Is Market Fear Persistent?
A Long-memory Analysis

June 2017

http://www.brunel.ac.uk/economics
IS MARKET FEAR PERSISTENT?
A LONG-MEMORY ANALYSIS

Guglielmo Maria Caporale*
Brunel University London, CESifo and DIW Berlin

Luis Gil-Alana**
University of Navarra

Alex Plastun
Sumy State University

June 2017

Abstract

This paper investigates the degree of persistence of market fear. Specifically, two different long-memory approaches (R/S analysis with the Hurst exponent method and fractional integration) are used to analyse persistence of the VIX index over the sample period 2004-2016, as well as some sub-periods (pre-crisis, crisis and post-crisis). The findings indicate that its properties change over time: in normal periods it exhibits anti-persistence (there is a negative correlation between its past and future values), whilst during crisis period the level of persistence is increasing. These results can be informative about the nature of financial bubbles and anti-bubbles, and provide evidence on whether there exist market inefficiencies that could be exploited to make abnormal profits by designing appropriate trading strategies.

Keywords: Market Fear, VIX, Persistence, Long Memory, R/S Analysis, Fractional Integration

JEL Classification: C22, G12

*Corresponding author. Department of Economics and Finance, Brunel University, London, UB8 3PH.

Email: Guglielmo-Maria.Caporale@brunel.ac.uk

** Luis A. Gil-Alana gratefully acknowledges financial support from the Ministerio de Ciencia y Tecnología (ECO2014-55236).
1. Introduction

According to an old saying on Wall Street the market is driven by just two emotions: fear and greed. Shefrin (2000) in his famous book “Beyond Greed and Fear” claims that they are the most important of a number of heuristic-driven biases influencing investors and resulting in market inefficiencies. Zweig (2007) points out that agents are often in the grip of emotions without even realising it and provides an interesting example: a survey of 1,000 investors suggested that there is a 51 percent chance that in any given year the US stock market will drop by one third, whilst on the basis of historical data the odds that US stocks will lose one third of their value in any given year are only around 2 percent. This misperception of reality is a direct result of fear and represents evidence that investors are not fully rational.

Shefrin (2000) points to behavioural anomalies in individual investors, institutional money managers, and corporate managers regardless of their training or experience; moreover, such anomalies can be observed in all market sectors, including equities, fixed income, foreign exchange, commodities, and options. Shiller (2003) argues that some changes in prices occur not for fundamental reasons, but because of mass psychology instead. Other behavioural finance studies have provided more evidence that is inconsistent with the Efficient Market Hypothesis (EMH), according to which investors are rational, and asset prices fully reflects all available information and therefore should follow a random walk (see Malkiel and Fama, 1970).

Analysing investor sentiment and fear in particular is therefore crucial. Surprisingly, to date there is no study investigating the long-memory properties of the latter. The present paper aims to fill this gap in the literature by examining the degree of persistence of the very popular VIX index, also known as CBOE Volatility Index. This can be interpreted as a “fear” index, since according to Whaley (2000) it is an ‘investor
fear gauge’ that reaches higher levels during periods of market turmoil. We use two different long-memory approaches (R/S analysis with the Hurst exponent method and fractional integration) to analyse persistence of the VIX over the sample period 2004-2016, as well as some sub-periods (pre-crisis, crisis and post-crisis) to see whether it varies over time depending on market conditions.

The results of our analysis are of interest to both academics and practitioners. They can be informative about the nature of financial bubbles and anti-bubbles, and provide evidence on whether there exist market inefficiencies that could be exploited to make abnormal profits by designing appropriate trading strategies. A better understanding of financial markets can be gained by applying quantitative methods to behavioural finance to analyse investor sentiment as in our study. Its layout is the following: Section 2 provides a brief review of the literature on market fear; Section 3 describes the data and outlines the methodology; Section 4 presents the empirical results; Section 5 provides some concluding remarks.

2. Literature Review

Fear is an unpleasant feeling of anticipation or awareness of danger. Sustained losses in financial markets can cause fear of further losses among investors. As a result phenomena such as financial bubbles (anti-bubbles) and their bust, volatility explosions, trends etc. may occur. When feeling fear investors move from risky assets (usually stocks) to less risky ones (money market securities, as well as so-called safe heavens, for example gold, the Swiss franc, the Japanese yen etc.). The mass exodus of investors from certain markets may cause market crashes or even financial crises. Fear is one of the reasons for market overreactions. A typical example is Black Monday (the Flash Crash in the US stock market on 19 October 1987): this single market drop mostly
caused by the emotional reactions of investors affected their behaviour for years (for more details see Keim and Madhavan, 2000). According to Baba Shiv et al. (2005) after incurring losses agents are less inclined to invest and prefer to stay out of the market. There is plenty of evidence that losing streaks influence their behaviour (see, e.g., Zalla et al., 2000; Breiter et al., 2001 etc.).

Losses or market instability make investors more vulnerable to fear, which often results in irrational behaviour and costly mistakes. Johnson and Tversky (1983) find that 50 percent of agents can recognise when they have been affected by a bit of negative news, but only 3 percent admit that this may influence their degree of risk aversion. Slovic (1987) proves that fast and finite dangers (fireworks, skydiving, train crashes, etc.) feel more “knowable” (and less worrisome) than vague, open-ended risks such as genetically modified foods or global warming. Agents underestimate the likelihood and severity of common risks, and overestimate those of rare risks (see Zweig, 2007).

A few studies have attempted to analyse market fear empirically using measures such as the VIX index (also known as the CBOE Volatility Index or Fear index), the CNN Money Fear & Greed Index, the IVX and the CBOE Skew Index. By far the most popular is the VIX, which is derived from the prices of S&P 500 options and yields the expected annualised change in the S&P 500 index over the following 30 days. It is an implied volatility index: the lower its level, the lower is demand from investors seeking to buy protection against risk and thus the lower is the level of market fear.

Most papers analysing the VIX have focused on its predictive power for future returns. Giot (2005) finds that high (low) levels of the VIX correspond to positive (negative) future returns. Guo and Whitelaw (2006) and Chow et al. (2014) also show that there is a positive relationship between market returns and the VIX. Heydon et al. (2000) find that global equity markets outperform bond markets after periods of
relatively high expected volatility in the US market and vice versa. Chow et al. (2016) estimate that approximately one-third of the VIX is attributable to the tail risk premium. Fleming et al. (1995) were the first to analyse the persistence of this index and found that its daily changes follow an AR(1) process, whilst its weekly changes exhibit mean reversion, and there is no evidence of seasonality. Long-memory behaviour in the VIX was also detected by Koopman et al. (2005) and Corsi (2009). Huskaj (2013) estimates GARCH, APARCH, FIGARCH and FIAPARCH models, and reports that long memory in the volatility has no significant impact on the prices of hypothetical VIX options. Jo-Hui and Yu-Fang (2014) apply ARFIMA and FIGARCH models to VIX-ETFS data and find no signs of long memory, whilst Fernandes et al. (2014) detect long-range dependence using a HAR model. Overall, the evidence on the properties of the VIX is rather mixed.

3. Data and Methodology


For robustness purposes we apply two different methods to measure persistence, namely R/S analysis and fractional integration (see Caporale et al., 2016 for an application to the Ukrainian stock market). Concerning the former, we use the following algorithm (see Mynhardt et al., 2014 for additional details):
1. A time series of length \( M \) is transformed into one of length \( N = M - 1 \) using logs and converting prices into returns:

\[
N_t = \log \left( \frac{Y_{t+1}}{Y_t} \right), \quad t = 1, 2, 3, \ldots (M - 1).
\]  

(1)

2. This period is divided into contiguous \( A \) sub-periods with length \( n \), so that \( A_n = N \), then each sub-period is identified as \( I_a \), given the fact that \( a = 1, 2, 3, \ldots, A \). Each element \( I_a \) is represented as \( N_k \) with \( k = 1, 2, 3, \ldots, N \). For each \( I_a \) with length \( n \) the average \( e_a \) is defined as:

\[
e_a = \frac{1}{n} \sum_{k=1}^{n} N_{k,a}, \quad k = 1, 2, 3, \ldots, N, \quad a = 1, 2, 3, \ldots, A.
\]  

(2)

3. Accumulated deviations \( X_{k,a} \) from the average \( e_a \) for each sub-period \( I_a \) are defined as:

\[
X_{k,a} = \sum_{i=1}^{k} (N_{i,a} - e_a).
\]  

(3)

The range is defined as the maximum index \( X_{k,a} \) minus the minimum \( X_{k,a} \), within each sub-period (\( I_a \)):

\[
R_{I_a} = \max(X_{k,a}) - \min(X_{k,a}), \quad 1 \leq k \leq n.
\]  

(4)

4. The standard deviation \( S_{I_a} \) is calculated for each sub-period \( I_a \):

\[
S_{I_a} = \left( \frac{1}{n} \sum_{k=1}^{n} (N_{k,a} - e_a)^2 \right)^{0.5}.
\]  

(5)

5. Each range \( R_{I_a} \) is normalised by dividing by the corresponding \( S_{I_a} \). Therefore, the re-normalised scale during each sub-period \( I_a \) is \( R_{I_a}/S_{I_a} \). In the step 2 above, adjacent sub-periods of length \( n \) are obtained. Thus, the average R/S for length \( n \) is defined as:
\[
(R/S)_n = \frac{1}{A} \sum_{i=1}^{A} \frac{R_{Ia}}{S_{Ia}}.
\]  

6. The length \( n \) is increased to the next higher level, \((M - 1)/n\), and must be an integer number. In this case, \( n \)-indexes that include the initial and ending points of the time series are used, and Steps 1 - 6 are repeated until \( n = (M - 1)/2 \).

7. The least square method is used to estimate the equation \( \log (R/S) = \log (c) + H \log (n) \). The angle of the regression line is an estimate of the Hurst exponent \( H \). (Hurst, 1951). This can be defined over the interval \([0, 1]\), and is calculated within the boundaries specified below:

- \( 0 \leq H < 0.5 \) – the data are fractal, the EMH is not confirmed, the distribution has fat tails, the series are anti-persistent, returns are negatively correlated, there is pink noise with frequent changes in the direction of price movements, trading in the market is riskier for individual participants.

- \( H = 0.5 \) – the data are random, the EMH cannot be rejected, asset prices may follow a random Brownian motion (Wiener process), the series are normally distributed; if the returns are uncorrelated, there is no memory in the series and they are white noise; traders cannot «beat» the market using any trading strategy. Nevertheless they can be weakly autocorrelated through stationary ARMA structures.

- \( 0.5 < H \leq 1 \) – the data are fractal, the EMH is not confirmed, the distribution has fat tails, the series are persistent, returns are highly correlated, there is black noise and a trend in the market.

This algorithm yields static estimates of market persistence. Its dynamic behaviour can be analysed by calculating the Hurst exponent for different windows. The procedure is the following: having obtained the first value of the Hurst exponent (for example, for the date 01.04.2004 using data for the period from 01.01.2004 to 31.03.2004), each of the following ones is calculated by shifting forward the “data
window”, where the size of the shift depends on the number of observations and a sufficient number of estimates is required to analyse the time-varying behaviour of the Hurst exponent. For example, if the shift equals 10, the second value is calculated for 10.04.2004 and characterizes the market over the period 10.01.2004 till 09.04.2004, and so on.

The second approach we follow to analyse persistence involves estimating parametric/semiparametric fractional integration or I(d) models. This type of models were originally proposed by Granger (1980) and Granger and Joyeux (1980); they were motivated by the observation that the estimated spectrum in many aggregated series exhibits a large value at the zero frequency, which is consistent with nonstationary behaviour; however, it becomes close to zero after differencing, which suggests over-differentiation. Examples of applications of fractional integration to financial time series data can be found in Barkoulas and Baum (1996), Barkoulas et al. (1997), Sadique and Silvapulle (2001), Henry (2002), Baillie et al. (2007), Caporale and Gil-Alana (2004), Gil-Alana and Moreno (2012), Abbritti et al. (2016) and Al-Shboul and Anwar (2016) among many others.

In the present study we adopt the following specification:

\[(1 - L)^d x_t = u_t, \quad t = 0, \pm 1, \ldots \tag{7}\]

where \(d\) can be any real value, \(L\) is the lag-operator \((Lx_t = x_{t-1})\) and \(u_t\) is I(0), defined for our purposes as a covariance stationary process with a spectral density function that is positive and finite at the zero frequency. Note that \(H\) and \(d\) are related through the equality \(H = d - 0.5\).

The most common semiparametric approach, which does not require any assumptions about \(u_t\), is based on the log-periodogram (see Geweke and Porter-Hudak, 1983), and was then extended and improved by Künsch (1986), Robinson (1995a),
Hurvich and Ray (1995), Velasco (1999a, 2000), Shimotsu and Phillips (2002) etc. For our purposes we use another method, which is essentially a local ‘Whittle estimator’ defined in the frequency domain using a band of high frequencies that degenerates to zero. The estimator is implicitly defined by:

\[ \hat{d} = \arg \min_d \left( \log C(d) - 2d \frac{1}{m} \sum_{s=1}^{m} \log \lambda_s \right) \], \hspace{1cm} (8)

\[ C(d) = \frac{1}{m} \sum_{s=1}^{m} I(\lambda_s) \lambda_s^{2d}, \quad \lambda_s = \frac{2\pi s}{T}, \quad \frac{m}{T} \to 0, \]

where \( m \) is a bandwidth parameter, and \( I(\lambda_s) \) is the periodogram of the raw time series, \( x_t \), given by:

\[ I(\lambda_s) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} x_t e^{i\lambda_s t} \right|^2, \]

and \( d \in (-0.5, 0.5) \). Under finiteness of the fourth moment and other mild conditions, Robinson (1995b) proved that:

\[ \sqrt{m} (\hat{d} - d_o) \to_d N(0, 1/4) \quad \text{as} \quad T \to \infty, \]

where \( d_o \) is the true value of \( d \). This estimator is robust to a certain degree of conditional heteroscedasticity and more efficient than other semiparametric alternatives; it has been further developed by Velasco (1999b), Velasco and Robinson (2000), Phillips and Shimotsu (2004, 2005) and Abadir et al. (2007).

The parametric estimation of \( d \) and the other model parameters can be carried out either in the frequency domain or in the time domain. For the former, Sowell (1992) analysed the exact maximum likelihood estimator of the parameters of the ARFIMA model, using a recursive procedure that allows a quick evaluation of the likelihood function. Other parametric methods for estimating \( d \) based on the frequency domain were put forward by Fox and Taqqu (1986), Dahlhaus (1989) etc. (see also Robinson,

In the following section we use both parametric (Robinson, 1994) and semiparametric (Robinson, 1995a, Abadir et al., 2007) techniques for testing and estimating the fractional differencing parameter $d$.

4. Empirical Results

The results of the R/S analysis on the return series for the whole sample and different sub-samples are presented in Table 1.

<table>
<thead>
<tr>
<th>Period</th>
<th>Daily frequency</th>
<th>Monthly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole sample</td>
<td>0.41</td>
<td>0.53</td>
</tr>
<tr>
<td>Pre-crisis (2004-2006)</td>
<td>0.44</td>
<td>0.36</td>
</tr>
<tr>
<td>Crisis (2007-2009)</td>
<td>0.46</td>
<td>0.70</td>
</tr>
<tr>
<td>Post-crisis (2010-2016)</td>
<td>0.41</td>
<td>0.47</td>
</tr>
</tbody>
</table>

As can be seen, market fear does not follow a random walk, and the estimates depend on the data frequency. For the daily data it is anti-persistent (returns are negatively correlated). However, in the case of monthly data during the crisis period it exhibits persistence (returns are positively correlated), which suggests that it becomes a self-driving feeling possibly resulting in price bubbles and anti-bubbles. Anti-persistence can be seen as evidence in favour of the Noise Market Hypothesis and implies that abnormal profits can be made through trading in financial markets.

The dynamic R/S analysis shows the evolution over time of persistence in the VIX. The daily and monthly results are shown in Figure 1 and 2 respectively. The number of steps and the data window size are selected such that the number of Hurst
exponent estimates equals the number of dividing periods (156 months for the case of Figure 1 or 26 half-year case of Figure 2).

Figure 1: Results of the dynamic R/S analysis for the daily data, 2004-2016 (step=55, data window=300)

Figure 2: Results of the dynamic R/S analysis for the monthly data, 2004-2016 (step=4, data window=48)

As can be seen the degree of persistence varies over the time, being higher during the crisis period. This confirms the previous static findings.
The results for the fractional integration methods and using the index prices are presented in Tables 2 and 3. First, we display in Table 2 the estimates of $d$ along with their corresponding 95% confidence interval based on a parametric method (Robinson, 1994), using both uncorrelated and autocorrelated (Bloomfield, 1973) errors.¹

Table 2: Results based on a parametric method with no autocorrelation

<table>
<thead>
<tr>
<th>Period</th>
<th>Daily frequency</th>
<th>Monthly frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole sample</td>
<td>0.83 (0.81, 0.86)</td>
<td>0.78 (0.66, 0.96)</td>
</tr>
<tr>
<td>Pre-crisis (2004-2006)</td>
<td>0.82 (0.76, 0.89)</td>
<td>0.32 (0.18, 0.54)</td>
</tr>
<tr>
<td>Crisis (2007-2009)</td>
<td>0.88 (0.78, 0.97)</td>
<td>0.96 (0.68, 1.52)</td>
</tr>
<tr>
<td>Post-crisis (2010-2016)</td>
<td>0.85 (0.81, 0.89)</td>
<td>0.52 (0.38, 0.78)</td>
</tr>
</tbody>
</table>

*: non-rejection cases of the I(1) hypothesis at the 5% level.

Table 3: Results based on a parametric method with autocorrelation

<table>
<thead>
<tr>
<th>Period</th>
<th>Daily frequency</th>
<th>Monthly frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole sample</td>
<td>0.81 (0.77, 0.86)</td>
<td>0.64 (0.52, 0.77)</td>
</tr>
<tr>
<td>Pre-crisis (2004-2006)</td>
<td>0.67 (0.61, 0.79)</td>
<td>0.35 (0.00, 0.71)</td>
</tr>
<tr>
<td>Crisis (2007-2009)</td>
<td>0.85 (0.79, 1.00)</td>
<td>0.59 (0.11, 1.02)</td>
</tr>
<tr>
<td>Post-crisis (2010-2016)</td>
<td>0.77 (0.74, 0.86)</td>
<td>0.29 (-0.12, 0.58)</td>
</tr>
</tbody>
</table>

*: non-rejection cases of the I(1) hypothesis at the 5% level.

It can be seen from Tables 2 and 3 that all the estimates are in the interval (0.5, 1), which implies non-stationarity and mean-reverting behaviour. This is corroborated by the semiparametric estimates displayed in Tables 3 and 4 (Abadir et al., 2007) for a selected group of bandwidth parameters.² However, when considering the three subsamples, i.e., pre-crisis, crisis, and post-crisis, it becomes apparent that there are

¹ Robinson’s (1994) method uses the Whittle function in the frequency domain, and the values reported in Tables 2 and 3 refer to the value of $d$ producing the lowest test statistic. An advantage of this method is that it remains valid in nonstationary contexts ($d \geq 0.5$).

² The optimal choice of the bandwidth parameters for semiparametric methods in the context of fractional integration is still unresolved; basically there is a trade-off between bias and variance.
substantial differences between them, the highest degrees of persistence being found in the crisis period, during which the values of \( d \) are higher than 0.8 in virtually all cases and the I(1) hypothesis cannot be rejected. This is in contrast with the results for the pre- and post-crisis periods where mean reversion (i.e. \( d < 1 \)) is found in all cases, which implies anti-persistent behaviour in the returns, consistently with the results from the R/S analysis.

Table 4a: Estimates of \( d \) based on a semiparametric method

<table>
<thead>
<tr>
<th>Daily frequency</th>
<th>( m = (T)^{0.4} )</th>
<th>( m = (T)^{0.5} )</th>
<th>( m = (T)^{0.6} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole sample</td>
<td>0.680</td>
<td>0.763</td>
<td>0.901</td>
</tr>
<tr>
<td>Pre-crisis (2004-2006)</td>
<td>0.588</td>
<td>0.578</td>
<td>0.652</td>
</tr>
<tr>
<td>Crisis (2007-2009)</td>
<td>0.958*</td>
<td>1.076*</td>
<td>0.990*</td>
</tr>
<tr>
<td>Post-crisis (2010-2016)</td>
<td>0.523</td>
<td>0.659</td>
<td>0.768</td>
</tr>
</tbody>
</table>

*: non-rejection cases of the I(1) hypothesis at the 5% level.

Table 4b: Estimates of \( d \) based on a semiparametric method

<table>
<thead>
<tr>
<th>Monthly frequency</th>
<th>( m = (T)^{0.4} )</th>
<th>( m = (T)^{0.5} )</th>
<th>( m = (T)^{0.6} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole sample</td>
<td>0.804</td>
<td>0.637</td>
<td>0.658</td>
</tr>
<tr>
<td>Pre-crisis (2004-2006)</td>
<td>0.485</td>
<td>0.500</td>
<td>0.436</td>
</tr>
<tr>
<td>Crisis (2007-2009)</td>
<td>0.719</td>
<td>0.861*</td>
<td>0.816*</td>
</tr>
<tr>
<td>Post-crisis (2010-2016)</td>
<td>0.443</td>
<td>0.387</td>
<td>0.490</td>
</tr>
</tbody>
</table>

*: non-rejection cases of the I(1) hypothesis at the 5% level.

5. Conclusions

Fear is one of the strongest emotions in financial markets; it can cause market crashes as well as price bubbles. However, its features are still relatively unexplored. We investigate its persistence properties by applying two different long-memory approaches (R/S analysis and fractional integration) to daily and monthly series for the VIX Index (the most commonly used quantitative measure of market fear) for the period from 2004
to 2016; using two different techniques, as well as analysing different data frequencies over a long sample including the most recent observations decreases the possibility of data snooping bias. Specifically, the R/S statistic is computed for the return series, whilst price indices are analysed applying I(d) techniques.

The results from the two approaches are consistent and indicate that market fear does not follow a random walk. It normally exhibits anti-persistence, but in crisis periods its persistence increases, which also suggests crowd effects. The fact that the long-memory properties of market fear are unstable and change over the time is an important finding that can lead to a better understanding of the behaviour of financial markets.
References


Hurst, H. 1951, Long-term storage of reservoirs, Transactions of the American Society of Civil Engineers, 116.


