18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

A New Procedure for Discriminating Between Long Memory and Shifting Means Alternatives

William Rea[1], Les Oxley[2] and Marco Reale[1]
[1]Department of Mathematics and Statistics, [2]Department of Economics and Finance, University of Canterbury, Private Bag 4800, Christchurch, New Zealand
Email: les.oxley@canterbury.ac.nz

Abstract: It is now recognized that long memory and structural change can easily be confused because the statistical properties of times series of lengths typical of financial and econometric series are similar for both models. The implications of long memory models or short memory with breaks for the pricing of financial assets (e.g. options) have been studied by other authorities who have concluded the pricing is significantly different between the two classes of models. Option pricing is also of interest to regulatory authorities. Thus it is of interest both from a finance theory and regulatory point of view whether the data are generated by a true long memory process or by something else.

We propose a new approach aimed at distinguishing between long memory and structural change. The approach, which utilizes the computational efficient methods based upon Atheoretical Regression Trees (ART), uses a null hypothesis of the data being a fractionally integrated series. It establishes through simulation the bivariate distribution of the fractional integration parameter, d, with regime length for simulated fractionally integrated series of given nominal d value and series length. Confidence intervals can be established empirically as contours on the simulated data. This bivariate distribution is then compared with the data for the time series under test. We apply these methods to the realized volatility series of 16 stocks in the Dow Jones Industrial Average in the 10 year period from the beginning of 1994 to the end 2003. We show that in these series the value of the fractional integration d parameter is not constant with time. Given that a constant d value is fundamental to the definition of a fractionally integrated process

Keywords: Long-range dependence, strong dependence, global dependence, Hurst phenomena.

1 Introduction

Long memory and fractionally integrated processes were introduced into the econometrics literature by Granger and Joyeux (1980). A major problem that arises is that long memory and structural change are easily confused. As Sibbertsen (2004) states, it is often the case that structural break methods report breaks when only long memory is present; similarly, long memory estimators often report long memory when only structural breaks are present.

Banerjee and Urga (2005) notes the potential confusion of ‘stochastic regime switching and long memory’ and cites a sample of the empirical literature in this area. However, the Diebold and Inoue (2001) concluded that ‘ . . . the temptation to jump to conclusions of “structural change producing spurious inferences of long memory” should be resisted, as such conclusions are potentially naïve. Even if the “truth” is structural change, long memory may be a convenient shorthand description, which may remain very useful for tasks such as prediction. Moreover, at least in the sorts of circumstances studied in this paper, “structural change” and “long memory” are effectively different labels for the same phenomenon, in which case attempts to label one as “true” and the other as “spurious” maybe of dubious value . . . ’. In effect Diebold and Inoue (2001) seem to propose, ‘it doesn’t really matter what you call it.’ To date it has been too difficult to distinguish between long memory and structural change. There exist econometric examples which highlight long memory vs. structural breaks, for example Smith (2005), but the overwhelming impression one gets from the literature is that ‘financial data exhibit long memory with a d (around) 0.4’ and if this is “spurious” it doesn’t much matter.

The economics of why highly efficient financial markets would exhibit long range dependence is never addressed in the literature. Mikosch and Starica (2003, p456) conclude “we have tried hard to find in the literature any convincing rational/economic argument in favor of long range dependent stationary log returns, but did not find any.”

There is, however, both a finance theory and regulator interest in long versus short memory modeling implications. As Jambee and Los (2005) highlight, the US Financial Accounting Standards Board (FASB), has an interest in how to value options because the traditional Black-Scholes formula is (for some reason) “not doing its job”. Bollerslev and Mikkelson (1996) simulated call option prices for the Standard & Poor’s 500 composite index and showed that the price of a call option is much higher when long memory in the volatilities is assumed compared to when it is not.

It therefore seems important, both for aspects of finance theory, econometric modeling and empirical consistency that we try to identify whether the data are generated by “true” long memory processes.

This paper’s main contribution to the literature is that we propose a method to identify breaks and to ascertain whether they are consistent with an FI(d) process.

2 Method

To motivate the new procedures consider the differences in the breaks reported in series generated by a true long memory process and those reported when there are structural breaks in the data. A time series which is generated by a true long memory process has a uniform data generating process (DGP) the entire series. Thus if a structural break location method is mistakenly applied to the series it may report a number of breaks where no breaks exist. These spurious breaks will yield a number of “regimes” of differing lengths but these “regimes” will only be subsamples of a single population. Thus the subsamples will have the same statistical properties as the full series because subsamples have been drawn from a single population. Any estimated differences will be the result of randomness and long range serial correlation not differences between the samples.

If a time series has one or more structural breaks then the series has one or more discontinuities in the DGP. In this case a structural break method will report a number of breaks which will divide the series into regimes which are different subpopulations. The statistical properties of these subpopulations within the regimes will need to be estimated. The estimated differences will be the result of actual differences between the samples.

Thus it seems reasonable to suspect that it may be possible to distinguish between an FI(d) process and series with structural breaks by examining the statistical properties of the reported regimes. For example, for an FI(d) series we can obtain through simulation the bivariate null distribution of d estimate with regime length by breaking a large number of simulated series into spurious regimes with a structural break location procedure and estimating d for each regime d. We then apply the same procedure to a real series and examine whether the bivariate data points for the real series fitted the null distribution obtained by simulation. If there was a good fit then we would not reject the null of an FI(d) series. If the fit was poor then we would accept the alternative that the series has one or more structural breaks.

The structural break method we use is Atheoretical Regression Trees (ART) because of its computational speed. ART was recently introduced by Cappelli et al. (2008), is based on least squares regression trees and is non-parametric.

In this paper we present a graphical procedure for distinguishing between true and spurious long memory. As indicated above the null hypothesis is that the series is an FI(d). The alternative is that the series has one or more structural breaks.

2.1 Using ART to estimate the bivariate distribution of d with regime length

1. Estimate d for the full series.
2. Through simulation obtain the bivariate probability distribution of d with regime length.
   (a) Simulate a large number of FI(d) series with the same d and series length as the series under test.
   (b) Use ART to break the series into "regimes".
   (c) Estimate d within these regimes.
   (d) Calculate the empirical median, 75%, 95%, and 99% confidence intervals for the bivariate d and regime length distribution.
   (e) Plot the bivariate distribution.
3. Apply ART to the full series to obtain the candidate break points.
4. Estimate d for each regime reported by ART.
5. For each regime overplot the d estimate and regime length on the previously determined empirical bivariate distribution.
6. Assess whether d is constant. Details of how the assessment was made are given in Section (4).

In step 2a of the method we found that typically it was necessary to simulate 10,000 to 25,000 FI(d) series to obtain a good estimate of the bivariate distribution.

Unless otherwise stated all d estimates were obtained using the estimator of Haslett and Raftery (1989) as implemented by in in the `R` software (R Development Core Team, 2005) package `fracdiff` of Fraley et al. (2006). FI(d) series were simulated with the function `farimaSim` in `fSeries` of Wuertz (2005).

The Beran test was evaluated using functions implemented in `longmemo` of Beran et al. (2006).

3 The Data Set

The data set comprised the realized volatility of 16 Dow Jones Industrial Average (DJIA) index stocks and were provided by Scharth and Medeiros (2007). The 16 stocks are Alcoa (AA), American International Group (AIG), Boeing (BA), Caterpillar (CAT), General Electric (GE), Hewlett

Stock	d Est	Beran	ORT	Graphical
AA	0.42	2.65×10^{-6}	2.68	1.54×10^{-5}
AIG	0.40	0.14	4.02	0.068
BA	0.40	0.02	1.63	0.0001
CAT	0.41	0.002	5.32	6.02×10^{-6}
GE	0.44	0.04	4.98	3.32×10^{-5}
GM	0.36	0.09	5.20	0.0004
HP	0.44	0.002	0.50	0.004
IBM	0.44	8.93×10^{-6}	1.13	0.02
INTC	0.46	3.59×10^{-7}	2.20	0.03
JNJ	0.40	0.16	3.62	0.02
KO	0.42	2.57×10^{-6}	4.40	3.74×10^{-10}
MRK	0.39	0.34	3.89	0.006
MSFT	0.46	0.008	11.24*	0.02
PFE	0.42	4.43×10^{-6}	3.88	0.0001
WMT	0.42	2.95×10^{-12}	4.15	0.001
XON	0.44	0.09	1.26	0.004

Table 1: For the 16 stocks in the sample column "d Est." reports d as estimated by the Haslett and Raftery (1989) estimator. Column "Beran" reports the p-value reported by the goodness-of-fit test of Beran (1992) applied to the full series. Column "ORT" is the test statistic from the method of Ohanissian et al. (2008) a single asterisk ($*$) marks the result which was significant at the 0.05 level. Column "Graphical" reports the p-value for the graphical technique for testing for a constant d presented in this paper.

Packard (HP), IBM, Intel (INTC), Johnson and Johnson (JNJ), Coca-Cola (KO), Merck (MRK), Pfizer (PFE), Wal-Mart (WMT), and Exxon (XON). The period of analysis was from January 3, 1994 to December 31, 2003. A detailed explanation of the dataset and how the realized volatilities were calculated can be found in Scharth and Medeiros (2007). It should be noted that because all 16 are part of the DJIA they cannot be considered to be independent series.

In this paper we will present results for American International Group (AIG) and, where practical, state the results for the other 15 stocks.

4 Results

The procedure begins with the estimation of d. These results are presented in Column "d Est" of Table (1). A d estimate close to 0.4 is appropriate for most series. Realized volatilities usually present a long memory behavior and a visual examination of the series, ACF and periodogram (not presented for reasons of space) suggested all 16 series are of the long memory type.

A simple alternative model which has been well studied is the two-state Markov switching model. We compared the empirical bivariate distribution of estimated d against regime length for a two-state Markov switching model with seven switches and a FI(0.4) series.

Panel A of Figure (1) presents the empirically estimated bivariate distribution from 1000 replications of the two-state Markov switching series with eight regimes (i.e. seven switches) and 2500 data points. Panel B presents the results for the 1000 FI(0.4) series.

The bivariate distribution for the Markov switching series reflected the fact that there was no long memory in these series. In general, ART correctly located the switchpoints and hence divided the series into regimes consisting of random numbers with the same mean. The Haslett-Raftery estimator was then applied to these regimes and it correctly reported a d value of zero or close to zero. For the FI(0.4) series (Panel B of Figure 1), splitting the series up into "regimes" with ART did not mask the fact that the series had long memory.

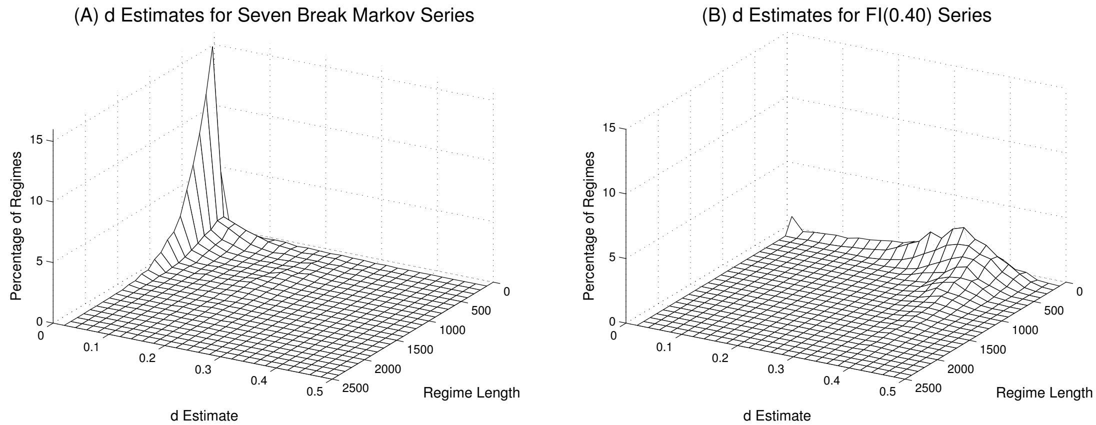

Figure 1: (A) Bivariate distribution of d estimates with regime length for a 2500 data point Markov switching series with seven switches. (B) Bivariate distribution of d estimates with regime length for FI(0.40) series.

The results for AIG are presented in Figure (2). The figure presents the empirical conditional bivariate distribution of d given regime length when the null hypothesis of FI(0.4) was known to be true. The solid and three sets of dashed lines present the empirically determined median, 75%, 95% and 99% confidence intervals. The "A" symbols represent the seven AIG data points. Visual inspection of this graph shows the 1168 trading day regime (right most "A" in Figure 2) has a statistically significant d value to the full series. We can use a binomial distribution to obtain a conservative p-value of whether the null is likely to be true.

Column "Graphical" of Table (1) presents the calculated p-values for this method for all 16 stocks. The null hypothesis is not accepted for 15 of the 16 stocks. The null hypothesis was not rejected for AIG. The calculation of the p value took no account of how far the points were outside the empirical confidence intervals. A visual inspection of Figure (2) suggests the null hypothesis would not be accepted for AIG if a more sophisticated calculation was carried out.

For comparison purposes we present the results of the Beran (1992) and Ohanissian et al. (2008)test, columns "Beran" and "ORT" of Table (1) respectively. Because the asymptotic properties of the Beran test are unknown (Deo and Chen, 2000) we subjected it to empirical testing through simulation. This testing showed the Beran goodness of fit test over rejected the null by a small amount and that a significance level of 0.04 reported by the test was approximately equal to a true level of 0.05. The Beran test rejected the null hypothesis for 11 of the 16 stocks at the 0.05 level. The Ohanissian et al. (2008) test only rejected the null of a FI(d) for the MSFT stock.

5 Discussion

In Section (2) we noted that when breaks are reported by a structural break location method in FI(d) series that they are fundamentally different from breaks reported in, say, a Markov switching series. We recall that ART finds candidate breakpoints by detecting shifts in the mean. A change in d and a shift in mean are only related if ART has located a genuine structural break in the series.

The results presented in Figure (1) showed that our intuition about the differences between the two types of breaks was indeed correct when our procedure was applied to simulated data. ART usually correctly located the switch points in the two state Markov switching series and the d estimator then correctly reported an estimate of d of zero or close to zero for the regimes. The breaks reported by ART for the FI(d) series were different. They were caused by the long excursions away from the mean exhibited by FI(d) series. All of them are spurious. Consequently the d estimates in panel B of Figure (1) reflect the fractional integration of the whole series. These

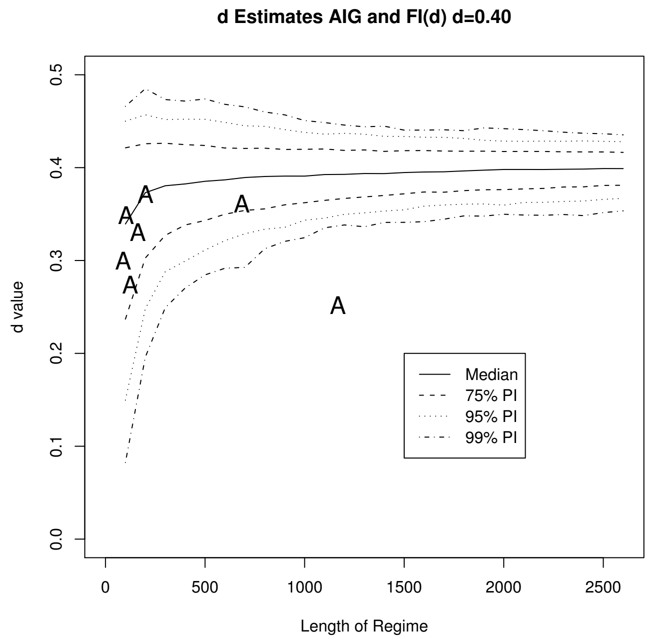

Figure 2: Conditional bivariate distribution of d given regime length. The data points for the seven AIG regimes are represented with the symbol "A".

represent two extreme cases; series with mean switches but no other serial correlation and pure FI(d) series.

Because a parameter only has meaning in the context of a model, if d is allowed to be a function of time, as demanded by the evidence, we are led to the so-called multifractal models. However, this leads to two problems. First, a time-varying d nullifies one of the great strengths of the FI(d) models (or FGNs), namely that a single parameter can model the long-range dependence properties of the whole series. Once a time-varying d is admitted, then any model for the data must of necessity be non-stationary.

Second, as far as we are aware, there are no estimators of d which can handle a non-constant d. All assume that d is constant as required by the FI(d) model. Thus is it not clear what, if any, meaning can be assigned to a d estimate for a series when d itself changes throughout the series.

We applied the tests of Beran (1992) and Ohanissian et al. (2008). These results are presented in the columns labeled "Beran" and "ORT" in Table (1). For the test of Ohanissian et al. (2008) we chose four levels of aggregation and the critical values for the χ^2 distribution with three degrees of freedom are 7.82 for $p = 0.05$ and 11.35 for $p = 0.01$. The graphical method does not accept the null for more series than either the Beran (1992) or Ohanissian et al. (2008) tests.

6 Conclusions and Future Research

It is now well established that long memory and structural change are easily confused, however, most researchers, particularly in the financial econometrics area, choose to ignore the problem, or simply find it too difficult, empirically, to distinguish between them.

The main contribution of the paper was to propose a new approach that can quickly locate potential breaks in the series and test whether the breaks identified are consistent with an FI(d) process. This is a fundamentally different way to approach the problem of distinguishing between long memory and structural change and has not previously been proposed. The particular structural break-test method used here, based upon Atheoretical Regression Trees, is fast enough to be practical with the large sized datasets typical in the financial econometrics area. The new approach was then applied to 16 financial data series to examine whether mean shifts and long memory were linked in realized volatility series. The data were also examined using some existing tests including that of Beran (1992) and Ohanissian et al. (2008).

References

Banerjee, A. and Urga, G. (2005). Modelling structural breaks, long memory and stock market volatility: an overview. *Journal of Econometrics*, 129:1–34.

Beran, J. (1992). A Goodness-of-Fit Test for Time Series with Long Range Dependence. *Journal of the Royal Statistical Society B*, 54(3):749–760.

Beran, J., Whitcher, B., and Maechler, M. (2006). *longmemo: Statistics for Long-Memory Processes (Jan Beran) – Data and Functions*. R package version 0.9-3.

Bollerslev, T. and Mikkelson, H. (1996). Modeling and pricing long memory in stock marker volatility. *Journal of Econometrics*, 73:151–184.

Cappelli, C., Penny, R. N., Rea, W. S., and Reale, M. (2008). Detecting multiple mean breaks at unknown points with Atheoretical Regression Trees. *Mathematics and Computers in Simulation*, 78(2-3):351–356.

Deo, R. S. and Chen, W. W. (2000). On the integral of the squared periodogram. *Stochastic Processes and their Applications*, 85:159–176.

Diebold, F. X. and Inoue, A. (2001). Long Memory and Regime Switching. *Journal of Econometrics*, 105:131–159.

Fraley, C., Leisch, F., Maechler, M., Reisen, V., and Lemonte, A. (2006). *fracdiff: Fractionally differenced ARIMA aka ARFIMA(p,d,q) models*. R package version 1.3-0.

Granger, C. W. J. and Joyeux, R. (1980). An Introduction to Long-range Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, 1:15–30.

Haslett, J. and Raftery, A. E. (1989). Space-time Modelling with Long-memory Dependence: Assessing Ireland's Wind Power Resource (with Discussion). *Applied Statistics*, 38(1):1–50.

Jambee, S. and Los, C. (2005). Long-memory options: Valuation. mimeo.

Mikosch, T. and Starica, C. (2003). Long-Range Dependence Effects and ARCH. In Doukhan, P., Oppenheim, G., and Taqqu, M., editors, *Theory and Applications of Long-Range Dependence*, pages 439–459. Birkhaüser.

Ohanissian, A., Russell, J. R., and Tsay, R. S. (2008). True or Spurious Long-Memory? A New Test. *Journal of Business and Economic Statistics*, 26(2):161–175.

R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Scharth, M. and Medeiros, M. C. (2007). Asymmetric Effects and Long Memory in the Volitility of Dow Jones Stocks. working paper.

Sibbertsen, P. (2004). Long memory versus structural breaks: An overview. *Statistical Papers*, 45(4):465–515.

Smith, A. D. (2005). Level Shifts and the Illusion of Long Memory in Economic Time Series. *Journal of Business and Economic Statistics*, 23(3):321–335.

Wuertz, D. (2005). *fSeries: Financial Software Collection*. R package version 220.10063.