Computer Science
and

[‡]Department of Economics and Statistics
Universitätsstr. 10, Postfach 5560
78457 Konstanz, Germany
University of Konstanz

January 25, 1999

¹corresponding author. E-mail: jberan@iris.rz.uni-konstanz.de

Abstract

By applying SEMIFAR models (Beran, 1999), we examine 'long memory' in the volatility of worldwide stock market indices. Our analysis yields strong evidence of 'long memory' in stock market volatility, either in terms of stochastic long-range dependence or in form of deterministic trends. In some cases, both components are detected in the data. Thus, at least partially, there appears to be even stronger and more systematic 'long memory', than suggested by a stationary model with long-range dependence.

Key words: SEMIFAR model, ARCH models, trend, long-range dependence, short-range dependence, volatility, semiparametric model, kernel estimation, bandwidth selection, maximum likelihood estimation, power transformation.

1 Introduction

Modeling (conditional) variances has been one of the most important topics in the stochastic analysis of financial time series, as can be seen by the extensive amount of research on autoregressive conditional heteroskedastic (ARCH) models (Engle, 1982) and their extensions, in particular generalized ARCH (GARCH), exponential GARCH (EGARCH) and integrated GARCH (IGARCH). These models are readily interpreted as ARMA- and ARIMA-type models of the (conditional) variance (Bollerslev & Mikkelsen, 1996). That is, they possess exponentially decaying summable correlations. Also, long-range dependence in the (conditional) variance of financial time series, in particular stock market indices, has recently attracted considerable attention in the literature (see e.g. Bollerslev & Mikkelsen, 1996; Crato & de Lima, 1994; Ding & Granger, 1996; Ding, Granger & Engle, 1993). The new approaches, such as fractionally integrated GARCH (FIGARCH) and long memory ARCH (LM-ARCH), allow to model a certain kind of volatility persistence, typically detected via a slow hyperbolic decay of the correlations of an appropriate volatility measure. In particular, Ding, Granger & En-

gle (1993) found substantially high correlation between absolute returns and power transformed absolute returns of some stock market indices for long lags. Independently Baillie, Bollerslev & Mikkelsen (1996) came to similar results, namely long memory in volatility series. Both studies appear to argue against short-range dependent ARCH type specifications of the (conditional) variance based upon squared return series.

In this paper, the potential usefulness of SEMIFAR models (Beran, 1999; Beran & Ocker, 1999; Beran, Feng & Ocker, 1998) is explained and their application to volatility series of worldwide nominal stock market indices is discussed. These models include a nonparametric trend function as well as a fractional differencing parameter. This allows for data-driven distinction of long-range dependence, difference-stationarity and deterministic trends.

The paper is organized as follows. In section 2, we give a brief description of Beran's (1999) SEMIFAR model. Most of this section is based on results in Beran (1999); other preprints are also included. A data-driven algorithm is given in section 3. The application of SEMIFAR models to volatility series of nineteen nominal stock market indices is discussed in section 4. Some final remarks are given in section 5. Tables and figures are provided in the appendix.

2 The SEMIFAR model

A *SEMIFAR* model is a Gaussian process Y_i with an existing smallest integer $m \in \{0, 1\}$ such that

$$\phi(B)(1 - B)^\delta \{(1 - B)^m Y_i - g(t_i)\} = \epsilon_i, \quad (1)$$

where $t_i = (i/n)$, $\delta \in (-0.5, 0.5)$, g is a smooth function on $[0, 1]$, B is the backshift operator, $\phi(x) = 1 - \sum_{j=1}^p \phi x^j$ is a polynomial with roots outside the unit circle and ϵ_i ($i = \dots, -1, 0, 1, 2, \dots$) are iid zero mean normal with $\text{var}(\epsilon_i) = \sigma_\epsilon^2$. Here, the fractional difference $(1 - B)^\delta$ introduced by Granger

and Joyeux (1980) and Hosking (1981) is defined by

$$(1 - B)^\delta = \sum_{k=0}^{\infty} b_k(\delta) B^k \quad (2)$$

with

$$b_k(\delta) = (-1)^k \frac{\Gamma(\delta + 1)}{\Gamma(k + 1)\Gamma(\delta - k + 1)}. \quad (3)$$

The motivation for this definition can be summarized as follows: We wish to have a model that may be decomposed into an arbitrary deterministic (possibly zero) trend and a random component that may be stationary or difference stationary. Moreover, short-range and long-range dependence as well as antipersistence should be included. Here, a stationary process Y_i with autocovariances $\gamma(k) = \text{cov}(Y_i, Y_{i+k})$ is said to have long-range dependence, if the the spectral density $f(\lambda) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} \exp(ik\lambda)\gamma(k)$ has a pole at the origin

$$f(\lambda) \sim c_f |\lambda|^{-\alpha} \quad (|\lambda| \rightarrow 0) \quad (4)$$

for a constant $c_f > 0$ and $\alpha \in (0, 1)$, where ' \sim ' means that the ratio of the left and right hand side converges to one (Mandelbrot, 1983; Hampel, 1987; Künsch, 1987; Beran, 1994 and references therein). In particular, this implies that, as $k \rightarrow \infty$, the autocovariances $\gamma(k)$ are proportional to $k^{\alpha-1}$ and hence they are not summable. On the other hand, a stationary process is called antipersistent, if (4) holds with $\alpha \in (-1, 0)$. This implies that the sum of all autocovariances is zero. Note that for usual shot-memory processes, such as stationary ARMA processes, (4) holds with $\alpha = 0$, and the autocovariances sum up to a nonzero finite value.

To model long-range dependence and to avoid overdifferencing, which is often encountered in the usual Box-Jenkins setting, Granger & Joyeux (1980) and Hosking (1981) introduced fractional ARIMA processes. There, the differencing parameter d is restricted to the stationary range $(-.5, .5)$. In a direct extension, Beran (1995) defines an arbitrary differencing parameter $d > -.5$ such that $(1 - B)^m Y_t$ is a stationary fractional ARIMA(p, δ, q) process, $m = [d + .5]$ is the integer part of $d + .5$ and $\delta = d - m$. This corresponds to equation (1) with a constant function $g \equiv \mu$. Since the integer differencing parameter m assumes integer values only and the fractional

differencing parameter δ is in $(-.5, .5)$, both differencing parameters can be recovered uniquely from the 'overall differencing parameter' $d = m + \delta$. If $d > .5$, then we have a nonstationary fractional ARIMA process. It should be noted, in particular, that this parametrization allows for maximum likelihood estimation of d . Thus not only δ , but also m can be estimated from the data and confidence intervals can be given for both differencing parameters (see Beran, 1995). Excluded are, however, deterministic trends with stationary errors ($m = 0$) and other than polynomial trends. SEMIFAR models extend the definition of fractional ARIMA models with arbitrary $d = m + \delta$ by including an arbitrary deterministic trend function g satisfying certain smoothness assumptions.

More specifically, for SEMIFAR models, $Z_i = \{(1 - B)^m Y_i - g(t_i)\}$ is a stationary (possibly) fractional autoregressive process. Thus, the spectral density of Z_i is proportional to $|\lambda|^{-2\delta}$ at the origin so that the process Z_i has long-memory if $\delta > 0$, antipersistence if $\delta < 0$ and short memory if $\delta = 0$. 1 generalizes stationary fractional AR-processes to the nonstationary case, including difference stationarity and deterministic trend. Four special cases of model (1) are:

- (a) $Y_i =$ no deterministic trend + stationary process with short- or long-range dependence;
- (b) $Y_i =$ deterministic trend + stationary process with short- or long-range dependence;
- (c) $Y_i =$ no deterministic trend + difference-stationary process, whose first difference has short- or long-range dependence;
- (d) $Y_i =$ deterministic trend + difference-stationary process, whose first difference has short- or long-range dependence.

Observe that alternative (c) includes stochastic trends which are typically generated by purely stochastic nonstationary processes ($m = 1$) such as random walks or integrated ARIMA models. In addition to nonstationary models, stationary long memory processes often exhibit local spurious trends

which may be hard to distinguish from deterministic and/or purely stochastic trends in nonstationary time series. Here, alternative (a) allows the possibility of local spurious trends ($m = 0, \delta > 0$). Also, (c) allows a combination of stochastic and local spurious trends ($m = 1$ and $\delta > 0$), whereas (b) is a mixture of deterministic and local spurious trends. Alternative (d) includes a mixture of all three kinds of trends.

In practical applications, it is often very difficult to find the 'right' model and, in particular, to decide whether a series is stationary, has a deterministic or stochastic trend, or whether there may be long-range correlations. (in fact, often, a combination of these may be present.) A possible approach to resolving the problem is given by the SEMIFAR model. The model provides a unified data-driven semiparametric approach that allows for simultaneous modeling of and distinction between deterministic trends, stochastic trends and stationary short- and long-memory components. Within the given framework (1), the approach helps the data analyst to decide which components are present in the observed data.

Briefly speaking, a SEMIFAR model is a fractional stationary or non-stationary autoregressive model with a nonparametric trend. This extends Box-Jenkins ARIMA models (Box & Jenkins, 1976), by using a fractional differencing parameter $d > .5$, and by including a nonparametric trend function g . The trend function can be estimated, for example, by kernel smoothing (see Beran, 1999). The parameters may be estimated by an approximate maximum likelihood introduced in Beran (1995) (see also Beran, Bhansali & Ocker, 1999). Note in particular that, with this method the integer differencing parameter is also estimated from the data. A data-driven algorithm for estimating SEMIFAR models, which is a mixture of these two approaches is presented in the following section. Confidence intervals and tests are given in Beran (1999).

3 A data-driven algorithm

The following algorithm, proposed in Beran (1999), is an adaptation of that in Beran (1995) by replacing $\hat{\mu}$ by a kernel estimate of g . The algorithm makes use of the fact that d is the only additional parameter, besides the autoregressive parameters, so that a systematic search with respect to d can be made. The optimal bandwidth is estimated by an iterative plugin method similar to the one in Herrmann, Gasser & Kneip (1992) and Ray & Tsay (1997). The steps of the algorithm are defined as follows:

Step 1: Define L =maximal order of $\phi(B)$ that will be tried, and a sufficiently fine grid $G \in (-0.5, 1.5)$. Then, for each $p \in \{0, 1, \dots, L\}$, carry out steps 2 through 4.

Step 2: For each $d \in G$, set $m = [d+0.5]$, $\delta = d-m$, and $U_i(m) = (1-B)^m Y_i$, and carry out step 3.

Step 3: Carry out the following iteration:

Step 3a: Let $b_o = \Delta_o \min(n^{(2\delta-1)/(5-2\delta)}, 0.5)$ with $0 < \Delta_o < 1$ and set $j = 1$.

Step 3b: Set $b = b_{j-1}$.

Step 3c: Calculate $\hat{g}(t_i; m)$ using the bandwidth b . Set $\hat{X}_i = U_i(m) - \hat{g}(t_i; m)$.

Step 3d: Set $\tilde{e}_i(d) = \sum_{j=0}^{i-1} b_j(\delta) \hat{X}_{i-j}$, where the coefficients b_j are defined by (4).

Step 3e: Estimate the autoregressive parameters ϕ_1, \dots, ϕ_p from $\tilde{e}_i(d)$ and obtain the estimates $\hat{\sigma}_\epsilon^2 = \hat{\sigma}_\epsilon^2(d; j)$ and $\hat{c}_f = \hat{c}_f(j)$. Estimation of the parameters can be done, for instance, by using the *S-Plus* functions *ar.burg* or *arima.mle*. If $p = 0$, set $\hat{\sigma}_\epsilon^2$ equal to $n^{-1} \sum \tilde{e}_i^2(d)$ and \hat{c}_f equal to $\hat{\sigma}_\epsilon^2/(2\pi)$.

Step 3f: Set $b_2 = b^{(5-2\delta)/(9-2\delta)}$ and estimate g'' by

$$\hat{g}''(t) = \frac{1}{nb_2^3} \sum_{j=1}^n \tilde{K}\left(\frac{t_j - t}{b_2}\right) U_j(m)$$

where $\tilde{K} : R \rightarrow R$ is a polynomial symmetric kernel such that $\tilde{K}(x) = 0$ for $|x| > 1$, $\int \tilde{K}(x)dx = 0$ and $\int \tilde{K}(x)x^2dx = 2$. Calculate $I(\hat{g}'')$.

Step 3g: Calculate V and C_{opt} from δ and the estimated parameters obtained in Step 3f. Set

$$b_j = C_{opt} n^{(2\delta-1)/(5-2\delta)}.$$

Step 3h: Increase j by one and repeat steps 3b through 3g 4 times. This yields, for each $d \in G$ separately, the ultimate value of $\hat{\sigma}_\epsilon^2(d)$, as a function of d .

Step 4: Define \hat{d} to be the value of d for which $\hat{\sigma}_\epsilon^2(d)$ is minimal. This, together with the corresponding estimates of the AR parameters, yields an automatic model selection criteria such as the $AIC_\alpha(p)$ (as a function of p) and the corresponding values of $\hat{\theta}$ and \hat{g} for the given order p .

Step 5: Select the order p that minimizes $AIC_\alpha(p)$. This yields the final estimates of θ and g .

The factor $(5 - 2\delta)/(9 - 2\delta)$ in step 3f inflates the bandwidth b to a bandwidth b_2 , which is optimal for estimating g'' in the case of $\delta = \delta^\circ$. The estimated parameters, the selected bandwidth \hat{b} as well as the estimated trend $\hat{g}(t)$, $t \in [0, 1]$, by the above algorithm are all consistent.

4 Volatility of stock market indices

4.1 The Data

The data include nineteen nominal stock market closing indices for the period January 1, 1992 to November 10, 1995. They are, according to definition of the IFC (1997), indices for ten developed markets (DMs: Australia, Belgium, Canada, France, Germany, Hong Kong, Italy, Switzerland, United Kingdom,

and United States), and nine emerging markets (EMs: Brazil, Chile, Greece, Hungary, Malaysia, Mexico, Poland, South Korea, and Thailand). Table 1 presents the names and the exchanges for these indices, together with global ranking by market capitalization in US\$ terms as of end-1995 (Euromoney, 1996).

The indices are expressed in local currencies and, overall, are neither adjusted for dividends nor for inflation. Figure 1 shows daily values of the indices (weekdays only). Besides the large number of infrequent local spikes, which are often related to as heteroskedasticity, most indices exhibit an apparent high magnitude around the middle part of the period under consideration. Note the impact of the Mexican currency and banking crisis, beginning in the last quarter of 1994, and the corresponding (retarded) spillovers to Brazil and Chile. Observe also the low level of the Easteuropean indices at the beginning of the period under consideration. The two stock markets in Hungary and Poland were re-established in 1990 and 1991 respectively, resulting in a low degree of activity in 1992. Finally, observe the smooth sample path of the Brazilian index, which is due to several rebasements during the sample period.

To study volatility, we analyze the power transformed absolute differences $Y_t = |I_t - I_{t-1}|^{.25}$, where I_t denotes the original index. The corresponding series are shown in figure 2 (weekdays only, excluding holidays). When evaluating multiple assets from different countries within an a multivariate framework, the handling of holidays becomes an issue. Sophisticated statistical optimization methods may be required to specify stochastic models. In the current univariate study, we take a simple pragmatic approach. In a first step, missing values in the original index series are replaced by the closest, previous closing value, resulting in zero increments. In a second step, zero values of Y_t were omitted and the series are treated as equidistant.

The reason for taking the fourth root of the increments is that the marginal distribution of the resulting series is very close to normal (see the normal probability plots in figure 3). A similar transformation approach is used, for instance, by Ding, Granger & Engle (1993). As in their study, the correlograms of Y_t in figure 4 do indeed indicate slowly decaying autocor-

relations (with the exception of the three DMs Belgium, France and US). The question arises, whether this behaviour may be explained by long-range dependence in the stochastic component and/or a (nonparametric) deterministic trend. A non-parametric deterministic (and essentially arbitrary) trend function as an additional building block can, apart from d , explain long-term fluctuations. A smooth deterministic function can be interpreted as an even stronger (and more systematic) degree of temporal dependence than stationarity with slowly decaying correlations.

4.2 Empirical Results

Table 2 summarizes the essential features of the fitted SEMIFAR models for the daily volatility series. The corresponding 95%-confidence intervals are given in brackets. The models were selected using the BIC.

The estimated value of d and the confidence intervals suggest that the stochastic part of all series is stationary ($d < .5$). This is not very surprising in view of the general visual impression given in figures 2. For the EMs (except Brazil) and the two small DMs Belgium and Italy, $d = 0$ is not, or almost not (for Belgium), contained in the 95%-confidence interval. Thus, the estimates indicate that there is long-range dependence in the stochastic component of daily volatility series of EMs and small DMs. For these stock markets, as a finding, the degree of persistence becomes stronger the smaller the market (see figure 7). Applying Spearman's rank correlation we found that $\hat{\rho} = .77$ (p-value=.022).

Substantial short-term dependence, which is typically assumed in traditional ARCH specifications, was only found in one series (namely for Thailand) in form of small AR(1) term.

For all DMs (except Belgium), a significant deterministic trend is found. For the EMs, only five out of nine markets have a significant trend. However, for them, stochastic long-range dependence is found (except Brazil). Figure 2 shows the volatilities Y_t with the fitted trends and upper and lower 5% critical limits for testing significance of the trends. The results indicate that there

are relatively long periods where volatility is high/low systematically for the DM series. This is, in particular, apparent for the DMs Australia, Canada, Germany, Hong Kong, Italy, Switzerland and UK, where a significant trend is detected due (at least a posteriori) to the relatively long period of high volatility around the middle part of the considered time period. Observe in particular the similarity between the trends for Hong Kong and Switzerland. These findings are less evident for the US and France (which also corresponds to their correlograms, see figure 4). Some EMs (Brazil, Chile, Malaysia, Mexico and Poland) also exhibit periods with high/low volatility in form of a (local) significant deterministic trend. In particular, the stock markets of Brazil, Malaysia and Poland show highly deterministic volatility patterns. Note the extreme behaviour of the Brazilian series which may be due to several rebasements during the period under consideration. For the other EMs, apparent local trends do not persist long enough, and can therefore be explained as spurious.

The satisfactory fits of the models are demonstrated by the normal probability plots and correlograms of the residuals in figures 5 and 6. Slight departures from normality can be observed for Belgium and Brazil. Note, however, that normality of the residuals is not required in order that the theoretical results hold (Beran, 1999).

Overall, the estimates indicate that there is 'long memory' in the volatility of stock market indices, either in form of local deterministic trend (for the DMs and some EMs) or in form of long-range dependence in the stochastic component (for the EMs and small DMs) resulting in local spurious trends. In some cases, both components are present in the data. In contrast, there is almost no evidence for short memory as it is typically assumed in traditional ARCH specifications. Moreover, the significant trends fitted to the volatility series indicate that there may be even stronger and more systematic 'long memory' in volatility than suggested by a stationary model with long-range dependence.

5 Final remarks

In this paper, we illustrated the potential usefulness of SEMIFAR models for volatility analysis by several data examples. We found strong evidence of 'long memory' in power transformed absolute return series. 'Long memory' is understood here as stochastic long-range dependence and/or deterministic trends. A deterministic trend as an additional building block can, apart from d , explain long-term fluctuations.

'Long memory' in the volatility of stock market indices has some important implications:

- If 'long memory' is indeed present in the data, statistical inferences concerning asset pricing models based on traditional testing procedures may no longer be valid (see e.g. Mandelbrot, 1971, Bollerslev & Mikkelsen, 1996).
- In addition, the discovery of 'long memory' suggests possibilities for improved volatility forecasting performance, especially over longer forecasting horizons (see e.g. Ocker, 1999; Beran & Ocker, 1999; Granger & Joyeux, 1980; Geweke & Porter-Hudak, 1983).
- Also, many financial time series are available in temporarily aggregated form. Long-range dependence is, in contrast to traditional short memory, robust with respect to temporal aggregation (see e.g. Ocker, 1999; Beran & Ocker, 1999). Realistic models should therefore include the possibility of 'long memory' (stochastic and deterministic).

Our results indicate that traditional short-memory ARCH type specifications may not be appropriate for modelling volatility of stock market indices. Our findings suggest that there may be even stronger and more systematic temporal dependence in volatility than suggested by a stationary ARCH model with stochastic long-range dependence. A more sophisticated analysis of volatility may be obtained by applying GARCH-type extensions of SEMIFAR models to the original index series I_t . The mathematical theory necessary for such extensions is subject to current research. For fractional

models that do not include deterministic trend functions Ling & Li (1997) extend the maximum likelihood method of Beran (1995) to fractional GARCH models. Also, an extension to moving average terms (which may be called 'SEMIFARIMA models') is obvious.

6 Acknowledgements

This research was supported in part by an NSF grant to MathSoft, Inc. (Seattle), and by the Center of Finance and Econometrics, University of Konstanz, Germany. Finally, we would like to thank Dr. Elke M. Hennig (Citibank, Frankfurt) for the stock market series.

BIBLIOGRAPHY

- Baillie, R.T, Bollerslev, T. and Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 3-30.
- Beran, J. (1994). *Statistics for long-memory processes*. Chapman and Hall, New York.
- Beran, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short- and long-memory ARIMA models. *J. Roy. Statist. Soc. B* 57, No. 4, 695-672.
- Beran, J. (1999). SEMIFAR models – a semiparametric fractional framework for modelling trends, long-range dependence and nonstationarity. Preprint, University of Konstanz.
- Beran, J., Bhansali, R.J. and Ocker, D. (1999). On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes. *Biometrika*. (in press).
- Beran, J., Feng, Y. and Ocker, D. (1998). SEMIFAR models. Preprint.
- Beran, J. and Ocker, D. (1999). SEMIFAR forecasts, with applications to foreign exchange rates. To appear in *Journal of Statistical Planning and Inference*.
- Beran, J., Feng, Y., Franke, G., Hess, D. and Ocker, D. (1999). SEMIFAR models, with applications to commodities, exchange rates and the volatility of stock market indices. Preprint.
- Beran, J. and Ocker, D. (1999). Temporal aggregation of stationary and nonstationary FARIMA($p, d, 0$) models. Preprint.
- Bollerslev, T, and Mikkelsen, H.O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics* 73, 151-184.

- Box, G.E. and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control. Holden Day, San Francisco.
- Crato, N. and de Lima, P.J.F. (1994). Long-range dependence in the conditional variance of stock returns. *Economics Letters* 45, 281-285.
- Ding, Z., Granger, C.W.J., and Engle, R.F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1, 83-106.
- Ding, Z., and Granger, C.W.J. (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of Econometrics* 73, 185-215.
- Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50, 987-1008.
- Euromoney (1996). The LGT guide to world equity markets 1996. Euromoney, London.
- Geweke, J., and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *J. Time Series Anal.* 4, 221-238.
- Granger, C.W.J., Joyeux, R. (1980). An introduction to long-range time series models and fractional differencing. *J. Time Series Anal.* 1, 15-30.
- Hampel, F.R. (1987). Data analysis and self-similar processes. In: Proceedings of the 46th Session of ISI, Tokyo, Book 4, 235-254.
- Herrmann, E., Gasser, T. and Kneip, A. (1992). Choice of bandwidth for kernel regression when residuals are correlated. *Biometrika* 79, 783-795.
- Hosking, J.R.M. (1981). Fractional differencing. *Biometrika* 68, 165-176.
- International Finance Corporation (IFC) (1997). Emerging stock markets factbook 1996. IFC, Washington.
- Künsch, H (1986). Statistical aspects of self-similar processes. Invited paper, Proc. First World Congress of the Bernoulli Society, Tashkent, Vol. 1, 67-74.

- Ling, S.-Q. and Li, W.K. (1997). Fractional ARIMA-GARCH times series models. *JASA* 92, 1184-1194.
- Mandelbrot, B.B. (1971). When a price be arbitrated efficiently? a limit to the validity of the random walk and Martingale models. *Review of Economics and Statistics* 53, 225-236.
- Mandelbrot, B.B. (1983). *The fractal geometry of nature*. New York: Freeman.
- Ocker, D. (1999). *Stationary and nonstationary fractional ARIMA models - model choice, forecasting, aggregation and intervention*. Unpublished PhD thesis. University of Konstanz.
- Ray, B.K. and Tsay, R.S. (1997). Bandwidth selection for kernel regression with long-range dependence. *Biometrika* 84, 791-802.

Appendix

Table 1: Stock indices of developed and emerging markets

| Countries | Exchange | Index | Ranking |
|------------------|--------------|------------------------|---------|
| <i>DM series</i> | | | |
| Australia | Sydney | All Ordinaries | 10 |
| Belgium | Brussels | Stock Index | 20 |
| Canada | Toronto | TSE 300 | 7 |
| France | Paris | CAC 40 | 5 |
| Germany | Frankfurt | DAX | 4 |
| Hong Kong | Hong Kong | Hang Seng | 9 |
| Italy | Milan | DS General | 12 |
| Switzerland | Zurich | Swiss Bank Corporation | 6 |
| United Kingdom | London | FTSE 100 | 3 |
| United States | New York | S&P 500 | 1 |
| <i>EM series</i> | | | |
| Brazil | Sao Paulo | BOVESPA | 17 |
| Chile | Santiago | IGPA | 22 |
| Greece | Athens | General Index | 38 |
| Hungary | Budapest | BUX | 49 |
| Malaysia | Kuala Lumpur | KLSE | 11 |
| Mexico | Mexico City | IPC | 21 |
| Poland | Warsaw | WIG | 44 |
| South Korea | Seoul | KOSPI | 18 |
| Thailand | Bangkok | Book Club | 23 |

Table 2: Estimation results

| Countries | \hat{d} | 95%-c.i. d | $\hat{\phi}_1$ | 95%-c.i. ϕ_1 | significant trend |
|------------------|-----------|----------------|----------------|-------------------|-------------------|
| <i>DM series</i> | | | | | |
| Australia | .025 | [-.024, .074] | - | - | yes |
| Belgium | .048 | [-.002, .097] | - | - | no |
| Canada | .035 | [-.014, .084] | - | - | yes |
| France | -.085 | [-.135, -.036] | - | - | yes |
| Germany | -.020 | [-.069, .029] | - | - | yes |
| Hong Kong | -.004 | [-.054, .046] | - | - | yes |
| Italy | .052 | [.003, .101] | - | - | yes |
| Switzerland | .036 | [-.013, .085] | - | - | yes |
| UK | -.025 | [-.074, .024] | - | - | yes |
| US | -.041 | [-.090, .008] | - | - | yes |
| <i>EM series</i> | | | | | |
| Brazil | -.062 | [-.112, -.012] | - | - | yes |
| Chile | .248 | [.198, .298] | - | - | yes |
| Greece | .221 | [.172, .271] | - | - | no |
| Hungary | .283 | [.233, .333] | - | - | no |
| Malaysia | .108 | [.058, .158] | - | - | yes |
| Mexico | .104 | [.055, .153] | - | - | yes |
| Poland | .233 | [.173, .293] | - | - | yes |
| South Korea | .142 | [.067, .217] | -.085 | [-.180, .011] | no |
| Thailand | .208 | [.136, .280] | -.179 | [-.269, -.089] | no |