Accepted Manuscript

Investigating detrended fluctuation analysis with structural breaks

Rui Menezes, Álvaro Oliveira, Sofia Portela

PII: S0378-4371(18)31511-5
DOI: https://doi.org/10.1016/j.physa.2018.12.006
Reference: PHYSA 20392

To appear in: *Physica A*

Please cite this article as: R. Menezes, Á. Oliveira and S. Portela, Investigating detrended fluctuation analysis with structural breaks, *Physica A* (2018), https://doi.org/10.1016/j.physa.2018.12.006

Investigating Detrended Fluctuation Analysis with Structural Breaks

HIGHLIGHTS

- Quadratic functions with negative concavity track integrated price returns
- Squared residuals yield better fits than average variances about the detrended walk
- Endogenous structural breaks yield better forecasts than equal length alternatives
- Data fitted by power-law distributions after the first few observations
- Crossover value of 0.5 reached for samples of at least four years of daily data

Investigating Detrended Fluctuation Analysis with Structural Breaks

Rui Menezes[†] (corresponding author) (rui.menezes@iscte-iul.pt)

Álvaro Oliveira[‡] (ad.oliveira@sapo.pt)

Sofia Portela[†] (sofia.portela@iscte-iul.pt)

[†] Department of Quantitative Methods for Business and Economics and BRU-IUL, ISCTE Business School, Av das Forças Armadas 1649-026 Lisboa, Portugal

[‡] Department of Accountancy and BRU-IUL, ISCTE Business School, Av das Forças Armadas 1649-026 Lisboa, Portugal

ABSTRACT

Detrended Fluctuation Analysis has been used in several fields of science to study the statistical properties of trend stationary and nonstationary time-series. Its application to financial data has produced important results concerning long-range correlations and long-memory. However, these results may be contaminated if the researcher attributes to nonstationary trends the effect of stationary trends with endogenous structural breaks. Our paper proposes a modified DFA model where boxes to determine local trends are replaced by endogenous structural break windows. We also allow local trends fitted by quadratic functions and use squared residuals in place of patchy standard deviations to study the magnitude of the power-law exponent. The results show that our modified DFA model performs better than the fixed length alternatives originally proposed, and is, therefore, a suitable model to fit with financial data. Consistently with previous findings, our results show positive long-range correlation in all indices with the higher value for emerging markets.

KEYWORDS

Detrended fluctuation analysis; detrended walk; structural break; forecast accuracy; power-law

PACS number(s): 05.40.−a, 89.65.Gh

1. Introduction

Many researchers have attempted to model the statistical properties of stock market prices and returns. This has led to the proposal of a huge number of models and techniques capable of dealing with the peculiarities of financial data. One of these techniques, born in the context of statistical physics, is called Detrended Fluctuation Analysis. This technique has gained some popularity among researchers given its simplicity and the successful results achieved. In spite of this, the use of this technique with financial data still needs some improvements, given the specificities of these data. One such specificity is the occurrence of structural breaks in time series that are relatively long. Structural breaks usually occur at irregular intervals and should not be defined exogenously in relation with the chronological sequence of the series.

In this paper we propose a modified DFA model that accommodates structural breaks which are determined endogenously on the basis of the observed data. After the computation of break dates, we define the patches on the basis of each irregular break interval then obtained. The forecasts are then computed within each patch using a nonlinear deterministic trend to fit the data. Instead of using the residuals' standard deviation, we use squared residuals themselves to evaluate the power law assumption. Finally, in order to evaluate the performance of our model in comparison with alternative models similar to the conventional DFA, we perform both in-sample and out-of-sample forecast accuracy tests. Our model was tested with four stock market aggregated indexes published by Morgan Stanley Capital International (MSCI). The level of aggregation varies since the main purpose of this paper is to analyze and select the best fit in different contexts and levels of data aggregation. As we shall see, our modified DFA model yields in general better results than the fixed-length models tested as alternatives, confirming that accounting for endogenous structural breaks improves the overall fit of our data. Our results indicate the presence of positive long-range correlation in all indices.

The rest of the paper is organized as follows. In the next section we provide the background to the study and summarize some features of conventional detrended fluctuation analysis. Next, we present in three steps our model of integrated price returns with structural breaks and the data used at each step for the four indexes under analysis. The following section discusses the main empirical results and section 5 compares the forecast accuracy of the models, both in-sample and out-of-sample, in order to ascertain whether our model performs better than the conventional model or not. In the final section we draw some conclusions of our study.

2. Background

Since the influential paper of Peng et al. [1] which introduced Detrended Fluctuation Analysis (DFA) to study the properties of DNA nucleotides, many empirical researchers have used this technique with applications to several fields of science. In finance there have been some applications too, but there is no evidence that the original model has been modified to accommodate financial data oddities. One such oddity concerns the occurrence of structural breaks that may be mistaken by stochastic non-stationarities in the data and lead to spurious results.

Peng et al. [1] argue that "DFA permits the detection of long-range correlations embedded in a patchy landscape and avoids the spurious detection of apparent long-range correlations that are an artifact of patchiness". This mention is related to the replacement of the unique long-run trend embedded in the data by local trends within patches, but in fact stochastic trends and structural breaks are overlooked in this process. Others argue that DFA allows the distinction of intrinsic autocorrelation from that imposed by external nonstationary movements. Intrinsic autocorrelation is associated with memory effects in the underlying dynamic system, but the existence of nonstationary movements on their own should not be ignored. Therefore, rather than computing local forecasts within patches simply on the basis of deterministic trends, and ignoring local stochastic trends, we may utilize a general nonstationary framework and test whether the residuals obtained from regressing actual on fitted local values are stationary, i.e. actual and fitted values must be cointegrated both in mean and in variance.

Another advantage of detrended fluctuation analysis is that it can be used to assess multiscale autocorrelation [2]. Multiscale in this context refers to different data frequencies or resolutions, where higher frequencies are associated with finer scales and lower frequencies are associated with coarser scales. Fine scales typically involve noise problems that need to be purged while coarse scales may suffer from time aggregation problems. However, in the context of fractal geometry, detrended fluctuation analysis has been used to assess the statistical self-affinity of a time series at different times. Self-affinity occurs when a set can be decomposed into subsets that can be linearly mapped into the full set, with special cases being called self-similarity. In this way, fractal objects are defined as shapes made of parts similar to the whole in some way, such that different data scales would not lead to different results or structures, i.e. a fractal object is self-similar or self-affine at any scale.

In finance, Ref. [3] studies multifractal detrended fluctuation analysis in the German stock market using tick data for six stocks traded in the *Deutsche Börse* and concludes that the scaling exponent lies around 0.5 or less, indicating that no significant autocorrelations exist. Ref. [4] analyses the multifractality degree of developed and emerging stock market indices. Their results show that the multifractality degree is inversely associated with the stage of market development, with most developed markets exhibiting scaling exponents around 0.5 or less and emerging markets showing scaling exponents typically above 0.5, and imply that the multifractality degree may be used to assess stages of stock market development. Ref. [5] also uses multifractal detrended fluctuation analysis, along with multifractal spectrum analysis, to analyze 10-minute closing prices of the Chinese stock index futures market, and found that this market exhibits long-range correlations and multifractality, consistent with a scaling exponent higher than 0.5. Similar results for the Asian stock markets were obtained by [6] using coupling detrended fluctuation analysis. Finally, Ref. [7] combines detrended fluctuation analysis and mutual information to study frontier markets' efficiency. The results for the scaling exponent show that it exceeds 0.5 in most markets and the higher the value the "less efficient" is the underlying market. Therefore, we may conclude that emerging markets tend to be less efficient than developed markets but, although the empirical evidence is consistent across different studies, none of these actually proposes modifications to the underlying basic model to accommodate singularities of the financial market data.

The use of the DFA technique leads to objects that can be statistically treated as power-laws where the power-law exponent describes the type of long-range correlation and memory effects that characterize the specific data under study [8, 9]. Taking full advantage of this relationship, one can define with precision what kind of time-series we are dealing with and what are the main features of our data. Under cointegration of the detrended walk, DFA can be used either with stationary and nonstationary data without misleading results [10, 11, 12].

Conventional detrended fluctuation analysis (DFA) involves the following steps [1]:

1. Integrate the original series (cumulative sum or profile) after computing the difference to the mean;
2. Divide the integrated series into patches of equal length l;
3. For each patch compute the forecast of the integrated series and the corresponding in-sample errors;
4. For each patch compute the square root of the average sum of squared errors;
5. Test the power-law assumption by running a regression of the log of the above defined standard deviation on the log of n, where n denotes the number of patches.

As we show in the next section, this paper proposes a modified Detrended Fluctuation Analysis model, where the above-mentioned steps are going to be changed in order to accommodate structural breaks and obtain better estimates of the model parameters.

3. Model and Data

3.1. Historical Price Returns

As noted above, the main purpose of this study is to analyze and select the best fit of the distributional signal of prices and returns using four stock market aggregated indexes published by Morgan Stanley Capital International (MSCI): World, Emerging Markets, Europe and Pacific. The first step is to examine the chronological profile of the signals. For returns, this profile can be seen in Figure 1.

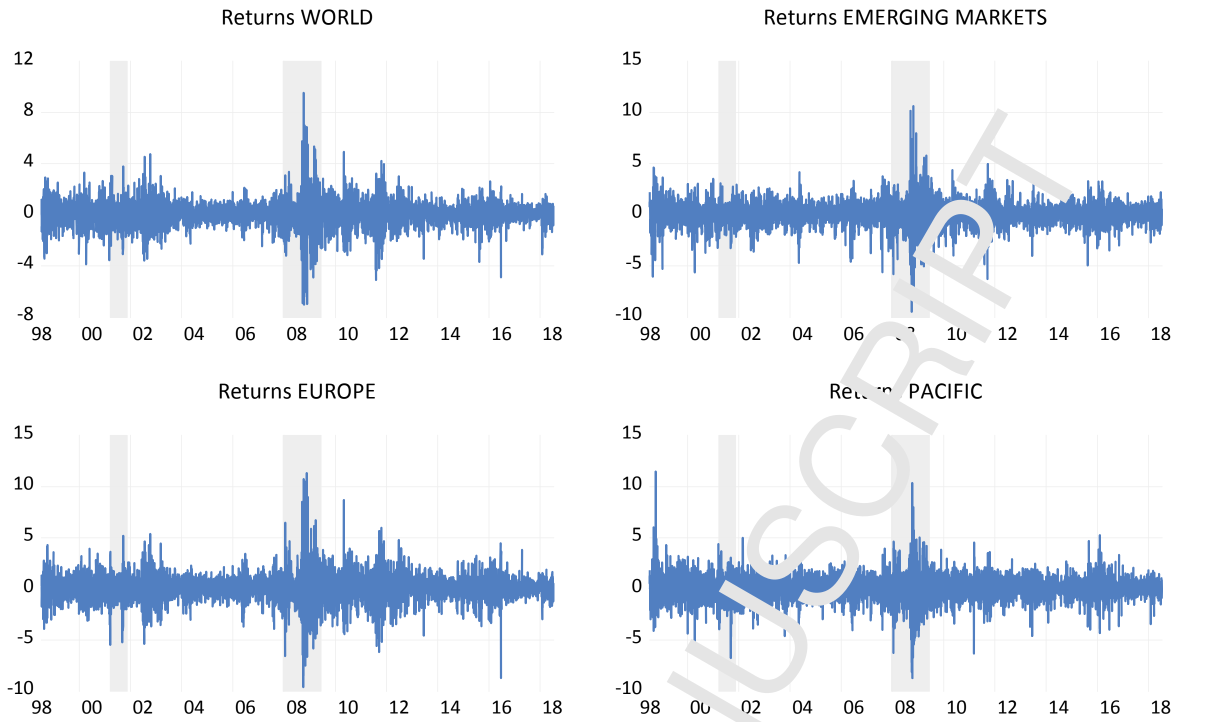


Figure 1: Returns signal. Source: Morgan Stanley Capital International (MSCI) daily price index, 07/13/1998 – 07/13/2018, 5220 observations. Shaded bars represent economic recession periods in the US.

Returns were calculated from the original prices using $x_t = 100 \times (P_t - P_{t-1})/P_{t-1}$, where P_t denotes the published price series and x_t represents the calculated return at time t.[1] The series of returns are substantially linearly correlated. The European index has a correlation of 83% with the World index and a correlation of 66% with the Emerging Markets index.[2] The Emerging Markets index, in turn, has a correlation of 68% with the World index and 59% with the Pacific index. The lowest correlation occurs between the Pacific and the European indexes (35%) and between the Pacific and the World indexes (40%). All the correlations are significant at the 1% level. These results are not really unexpected, given the increasing degree of volatility contagion observed in financial markets over the past decades, especially in the western economies, and the bulky contribution of European stock markets to the global indexes [13].

The series depicted in Figure 1 show that the volatility of returns varies with time and there is a propensity to cluster the periods of high volatility, that is, high volatility periods follow other periods of high volatility and vice-versa (volatility clustering). Another stylized fact of

[1] In other studies, the returns were calculated as the difference of the logarithm of prices at two adjacent points in time, with a similar interpretation.
[2] Throughout this section, and unless otherwise stated, correlation refers to the Pearson correlation coefficient.

many financial time-series is the negative correlation between the volatility of returns and prices, that is, high volatility occurs more frequently when prices drop (leverage effect). In our case, the linear correlation between returns' volatility and prices is negative and statistically significant at the 1% level for the four aggregated indexes: World (−34%), Emerging Markets (−11%), Europe (−30%) and Pacific (−27%).

Asymmetry, and especially leptokurtosis, are also frequently found in the empirical distribution of stock market returns (leptokurtosis leads to fat or heavy tails). Kurtosis lies around 10 in all cases and, except for the European index, all our series exhibit negative skewness. Negative skewness means that the returns distribution has a long left-tail and, although positive returns are more likely to occur than negative returns, the tail is greater on the left-side. A Jarque-Bera [14] test of the null hypothesis of gaussianity is rejected at the 1% level in the distribution of returns of all our four indexes. Thus, alternative distributions that account for thicker tails may describe better the distribution of returns. The main descriptive statistics of the return's series are presented in Table 1.

	World	Emerging Markets	Europe	Pacific
Mean	0.017	0.029	0.013	0.020
Maximum	9.523	10.598	11.291	11.431
Minimum	-7.063	-9.511	-9.677	-8.773
Std. Dev.	0.991	1.194	1.330	1.233
Skewness	-0.234	-0.339	0.004	-0.023
Kurtosis	10.746	10.522	10.221	8.723
Jarque-Bera	13096.04 **	12404.57 **	11339.81 **	7122.33 **

Table 1: Descriptive statistics of the returns signal. Source: Morgan Stanley Capital International (MSCI) daily price index, 07/13/1998 – 07/13/2018, 5220 observations.

The four time-series outlined before follow, in general, the same chronological pattern of volatility, where higher volatility occurs during periods of price decrease (crisis spells). There are two relevant periods of financial crisis during the whole span analyzed, each one associated with a period of economic recession in the US too. The first financial downturn occurred from 2000 to 2003 (Dot-Com Crash). During this period, an economic downturn known as the Early 2000s Recession (Mar.-Nov. 2001) also took place in the US, motivated by the above-mentioned Dot-Com crash in combination with a fall in business outlays and investments, and the September 11th attacks [15].

Later, a second and much more severe financial downturn occurred in 2007-2009, which is known as the Global Financial crisis. During the same time span, an economic downturn known as the Great Recession also happened, motivated by the subprime mortgage crisis and subsequent collapse of the US housing bubble, triggering a failure or collapse of many of the US largest financial institutions (e.g. Lehman Brothers *inter alia*) as well as a crisis in the automobile industry.

The above-mentioned financial and economic crises in the US produced strong spillover effects across the global economy, provoking a cascade of bailouts and bankruptcies, particularly in Europe, leading to the European Sovereign Debt Crisis and other slump economic events.

Overall, the behavior of the aggregated indexes analyzed is rather synchronized, particularly during the episodes of price downturn 2000-2003 and 2007-2009, characterized by steep negative slopes commencing, respectively, in 2000 and 2007, and lasting until the end of the respective crisis, and with a relatively similar magnitude in all the cases.

3.2. Integrated Series and Structural Breaks

The second step entails the computation of the integrated series of price returns of the four aggregated indexes analyzed. Integration of price returns means to compute the cumulative sum or profile of the series after taking the first difference to its mean: $X_t = \sum_{i=1}^{t}(x_i - \langle x \rangle)$, where $\langle x \rangle$ represents the mean value of the returns time-series. Hence, the standardized price series itself is the integrated signal of returns, or just called the integrated signal.

Denoting by d the number of times that a series needs to be differenced in order to remove all the stochastic trends that may exist in the data, then returns are stationary if they are integrated of order d equal to zero. Under these circumstances, prices, that is the cumulative sum of returns, are integrated of first-order, and are said to be nonstationary or unit root processes. The non-stationarity of prices implies that they have stochastic trends which can be captured, for example, using Perron [16] unit root tests with structural breaks. In our case, the Perron tests confirm that all the four price index series are nonstationary and integrated of first-order, which means that they hold one stochastic trend. The series of returns, on the other hand, are all trend stationary. Bai-Perron [17] methodology to determine m endogenous structural breaks was also applied and, within each break window (subsample), conventional ADF unit root tests were performed. The results are presented in Table 2.

In his seminal paper on unit roots with structural breaks, Perron [20] points out that the conventional unit root tests (e.g. ADF) are biased toward a false unit root null when the data are trend stationary with a structural break. A number of modified ADF unit root tests were then proposed by the researchers which allow for levels and trends that differ across single or multiple break dates.

Breaks can occur in different situations, such as: i) gradually (innovational outliers) or immediately (additive outliers), ii) consisting on a level shift, a trend break, or both, iii) the break date is known (exogenous) or unknown and estimated from the data (endogenous), and iv) the data are non-trending or trending. One can thus define an intercept break variable termed $DU_t(T_b) = 1\ (t \geq T_b)$ that takes the value 0 for all dates prior to the break, and 1 thereafter, a trend break variable named $DT_t(T_b) = 1\ (t \geq T_b) \cdot (t - T_b + 1)$ which takes the value 0 for all dates prior to the break and is a break date re-based trend for all subsequent dates, and a one-time break dummy variable given by $D_t(T_b) = 1\ (t = T_b)$ which takes the value 1 only on the break date and 0 otherwise. $1(\cdot)$ denotes the indicator function and T_b the break date.[3]

[3] In this paper we define the break date as the first date of the new regime.

Variable	World		Emerging Markets		Europe		Pacific	
	Breaks: 9/23/2002, 10/07/2005, 10/07/2008, 5/17/2012, 7/16/2015		Breaks: 11/08/2001, 3/24/2005, 10/03/2008, 9/13/2012		Breaks: 11/28/2002, 9/29/2008, 9/16/2013		Breaks: 7/12/2001, 7/21/2004, 10/03/2008, 1/02/2013	
	Coeff.	ADF-Stat.	Coeff.	ADF-Stat.	Coeff.	ADF-Stat.	Coeff.	ADF-Stat.
	7/14/1998 - 9/20/2002 -- 1094 obs		7/14/1998 - 11/07/2001 -- 867 obs		7/14/1998 - 11/27/2002 -- 1142 obs		7/14/1998 - 7/11/2001 -- 782 obs	
constant	-9.943 **		-33.798 **		-12.025 **		-19.391 **	
t	0.100 **	0.851	0.250 **	-0.376	0.066 **	[illegible]110	0.276 **	-0.714
t^2	0.000 **		0.000 **		0.000 **		0.000 **	
	9/23/2002 - 10/06/2005 -- 794 obs		11/08/2001 - 3/23/2005 -- 880 obs				7/12/2001 - 7/20/2004 -- 789 obs	
constant	-229.993 **		109.229		-		248.883 **	
t	0.226 **	-1.965 *	-0.257 **	-1.527	-	-	-0.481 **	-0.874
t^2	0.000 **		0.000 **		-		0.000 **	
	10/07/2005 - 10/06/2008 -- 782 obs		3/24/2005 - 10/02/2008 -- 921 obs		11/28/2002 - 9/26/2008 -- 1522 obs		7/21/2004 - 10/02/2008 -- 1097 obs	
constant	-740.659 **	1.074	-595.569 **		[illegible]		-403.179 **	
t	0.646 **		0.521 **	-0.418	0.179 **	-1.897	0.383 **	-0.851
t^2	0.000 **		0.000 **		0.000 **		0.000 **	
	10/07/2008 - 5/16/2012 -- 942 obs		10/03/2008 - 9/12/2012 -- 1029 obs		9/29/2008 - 9/13/2013 -- 1295 obs		10/03/2008 - 1/01/2013 -- 1108 obs	
constant	-977.244 **		-1976.78 **		-453.584 *		-862.523 **	
t	0.571 **	-0.739	1.241 **	-0.783	0.248 *	-3.307	0.529 **	-2.139 *
t^2	0.000 **		0.000 **		0.000		0.000 **	
	5/17/2012 - 7/15/2015 -- 825 obs		9/13/2012 - 7/13/2018 -- 1522 obs		9/16/2013 - 7/13/2018 -- 1260 obs		1/02/2013 - 7/13/2018 -- 1443 obs	
constant	-1189.00 **		957.77[illegible] **		1055.49 **		468.690 *	
t	0.553 **	-2.115 *	-0.406 **	-1.741	-0.453 **	-1.942 *	-0.208 *	-2.520 *
t^2	0.000 **		0.000 **		0.000 **		0.000 *	
	7/16/2015 - 7/13/2018 -- 782 obs							
constant	825.136		-		-		-	
t	-0.373 *	-3.695 *	-	-	-	-	-	-
t^2	0.000 *		-		-		-	
Perron (1997) unit root test with structural breaks		-3.[illegible]52		-3.336		-3.872		-4.032

Table 2. Structural Breaks and Unit Roots of integrated signals. Sample: MSCI daily price index, 07/13/1998 – 07/13/2018, 5220 observations. Break coefficients and dates computed using Bai-Perron [17] tests of l+1 vs. l globally determined breaks. Perron [16] unit root test with an endogenous structural break in both the intercept and trend. All integrated signals (standardized prices) have a unit root with breaks. Conventional ADF unit root tests [18] without breaks were applied to each patch using MacKinnon [19] one-sided p-values. The ADF null was rejected at standard levels in a number of patches. ** Significant at the 1% level; * significant at the 5% level.

Following Perron [20], the unit root tests with structural breaks are based on the equation $y_t = \mu + \theta DU_t(T_b) + \beta t + \gamma DT_t(T_b) + \delta D_t(T_b) + \alpha y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + \epsilon_t$, where α, β, μ and c_i, are parameters, θ, γ and δ are vectors of parameters with dimension equal to the number of breaks, similarly to the underlying dummies. For the innovational outlier (IO) model, the null and alternative hypotheses can be set up imposing a few restrictions on the parameters. For example, the null under the original Perron's model sets the trend parameter β to zero, the autocorrelation (or unit root) parameter $\alpha = 1$ and the γ-vector (trend break) to zero, while the alternative sets the δ-vector (one-time break) to zero and $\alpha = 0$. The break variables and the innovations may enter the model with the same dynamics given by a lag polynomial $\psi(L)$. The model is set up such as ϵ_t are i.i.d. innovations and the breaks are selected endogenously as suggested by Perron [16] and others.

Bai and Perron [17] advise a number of strategies for dealing with the choice of endogenous structural breaks using general to specific procedures rather than methods based on information criteria, such as AIC, because, they argue, the latter tend to select very parsimonious models leading to tests with sometimes serious size distortions and/or power losses with data in the class of ARMA processes. Actually, recent research conducted by Hall et al. [21] shows that the penalty terms incorporated in the structural breaks' information criteria of Yao [22] and LWZ [23] may underestimate by a factor of 3 the number of m true breaks that have an asymptotic effect on the minimized residual sum of squares and propose a modified penalty term for the information criteria in the context of structural break estimation.

We follow [17] to determine breakpoints and coefficients computed by LS with Breaks, using tests of l+1 vs. l globally determined breaks selected by sequential evaluation, trimming 0.15, HAC standard errors & covariance (Quadratic-Spectral kernel, Andrews bandwidth) and allowing heterogeneous error distributions across breaks. Breakpoints are the start of the next regime. Under the null, the variable has l endogenous structural breaks in both the intercept and trend. A useful strategy is to first look at the max tests [17] to see if at least one break is present. In this case, then the number of breaks can be decided based upon a sequential examination of the sup $F(l+1 \mid l)$ statistics constructed using global minimizers for the break dates, i.e. select m such that the tests are insignificant for $l \geq m$. As shown in

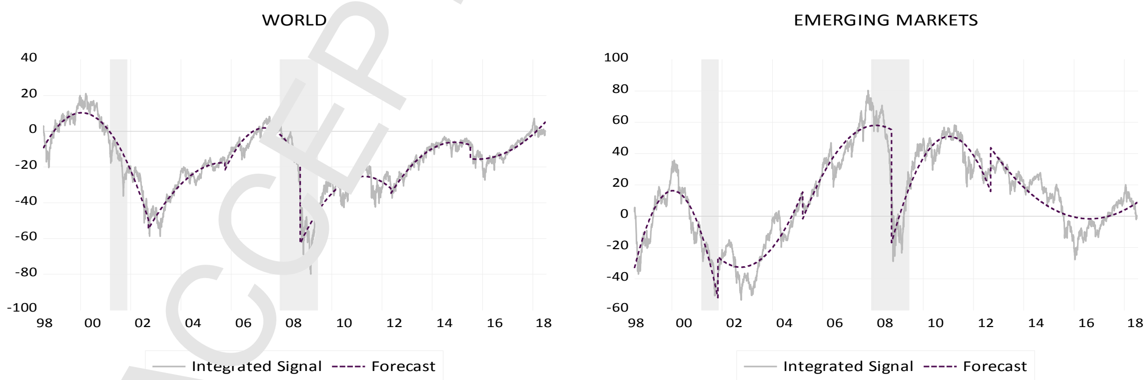
Table 2, all the integrated signals (standardized prices) have a unit root with breaks. Conventional ADF [18] unit root tests without breaks were also applied to each patch and the null was rejected at standard levels in some cases.

The ADF tests are based on the equation $\Delta y_t = \mu + \beta t + (\rho - 1)y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + \epsilon_t$, where μ and β denote, respectively, the intercept and slope of a linear deterministic trend, ρ is the autocorrelation parameter of y_t, and ϵ_t denotes a white noise process. The parameters c_i are included in the model to capture any remaining serial correlation up to order k on the left-hand side variable. If $\rho = 1$, then y_t is a difference stationary process. In addition, if $c_i = \beta = 0$, then y_t is a random walk with drift. A random walk process is a unit root or nonstationary process $I(d)$ with $d = 1$.

4. Results

4.1. Cumulative Sum of Price Returns

The integrated series of price returns for the four aggregated indexes here analyzed are depicted in Figure 2. Note that the periods of price decline correspond to the periods of higher volatility shown in Figure 1, upholding the presence of leverage effects in the stock market as mentioned before, together with the recent well-known economic recessions.



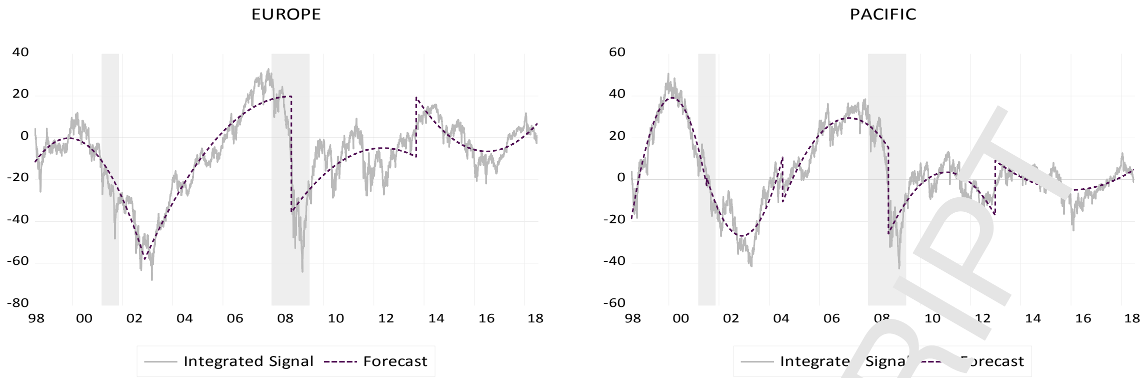


Figure 2: Integrated signal. Source: Morgan Stanley Capital International (MSCI) daily price index, 07/13/1998 – 07/13/2018, 5220 observations. Solid line denotes the actual series, dashed line denotes the fitted values. Shaded bars represent economic recession periods in the US according to the NBER definition.

For each aggregated index, Figure 2 depicts the actual values (solid line) and the fitted values (dashed line). The fitted values refer to an estimated 2nd-degree polynomial with intercept for each patch. Our results indicate the existence of six break patches for the World price index, five for Emerging Markets and Pacific, and four for Europe, where the number of patches is equal to the number of breakpoints plus one. Based on our results, the integrated signal of price returns for each patch tracks a sequence of partial quadratic functions, most of which having negative concavity, as shown in the various landscapes of Figure 2. The R^2 statistic for the polynomial fit with breaks varies between 0.83 (Europe) and 0.92 (World) and is significant at the 1% level in all cases.

The breakpoints we have identified using the Bai-Perron [17] methodology (see Table 2 and Figure 2) correspond, in some cases, to the end of spells of financial downturns such as the Dot-Com crash, the Global Financial crisis or the Chinese stock market crash, as described in subsection 3.1. Breakpoints in several financial markets were identified in the second semester of 2001 and 2002, in 2004 and 2005, in September and October of 2008, in 2012 and 2013 and in July 2015.

4.2. Squared Residuals

After generating the cumulative sum of price returns and the underlying 2nd-degree polynomial fits with structural breaks, we next proceed to the third step, wherein we compute the squared distance (or squared residuals) from the observed to the fitted values of the integrated series X_t.

The residuals (or detrended walk) were estimated from $u_t = X_t - \hat{X}_t$, where $\hat{X}_t$ represents the fitted series (local trend), that is, the piecewise sequence of quadratic regression fits for each patch. Residuals are stationary in mean, as indicated by conventional ADF tests without deterministic terms, not reported here.[4] Therefore, according to the Granger Representation Theorem [24] the actual and fitted series are cointegrated of first order, and hence we obtain nonspurious residuals. Despite this evidence, a Jarque-Bera [14] test of the null of gaussianity is rejected at the 1% level in the distribution of residuals in all cases. Alternative data distributions may be suggested by inspection of the squared residuals distributions represented in Figure 3.

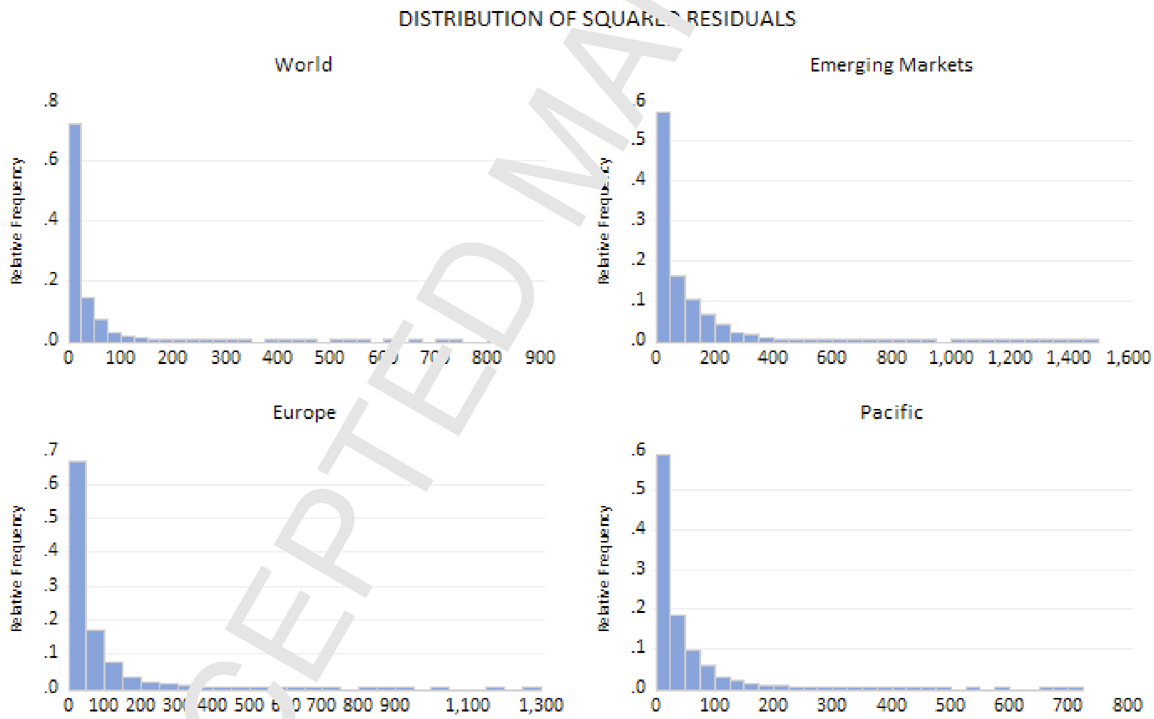


Figure 3: Distribution of squared residuals. MSCI daily price index, 07/13/1998 – 07/13/2018, 5220 obs.

4.3. Power-Law

[4] Outputs not reported are available from the authors upon request.

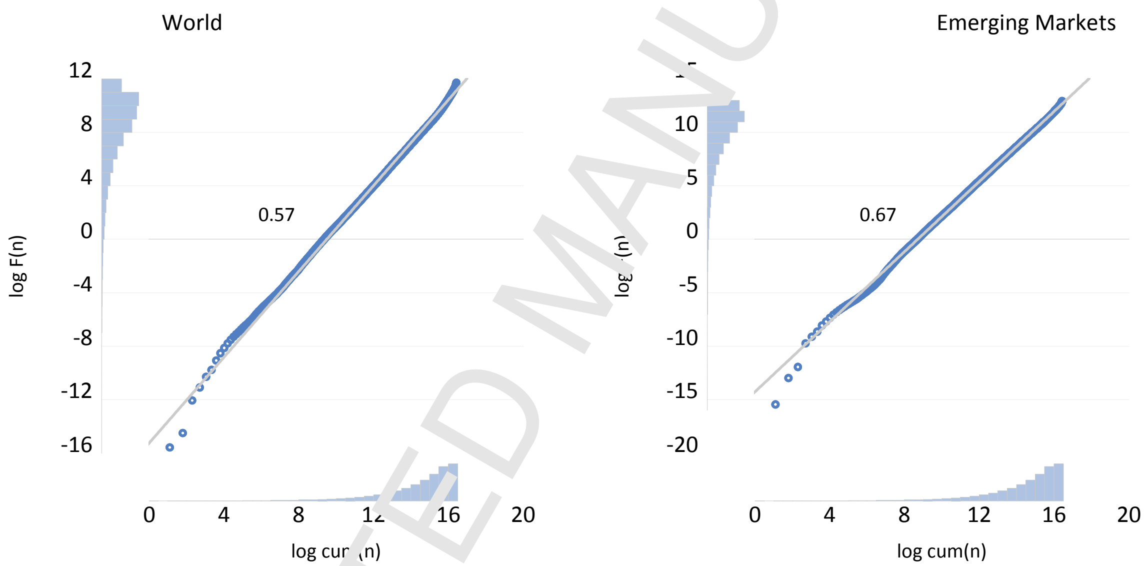
A natural candidate to replace the Gaussian distribution is the Zeta distribution modeled by Detrended Fluctuation Analysis (DFA) as suggested by Peng et al. [1]. This is also known as the power-law distribution [1, 10] with scaling parameter α where $F(n) \propto n^{\alpha}$ ($\alpha > 0$) is the cumulative distribution function of squared residuals. Note that $F(n) = \sum_{t=1}^{n} (X_t - \hat{X}_t)^2$ with $n \in \mathbb{N}$. Detrended Fluctuation Analysis follows the same route described above to obtain the integrated price return signal. However, the non-overlapped patches are determined in a very different way, since endogenous breakpoints occurring at different lengths (such as in our case) are not considered in DFA and there is no real rational behind the choice of equally lengthened boxes as suggested by [1]. Another difference in our procedure is that we do calculate $F(n)$ as the cumulative distribution function of the squared residuals rather than the average (over all the boxes) of the variances about the detrended walk for each box. In fact, the exponent α describes the scaling properties of the entire distribution, including the cumulative distribution function or some other relevant statistics such as the root mean squared error used by [1].

The scaling parameter α can be estimated as the slope of a normalized linear regression of the logarithm of $F(n)$ against the logarithm of $\sum n$. If $0 < \alpha < 0.5$ then the series is anti-persistent. Anti-persistence means that larger fluctuations are followed by smaller fluctuations and vice-versa. If $0.5 < \alpha < 1$ then the series has positive long-range dependence with persistent or long-memory behavior. If $\alpha = 0.5$ then the integrated price return series is white noise. This means that the autocorrelation function tends to zero and price returns have no significant long-memory. If $\alpha = 1$ then the integrated price return series is pink noise. The pink noise autocorrelation function does not decay exponentially as the lag length increases and memory tends to persist. Finally, when $\alpha > 1$, there is long-range correlation that cannot be explained by a power-law.[5] Therefore, the value of the scaling parameter α measures the degree of long-range correlation existing in the data [8].

Figure 4 depicts the scatter plots of the logarithm of the distribution function of the squared residuals for each integrated price returns index. This is displayed against the logarithm of the cumulative sum of n. The estimates of α obtained by running the regression of the log $F(n)$ on the log cum(n) are shown near each scatter plot. These estimates are the slopes of the regression lines also depicted in the figure (light grey line). As can be seen, the slopes are

[55] If an intercept is also estimated the interpretation of the α exponent should be adjusted.

very similar in the four market indexes ranging from 0.57 (World) to 0.67 (Emerging Markets). A Wald test for coefficient restrictions shows that all these estimates are significantly different from 0.5 and 1 at the 1% level. The estimates obtained for the scaling exponent indicate that squared residuals have positive long-range dependence with persistent or long-memory behavior. This behavior seems to be a little bit more mitigated in aggregated series such as the World index and aggravated in regional or sectorial indexes. Since squared residuals can be seen as a proxy to residuals' volatility, these results seem to indicate that fractional integrated models, such as the FIGARCH [25], may produce better estimates for long-memory volatility than those obtained by discrete integrated models, such as the I(0) (stationary) GARCH-type models, which assume no memory [25, 26].[6]



[6] However, a broader full discussion of this topic is outside of the scope of this paper.

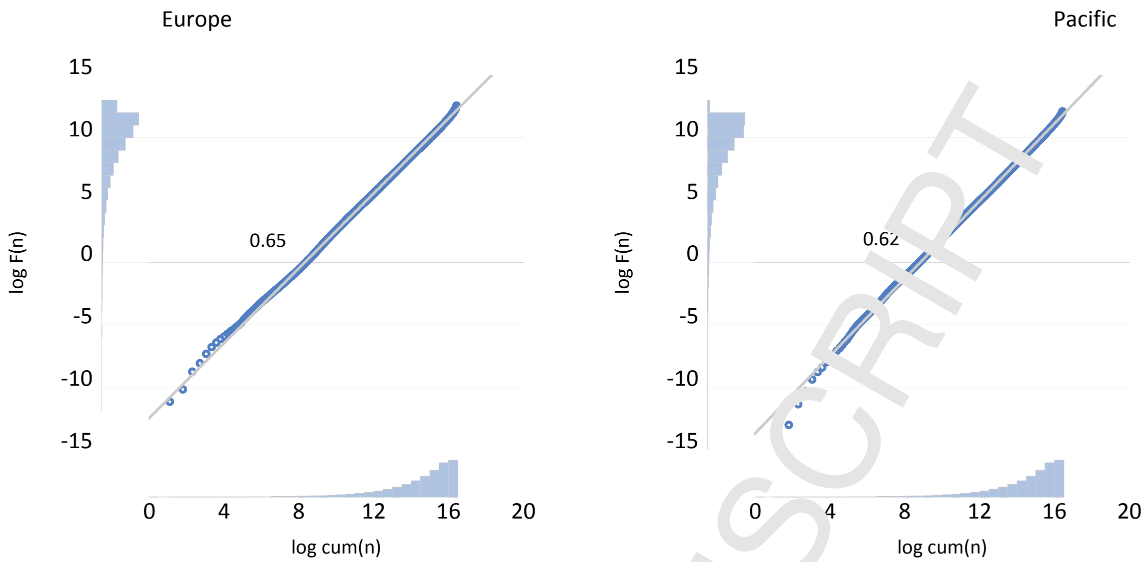


Figure 4: Detrended Fluctuation Analysis. MSCI daily price index 07/13/1998 – 07/13/2018, 5220 obs. The plot shows the scatter of log $F(n)$ on log cum(n) along with the underlying regression fit (grey line). The log $F(n)$ denotes the logarithm of the cumulative distribution function of squared residuals. The log cum(n) denotes the logarithm of the cumulative values of n. Histograms for each variable are shown near the axes. The estimated α exponent is reported near the scatter in each landscape.

Looking at the scatter plots in the four landscapes of Figure 4 we see that the data are very well fitted by the underlying regression line, except for the first few observations of each plot. The plot for Europe, however, yields smaller deviations between actual and fitted values, even for the first few observations, which is in line with the higher value of R^2 obtained for this fit (0.67) against 0.64 for the Emerging Markets fit, 0.63 for the Pacific fit and 0.57 for the World fit.

Another important issue concerns the size of the samples and the crossover value that triggers a shift in the classification of the long-range correlation in the data: for example, from anti-persistence to positive long-range correlation. To this end we perform a rolling regression of the log $F(n)$ on the log cum(n) where the sample size increases by one unit every fit and the first 260 estimations were discarded to avoid initial sample noisy effects. The results of the estimated α are displayed in Figure 5. All estimates lie within the range 0.16-0.67 and as the sample size increases, the α exponent tends asymptotically to its estimate obtained with the full sample (Figure 4 and 5). Starting values lie in the range 0.16-0.31. The α coefficient grows about 0.07-0.09 units per every 1000 days but this figure conceals a large percentage variation between the four indexes, ranging from 10% in Europe to 170% in the World series.

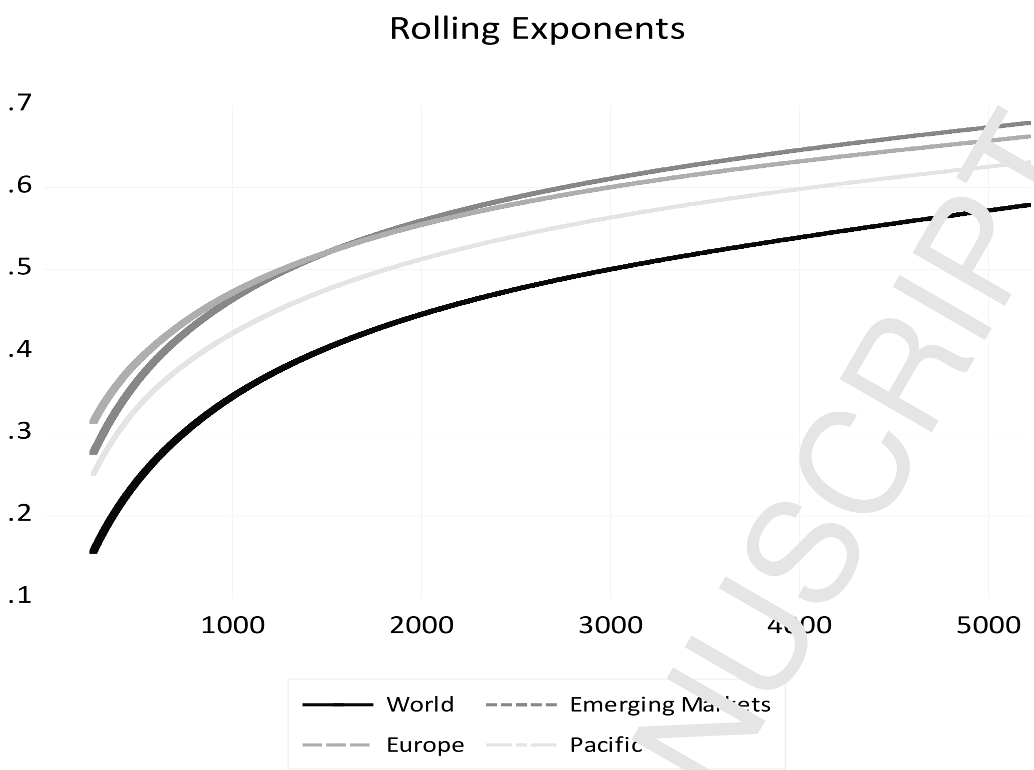


Figure 5: **Rolling α exponents.** MSCI daily price index, 07/13/1998 – 07/13/2018, 5220 obs. The plot shows the values of α when the regression sample anchored at start increases by one unit every fit. The first 260 estimations were discarded to avoid initial sample noise effects. The power-law exponent α varies between 0.16 and 0.31 at the start and 0.57 and 0.67 at the end. The cutoff point 0.5 occurs at observation #2988 (World), #1289 (Emerging Markets), #1259 (Europe) and #1807 (Pacific).

The crossover value of interest is 0.5. An α below this value indicates that the series is anti-persistent, that is large and small fluctuations alternate over time with relatively short-range correlations. Conversely, an α above this value indicates that the series has positive long-range correlations with long-memory behavior. The evidence of long-memory increases as $\alpha \to 1$. Finally, at the crossover value of 0.5, the series is a white noise random sequence. In our case, the World series "needs" about 10 years of daily data, excluding discarded estimations, to shift from short-range (or anti-persistence) to long-range correlation effects, the Pacific series needs about 6 years of data and Emerging Markets and Europe need about 4 years of data. The leading conclusion of these results is that researchers should be careful with sample size when dealing with detrended fluctuation analysis (or other models to assess long-range correlation), since small samples may bias the results and conclusions towards short-range correlation. Another problem may arise when the researcher opts to artificially split the full sample into a number of independent patches where the size of each patch is small. Therefore, using endogenous breaks to obtain non-overlapped and independent boxes may provide a relevant alternative to conventional detrended fluctuation analysis.

5. Forecast Accuracy

We now turn to evaluate the forecast accuracy of our modified model of detrended fluctuation analysis with endogenous structural breaks, using *ex post* in-sample and out-of-sample forecast accuracy gauges. We use three tests of in-sample forecast accuracy: Root Mean Squared Error (RMSE), Theil inequality coefficient (U) and Theil UII coefficient [27]. Since our variables are all standardized, we do not face problems of scale in the dependent variable and the results for different models can be directly compared. Under these circumstances, the smaller the error the better the forecasting ability of the model according to the RMSE test. Likewise, smaller Theil coefficients U and UII indicate better forecast ability of the model. For example, UII reaches its lower boundary of UII = 0 at perfect forecasts. In-sample tests use the full sample for forecasting purposes and then compare the forecasts with the actual values observed over the full sample.

On the other hand, out-of-sample tests use a sub-sample for model estimation and then compare the forecasts for the remaining sample with the actual values observed. The popular Diebold-Mariano test [28] will be employed in our case, where the null hypothesis states equal accuracy of the two models under consideration. A test of model adequacy using Monte Carlo simulation or otherwise could be important to perform in future work.

In order to perform these tests, we computed the four above mentioned gauges for our model with endogenous structural breaks and for three alternatives following the recommendation of Peng et al. [1] where we split the full sample into several equal length patches using an arbitrary size known *a priori* for each patch. In our case we used the following window sizes to obtain local trends: 100, 300 and 600. The results are presented in Table 3.

Model	World				Emerging Markets			
	out-of-sample	in-sample			out-of-sample	in-sample		
	DM	RMSE	Theil U	Theil UII	DM	RMSE	Theil U	Theil UII
l: Breaks		2.075	0.124	252.101		1.938	0.099	191.968
l = 100	15.767 **	2.249	0.160	421.073	29.437 **	2.22[illegible]	0.145	343.365
l = 300	6.912 **	2.159	0.137	323.919	25.823 **	2.022	0.115	249.343
l = 600	-5.555 **	2.069	0.122	224.074	-0.843	1.96[illegible]	0.101	188.990

Model	Europe				Pacific			
	out-of-sample	in-sample			out-of-sample	in-sample		
	DM	RMSE	Theil U	Theil UII	DM	RMSE	Theil U	Theil UII
l: Breaks		1.697	0.089	166.036		1.862	0.102	200.702
l = 100	31.006 **	2.215	0.150	326.691	24.529 **	2.177	0.149	336.533
l = 300	24.481 **	2.021	0.119	243.815	19.767 **	1.961	0.116	207.571
l = 600	21.290 **	2.077	0.118	298.446	7.330 **	2.035	0.112	220.106

Table 3. Forecast accuracy of integrated signals. Sample: MSCI daily price index, 07/13/1998 – 07/13/2018, 5220 observations. Forecast tests: *ex post* out-of-sample Diebold-Mariano (DM) [28] forecast evaluation test and in-sample RMSE, Theil U, and Theil UII tests [27]. Models to compare: forecasts using endogenous structural breaks against local forecasts using patch sizes of 100, 300 and 600 observations. DM null: the two forecasts have the same accuracy. Loss function: squared. DM forecast sample: 07/13/2012 – 07/13/2018. RMSE and Theil tests should be minimized. ** significant at the 1% level.

We first look at the Diebold-Mariano [28] out-of-sample forecast accuracy results displayed in the first column of each block in Table 3. The null hypothesis of equal forecast accuracy is rejected at the 1% level in all cases except one (test l: Breaks against l = 600 for the Emerging Markets index). Under rejection of the null it becomes evident that the models produce different forecast accuracies but the DM test itself does not indicate which model produces the best forecasts. In this case, after knowing that the models produce different levels of accuracy, we may rely on the in-sample tests to make the adequate selection.

Regarding the in-sample forecasts, the results presented above indicate that our model of DFA with endogenous structural breaks performs, in general, better than the alternatives of equal length patches (100, 300 and 600) except for the World index where the l = 600 model appears to perform slightly better.[7] Thus, in spite of some potential limitations caused by the level of World data aggregation, our model appears to outperform conventional models of DFA by imposing less restrictions and especially avoiding the prior selection of patch size.

[7] Nevertheless, the model with structural breaks always performs better than the l = 100 and l = 300 alternatives.

6. Conclusions

The use of statistical tools to determine the distributional shape of some financial time-series has become widespread in the literature since the recognition of a few stylized features of these data. These features include volatility clustering, long-memory, fat tails, and leverage effect, among others, which overturns the usual Gaussian assumption. In statistical physics, some researchers have used the Riemann Zeta distribution to explain the statistical behavior of these phenomena. This is known in the literature as the power-law distribution which describes a large number of real phenomena. The same occurs with financial data.

Some models and techniques to deal with power-laws have been proposed in the literature. One such technique is called Detrended Fluctuation Analysis and was originally suggested by Peng et al. [1] in a study of DNA nucleotides. Applications to financial data were also made but some oddities of these data may recommend a few modifications in the original DFA model. One such modification refers to the occurrence of structural breaks in financial data, which recommends that regularly spaced patches are replaced by irregularly spaced intervals where endogenously determined break dates are used as the starting date of a new window. Another modification refers to the replacement of within patch standard errors by daily observations on squared residuals, in conformity with the idea that squared residuals are a proxy for volatility. Finally, the selection of the "best" model in our case is based on forecast accuracy criteria instead of pure visual inspection of overlapped results. Our process, therefore, is more robust and independent of individual researcher judgement than in the conventional DFA model.

Our results show that the modified DFA model performs in general better than the conventional one, where this judgement is made on the basis of pure analytic in-sample and out-of-sample tests. Therefore, structural breaks seem to contain relevant information for the construction of boxes into which the full sample should be split. Another relevant information concerns the determination of the sample size required to reach the crossover value 0.5. Although it changes with the level of data aggregation, we can say that at least four years of daily data are required to obtain stable estimates of the power-law exponent. As in some previous studies, the results obtained exhibit positive long-range correlation, which degree decreases with the level of index aggregation.

Acknowledgements

This work was supported by Fundação para a Ciência e a Tecnologia, grant UID/GES/00315/2013

References

[1] C.-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Mosaic organization of DNA nucleotides, *Phys. Rev. E* 49 (1994) 1685-1689.

[2] M.A.R. Ferreira, M. Westy, H.K.H. Leez, D.M. Higdon, Multi-scale and hidden resolution time series models, *Bayesian Anal.* 1 (2006) 947-968, doi:10.1214/06-BA131.

[3] P. Oświęcimka, J. Kwapień, S. Drożdż, Multifractality in the stock market: price increments versus waiting times, *Phys. A* 347 (2005) 626-638, doi: 10.1016/j.physa.2004.08.025.

[4] L. Zunino, B.M. Tabak, A. Figliola, D.G. Pérez, M. Garavaglia, O.A. Rosso, A multifractal approach for stock market inefficiency, *Phys. A* 387 (2008) 6558-6566, doi:10.1016/j. physa.2008.08.028.

[5] X. Lu, J. Tian, Y. Zhou, Z. Li, Multifractal detrended fluctuation analysis of the Chinese stock index futures market, AUT University (NZ) Department of Economics, *Working Paper Series 2012-08* (2012).

[6] Q. Wang, Y. Zhu, L. Yang, R.A.H. Mul, Coupling detrended fluctuation analysis of Asian stock markets, *Phys. A.* 471 (2017) 337-350, doi:10.1016/j.physa.2016.12.076.

[7] W. Mohti, A. Dionísio, P. Ferreira, I. Vieira, Frontier markets' efficiency: mutual information and detrended fluctuation analyses, *J. Econ. Interact. Coord.* (2018) (forthcoming) doi:10.1007/s11403-018-0224-9.

[8] J.F. Eichner, E. Koscielny-Bunde, A. Bunde, S. Havlin, H.-J. Schellnhuber, Power-law persistence and trends in the atmosphere: A detailed study of long temperature records, *Phys. Rev. E* 68 (2003) 046133-1:5, doi:10.1103/PhysRevE.68.046133.

[9] A. Clauset, C.R. Shalizi, M.E.J. Newman, Power-law distributions in empirical data, *Siam Rev.* 51 (2009) 661-703, doi:10.1137/070710111.

[10] Z. Chen, P. Ch. Ivanov, K. Hu, H.E. Stanley, Effect of nonstationarities on detrended fluctuation analysis, *Phys. Rev. E* 65, 041107 (2002).

[11] Z. Chen, K. Hu, P. Carpena, P. Bernaola-Galvan, H.E. Stanley, P. Ch. Ivanov, Effect of nonlinear filters on detrended fluctuation analysis, *Phys. Rev. E* 71, 011104 (2005).

[12] D. Horvatic, H.E. Stanley, B. Podobnik, Detrended Cross-Correlation analysis for non-stationary time series with periodic trends, *Europhys. Lett.* (EPL) 94, 18007 (2011).

[13] B. Podobnik, Z.-Q. Jiang, W.-X. Zhou, H.E. Stanley, Statistical tests for power-law cross-correlated processes, *Phys. Rev. E* 84, 066118 (2011).

[14] C.M. Jarque, A.K. Bera, A test for normality of observations and regression residuals, *Int. Stat. Rev.* 55 (1987) 163-172 doi:10.2307/1403192.

[15] Y. Shachmurove, A historical overview of financial crises in the United States, *Global Fin. J.* 22 (2011) 217-231, doi:10.1016/j.gfj.2011.10.012.

[16] P. Perron, Further evidence on breaking trend functions in macroeconomic variables, *J. Econom.* 80 (1997) 355-385, doi:10.1016/S0304-4076(97)00049-3.

[17] J. Bai, P. Perron, Critical values for multiple structural change tests, *Econom. J.* 6 (2003) 72-78, doi:10.1111/1368-423X.00102.

[18] D.A Dickey, W.A. Fuller (1979), Distribution of the estimators for autoregressive time series with a unit root, *J. Am. Stat. Assoc.* 74 (1979) 427-431, doi:10.2307/2286348.

[19] J.G. Mackinnon, Numerical distribution functions for unit root and cointegration tests, *J. Appl. Econom.* 11 (1996) 601-618, doi:10.1002/jae.2295.

[20] P. Perron, The Great Crash, the oil price shock and the unit root hypothesis, *Econometrica* 57 (1989) 1361-1401, doi:10.2307/1913712.

[21] A.R. Hall, D.R. Osborn, N. Sakkas, Inference on structural breaks using Information Criteria, *Manchester Sch.* 81 (2013) 54-81, doi:10.1111/manc.12017.

[22] Y.C. Yao, Estimating the number of change-points via Schwarz' Criterion, *Stat. Prob. Lett.* 6 (1988) 181-189, doi:10.1016/0167-7152(88)90118-6.

[23] J. Liu, S. Wu, J.V. Zidek, On segmented multivariate regressions, *Stat. Sin.* 7 (1997) 497-525.

[24] R.F. Engle, C.W.J. Granger, Co-integration and error correction: Representation, estimation and testing, *Econometrica* 55 (1987) 251-276, doi:10.2307/1913236.

[25] R. Baillie, T. Bollerslev, H. Mikkelsen, Fractionally integrated generalized autoregressive conditional heteroscedasticity. *J. Econom.* 74 (1996) 3-30.

[26] J. Tenenbaum, D. Horvatic, S. Cosovic Bajic, B. Pehlivanovic, B. Podobnik, H.E. Stanley, Comparison between response dynamics in transition economies and developed countries, *Phys. Rev. E* 82, 046104 (2010).

[27] H. Theil, *Applied Economic Forecasting*, Amsterdam: North Holland Publishing Co.; Chicago: Rand McNally & Co. 1966.

[28] F.X. Diebold, R.S. Mariano, Comparing predictive accuracy, *J. Bus. Econ. Stat.* 20 (2002) 134-144, doi:10.1198/073500102753410444.