Introduction

In a recent paper, Ferreira and Dionísio (2016) investigated how long the memory of USA stock market is. With a 20 years sample, since 1995 to 2014 (with almost 5.000 observations), they computed, sequentially, both linear and nonlinear correlations between return rates and its lags. As usual, they confirmed that linear correlations, measured by Pearson coefficient, quickly tend to zero. However, using detrended cross-correlation analysis (DCCA) and its correlation coefficient, correlations remain significant until, approximately, 150th lag. In this paper, we extend the analysis of that paper by two different ways: firstly, we use an extended panel of countries, with all the G7 countries (Canada, France, Germany, Italy, Japan and United Kingdom, besides the referred USA); secondly we also extend the sample of analysis – beginning on 1972 and ending on 2015, in a total of 11479 observations.

This kind of analysis is inserted in a great group of studies that have the objective to study the behavior of financial markets. In fact, a larger amount of studies is dedicated to the Efficient Market Hypothesis (EMH), which is one of the most important hypotheses in financial economics. According to EMH, it is not possible to identify any deterministic pattern in its time series behavior (implying that EMH is verified in its weak form if). In other words, it means that, through arbitrage, agents could not obtain systematic abnormal profits using past information (Fama, 1970).

As previously referred, we can find in the literature a large amount of studies dealing with this problem. This analysis lasts for over a century: probably the study of Bachelier (1900) is the first one that tries to explain the random walk behavior of stock prices. Although with some interval span, some other studies corroborated this important finding: Fama (1963), Osborne (1964) or Granger and Morgenstein (1964) are just some of the most important studies on this theme, in the middle of the 20th century. These and other studies stated that, when linear autocorrelation exist between return rates, they quickly disappear.

Despite the importance of linear autocorrelations, some authors argue that return rates could suffer of other kind of dependence, namely long-range dependency. In this context, this kind of dependence could be interpreted as temporal or sectional dependence in return rates which could imply some capacity to predict financial series which, if true, could violate the EMH. The possibility of other kinds of dependence is known in the literature by stylized facts. The main stylized facts found are the existence of fat tails, asymmetries in gains and losses, volatility clustering behavior, leverage effect and the existence of autocorrelation in variance. The work of Cont (2001) makes a survey on several stylized facts. But that analysis remains interesting and later studies demonstrate it (see, for example, Malmsten and Teräsvirta, 2004 or Nystrup et al., 2015). Due to the large amount of papers that analyze EMH, it is not possible to make a simple literature review on it. For a more complete literature review on EMH see, for example, the work of Sewell, 2011.

The behavior of financial markets as complex systems, with large amount of available data, attracted the attention of physicists. In recent years, the behavior of financial markets was frequently studied, using measures and methodologies with origin on statistical physics. Inclusively, a new research field was born: Econophysics. In this new field, multidisciplinary researchers are able to study economic issues, especially in financial markets. Our objective is not to carry out an extensive literature review on Econophysics. You can find some reviews in works of Jovanovic and Schinckus (2013) or Schinckus (2013), among others.

Following the methodology of Ferreira and Dionísio (2016), the objective of this paper is to analyze the behavior stock markets in the G7 countries and find which of those countries is the first to reach levels of long-range correlations that are not significant. We carry out this analysis using detrended cross-correlation analysis and its correlation coefficient, to check for the existence of long-range dependence in time series. The existence of long-range dependence could be understood as a possibility of EMH violation. This analysis remains interesting because studies are not conclusive about the existence or not of long memory in stock return rates.

1. Methodology

The main goal of this paper is to analyse the presence of long range memory in returns. This result could be related with the existence of nonlinear dependence in return rates which may not be detected with the corresponding linear tests (see, for example, Darbellay, 1998 or Granger *et al.*, 2004).

In order to evaluate the global dependence in and within returns we use a nonlinear approach based on detrended fluctuation analysis (DFA). Created by Peng *et al.* (1994), this methodology studies the behavior of individual series and has several applications to financial markets (see, for example, the work of Cizeau et al., 1997, Ausloos et al., 1999 or Ferreira and Dionísio, 2014, among others). The great advantage of DFA is the fact that it can be used on both stationary and non-stationary series, while linear approaches can only be used on the former.

Besides DFA, detrended cross-correlation analysis (DCCA) can also be used to study long-range dependence. However, DCCA is used not to study the long-range behavior of one time series but the behavior between time series, namely its long-range cross-correlation. Created by Podobnik and Stanley (2008), it has the advantage of also being used in non-stationary time series.

DCCA gives us information about cross correlation between series but does not quantify that relation. In order to make that quantification, from the results of DCCA between x and y and DFA for each series, Zebende (2011) created the correlation coefficient given by $\rho_{DCCA} = \frac{F^2_{DCCA}}{F_{DFA\{xi\}} F_{DFA\{yi\}}}$. This coefficient has the general properties of one correlation coefficient, namely $-1 \leq \rho_{DCCA} \leq 1$. A value of $\rho_{DCCA} = 0$ means that there is no cross-correlation between series, while a positive or negative value means, respectively, evidence of cross-correlation or anti cross-correlation between series.

According to Podobnik et al. (2011), we can test the significance of this correlation coefficient. We use that methodology to estimate critical points of our test, which is considered robust (see, for example, Kristoufek, 2014). We can find in the literature some papers using DCCA or its variations (where the correlation coefficient could be included) analyzing financial markets: Cao et al. (2012) or Wang et al. (2013) are just two examples.

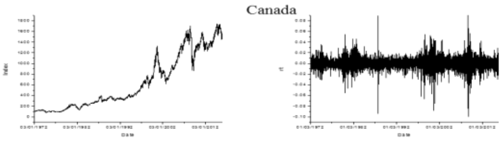
2. Results

We used information the Morgan Stanley Capital International (MSCI) stock indexes for G7 countries, between January 1972 and December 2015, with a total of 11479 observations[1]. Data was retrieved from Datastream. We chose the same kind of index for each country (the MSCI one) to better comparison. Each index was standardized to 100 for the first observation and, after this, we calculated the return rates for them, making the traditional difference between logarithms in two consecutive moments in time, i.e., $r_t = ln(I_t) - ln(I_{t-1})$, with I_t being the value of the index at moment *t*.

Figure 1 shows the evolution of each index (on the left) and of the correspondent return rate (at the middle). At the right, we have descriptive statistics of both variables. As expected, the mean of the return rate is very near to zero.

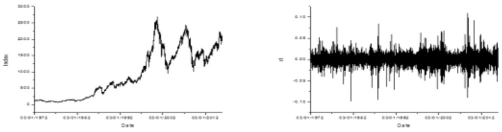
[1] For these countries, data exists since January 1970. However, until January 1972 the information was just updated on a weekly basis.

Canada



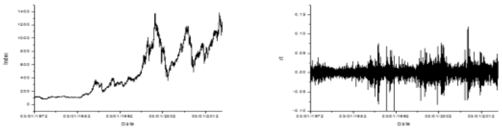
	Index	r_t
Mean	634,3368	0,0003
Std. Deviation	503,6012	0,0098
Maximum	1739,2265	0,1021
Minimum	79,4779	-0,0991

France



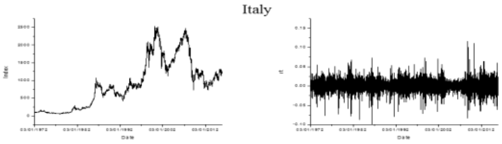
	Index	r_t
Mean	916,9599	0,0003
Std. Deviation	724,4427	0,0123
Maximum	2674,1455	0,1092
Minimum	65,4332	-0,0980

Germany



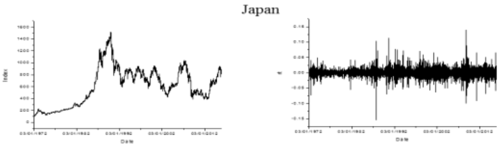
	Index	r_t
Mean	488,8095	0,0003
Std. Deviation	351,5585	0,0123
Maximum	1381,0465	0,1177
Minimum	79,0290	-0,1304

Italy



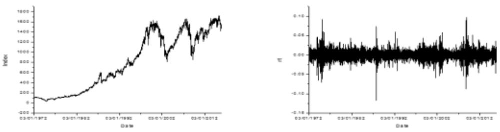
	Index	r_t
Mean	894,5644	0,0003
Std. Deviation	676,6202	0,0138
Maximum	2534,2758	0,1161
Minimum	54,8865	-0,1076

Japan



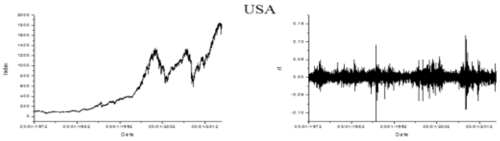
	Index	r_t
Mean	608,7686	0,0003
Std. Deviation	315,2616	0,0118
Maximum	1520,1532	0,1395
Minimum	100,0000	-0,1535

UK



	Index	r_t
Mean	789,3741	0,0003
Std. Deviation	552,4286	0,0113
Maximum	1713,8964	0,0971
Minimum	32,4813	-0,1178

USA



	Index	r_t
Mean	599,5450	0,0003
Std. Deviation	496,5932	0,0106
Maximum	1848,5577	0,1168
Minimum	57,5057	-0,2041

Figure 1: Stock indexes and return rates

Applying DFA to the return rate, we obtain the exponents presented on **Table 1**. Canada, France, Germany and Italy show some persistent patter, while UK and USA stock markets present some anti-persistence. The Japanese stock shows a behavior near to a random-walk. In fact, if we perform the hypothesis test for a parameter equal to 0.5, the conclusion, we do not reject that hypothesis. Some of these results are coherent with previous studies (see, for example, Ferreira and Dionísio, 2014, 2016).

Table 1: DFA results for stock market returns.

Country	DFA exponent
Canada	0.5144 ± 0.0037
France	0.5134 ± 0.0040
Germany	0.5166 ± 0.0051
Italy	0.5308 ± 0.0049
Japan	0.5047 ± 0.0049
UK	0.4672 ± 0.0057
USA	0.4716 ± 0.0038

With the return rates, we calculated successively the correlation between r_t and r_{t-i} , from $i = 1,...,200$. Firstly, we applied Pearson's correlation coefficient and we can conclude that linear correlations quickly turn non-significant in all countries, although in different lags[2]. However, our objective is to study the long-range

[2] Due to space constraints, we do not present every results but they would be available upon request. However, it is possible to say that for the USA and the German market the correlation is not significant immediately for the first lag, in Canada and Italy it is not significant in the third lag, Japan in the fourth, UK in the fifth and France in the sixth lag.

relationship of return rates. We calculated DCCA and the respective correlation coefficient between r_t and r_{t-i} .[3]

Figure 2 shows the behavior of the long-range correlation coefficient: the horizontal axis measures length boxes and on the vertical one the correlation coefficient. We chose some lags (from *t-1* to *t-200*) to show the results, comparing with lower and upper limits that allow the rejection of absence of correlation (with a 99% confidence level). If the correlation coefficient is outside the limits, then the correlation is statistically significant.

[3] The paper by Zebende et al. (2013) shows that this kind of analysis can be robust for simulated series, applying the ARFIMA process (long tail).

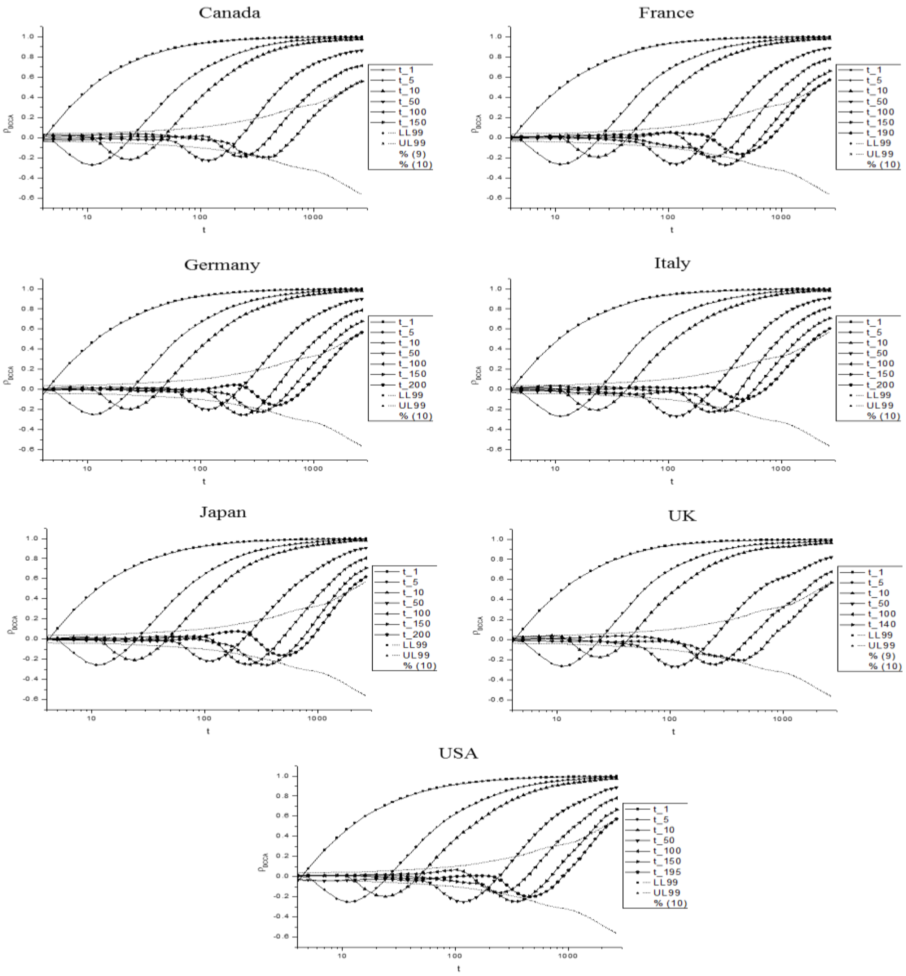


Figure 2: Detrended cross-correlation coefficients for the G7 return rates – t (days) is the time scale

Generically, all indexes have correlation coefficients that quickly become zero, for smaller boxes, the. However, for higher boxes, the correlation is highly significant. In first lags the correlation tends toward 1, meaning that, for shorter lags, there is evidence of strong long-range correlations. However, more distant is the lag, the lower is the correlation coefficient. This is an expected result, once is expected that return rates memory decay over time. Somehow surprisingly is the fact that the coefficient do not show evidence of quick decay, whatever the index we analyse.
The first index the reach the non-significant correlation, with a significance level of 1%, is the British one but just about at the 140th lag. It is followed by the Canadian index (150th lag), the French index (190th lag), the American index (195th lag), the German index (200th lag) and the Italian and Japanese indexes (both in the 210th lag). It means that some indexes have about ten months of memory in their return rates (we are working on a weekday base)[4].

Concluding Remarks

The verification or not of EMH is one of the most studied topics on finance. Recently, the advent of Econophysics increased the number of studies on this hypothesis, with the application of different methodologies that allow studying the long-range dependence of variables, even in the presence of non-stationary time series.

In this paper we propose to compare, between the G7 countries, which are those which have longer long-range dependence. With the implementation of DCCA and its correlation coefficient, robust methodologies to evaluate and analyse serial and nonlinear dependence, we find that all the seven indexes show evidence of (very) long dependence. The first index to reach the non-significance is the Canadian, while the Italian and the Japanese are those which have longer memory. Could we

[4] The mentioned study of Ferreira and Dionísio (2016) find that the memory of the USA stock market lasts for about 150 days. In this paper, the memory of that index goes to, approximately, 195 days. One possible explanation is the different sample of both studies.

conclude about (in) efficiency with those results? Probably no, first we should prove that this serial dependence promotes systematic and abnormal profits.

Acknowledgments

The authors are pleased to acknowledge financial support from Fundação para a Ciência e a Tecnologia (grant UID/ECO/04007/2013) and FEDER/COMPETE (POCI-01-0145-FEDER-007659)

References

Ausloos, M., Vandewalle, N., Boveroux, P., Minguet A., Ivanova, K. (1999). Applications of Statistical Physics to Economic and Financial Topics. *Physica A*.(274), 229-240.

Bachelier, L. (1964). "Theory of Speculation", in (ED. P. Cootner), The Random Character of Stock Prices. Cambridge: MIT Press.

Cao, G., Xu, L., Cao, J. (2012). Multifractal Detrended Cross-Correlations Between The Chinese Exchange Market and Stock Market. *Physica A*, (391), 4855–4866.

Cizeau P, Liu, Y., Meyer, M., Peng, C., Stanley, H. (1999). "Correlations in Economic Time Series". arXiv:condmat/9706021v1.

Cont, R. (2001). Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. *Quantitative Finance*, (1), 223-236.

Darbellay, G. (1998). An Adaptative Histogram Estimator for the Mutual Information. *UTIA Research Report,* (1889), Prague: Acad. Sc.

Darbellay, G., Wuertz, D. (2000). The Entropy as A Tool for Analysing Statistical Dependences in Financial Time Series. *Physica A*. (287), 429-439.

Fama, E. 1970. Efficient capital markets: A review of theory and empirical work, *The Journal of Finance* 25, 383–417.

Fama, E. 1963. Mandelbrot and the Stable Paretian Hypothesis. *Journal of Business*. 36: 420-429.

Ferreira, P., Dionísio A. (2014). Revisiting Serial Dependence in The Stock Markets of The G7 Countries, Portugal, Spain and Greece, *Applied Financial Economics*, (24), 319-331.

Ferreira, P., Dionísio, A. (2016). How Long is The Memory of The USA Stock Market?, *Physica A*, Doi:10.1016/j.physa.2016.01.080

Granger, C., and Morgenstein, O. (1964). "Spectral Analysis of New York Stock Market Prices", in (ed. P. Cootner) *The Random Character of Stock Prices.* Cambridge: MIT Press.

Granger, C., Maasoumi, E. and Racine, J. (2004). A Dependence Metric for Possibly Nonlinear Processes *Journal of Time Series Analysis*,(25), 649-669.

Jovanovic F.and Schinckus, C. (2013). The Emergence of Econophysics: A New Approach in Modern Financial Theory. *History of Political Economy*, (45), 443-474.

Kristoufek, L. (2014). Measuring Correlations Between Non-Stationary Series with DCCA Coefficient. *Physica A*(402), 291-298.

Malmsten, H., Teräsvirta, T. (2004). Stylized Facts of Financial Time Series and Three Popular Models of Volatility. *Department of Economic Statistics, Stockholm School of Economics Working Paper 563*.

Nystrup, P., Madsen, H. and Lindström, E. (2015). Stylised Facts of Financial Time Series and Hidden Markov Models in Continuous Time, *Quantitative Finance, 15*(9), 1531-1541. Doi: 10.1080/14697688.2015.1004801

Osborne, M. (1964). "Brownian Motion in The Stock Prices", in (Ed. P. Cootner), *The Random Character of Stock Prices* Cambridge: MIT Press.

Peng, C., Buldyrev, S., Havlin, S., Simons, M., Stanley, H. and Goldberger, A. (1994). Mosaic Organization of DNA Nucleotides. *Physical Review E*. (49), 1685-1689.

Podobnik, B., and Stanley, H. (2008). Detrended Cross-Correlation Analysis: A New Method for Analyzing Two Non-Stationary Time Series. *Physical Review Letters*. (100), 084102.

Podobnik, B., Jiang, Z., Zhou, W., and Stanley, H. (2011). Statistical Tests for Power-Law Cross-Correlated Processes, *Physical Review E*, (84), 066118.

Podobnik, B., Horvatic, D., Petersen A., and Stanley, H. (2009). Cross-Correlations Between Volume Change and Price Change. *Proc. Natl. Acad. Sci. U.S.A.* (106), 22079–22084.

Podobnik, B., Wang, D., Horvatic, D., Gosse, I., and Stanley, H. (2010). Time-Lag Cross-Correlations in Collective Phenomena. *EPL*, (90) 68001. Doi: 10.1209/0295-5075/90/68001

Schinckus, C. (2013). Introduction to Econophysics: Towards A New Step in The Evolution of Physical Sciences, *Contemporary Physics*, *54*(1), 17-32. http://dx.doi.org/10.1080/00107514.2013.774528

Sewell, M. (2011). History of The Efficient Market Hypothesis. *UCL Department of Computer Science Research Note RN / 11*(4).

Wang, G, Xie, C., Chen, Y., and Chen, S. (2013). Statistical Properties of the Foreign Exchange Network at Different Time Scales: Evidence from Detrended Cross-Correlation Coefficient and Minimum Spanning Tree. *Entropy*, (15), 1643-1662.

Zebende, G. (2011). DCCA Cross-Correlation Coefficient: Quantifying Level of Cross-Correlation, *Physica A*, (390), 614–618.

Zebende, G., Fernandes, M., and Machado, A. (2013). DCCA Cross-Correlation Coefficient Differentiation: Theoretical and Practical Approaches, *Physica A*, (392), 1756-1761.