Empir Econ
DOI 10.1007/s00181-010-0338-y

ORIGINAL PAPER

Are U.S. stock prices mean reverting? Some new tests using fractional integration models with overlapping data and structural breaks

Steven P. Clark · T. Daniel Coggin

Received: 3 May 2009 / Accepted: 24 November 2009

Abstract Using recently developed econometric models of fractional integration with overlapping data, this study examines the time series properties of real monthly U.S. stock returns over the period 1871–2003. Using 1-month and overlapping, long-horizon stock returns of 12, 24, and 36 months, we find that real U.S. stock returns are covariance stationary for this period before and after allowing for the presence of structural breaks. Our results imply that the permanent (random walk) component of stock prices *overwhelms* any temporary (mean reverting) component, producing a fractional d-value for returns indistinguishable from zero. We highlight the limitations of standard ACF models of overlapping returns, and suggest that the previously observed pattern of increasingly negative autocorrelations is largely an artifact of short-term ARMA dynamics. We confirm the result of Souza (J Time Ser Anal 28:701–722, 2007) that, holding the bandwidth constant, overlapping (and nonoverlapping) temporal aggregation should not affect semiparametric, frequency domain d-estimates such as the GPH and feasible exact local Whittle.

Keywords Fractional integration · Long-horizon stock returns · Mean reversion · Overlapping data · Structural breaks · Temporal aggregation

JEL Classification C14 · C32 · G10

S. P. Clark
Department of Finance, University of North Carolina at Charlotte, Charlotte, NC, USA
e-mail: spclark@uncc.edu

T. D. Coggin (✉)
Horizon Investments, LLC, Charlotte, NC, USA
e-mail: dcoggin@bellsouth.net

Published online: 07 February 2010

> We are impressed by the inexorable tendency for reversion to the mean in security returns. (Bogle and Malkiel 2006, p. A14).

1 Introduction

It is now widely assumed in the academic and practitioner literature that stock returns are mean reverting. In this context, it is important to distinguish between stock *prices* and stock *returns*, the latter of which (ignoring dividends) are defined as first-differences of the logs of stock prices. Specifically, it is argued that stock prices are the sum of two components: a permanent (possibly random walk) component influenced by fundamental factors and a temporary (mean reverting) component that tends to induce negative autocorrelation, typically concentrated in return horizons beyond 1 year. As noted by Summers (1986) and Fama and French (1988) and others, the mean reversion hypothesis begins by assuming that *stock prices are nonstationary I(1)*, implying that *stock returns are stationary I(0)*. The key question is then to test whether *returns* are (weakly) autocorrelated or not.

The mean reversion hypothesis is derived from the temporary (mean reverting) component of stock prices, which induces *predictability* (negative autocorrelation) in stock returns. Specifically, the permanent (random walk) component produces white noise in stock returns, and the temporary (stationary) component causes negative autocorrelation slowly converging to zero. The mean reverting model for stock prices thus implies that the volatility of returns *decreases* with the time horizon, as returns "revert" to a constant mean. That is, a special form of mean reversion is implied. As noted by Kim et al. (1991, p. 515):

> Thus, stock returns seem to be characterized by positive autocorrelation over intervals under a year and by negative autocorrelation over longer intervals. The latter finding has been interpreted as evidence of 'mean reverting' behavior in stock prices. Evidently, a given change in price tends to be reversed over the next several years by a predictable change in the opposite direction.

It is therefore important to note that in this model the *test* of mean reversion in stock prices centers on the predictability (negative autocorrelation) of the temporary component of stock returns for long-horizon returns (Fama and French 1988, p. 249). That is, the test of mean reversion in stock *prices* in the Fama and French model is a test for significant negative autocorrelation in stock *returns* at long horizons.

While not as popular, a competing model in financial economics is the *mean averting* model. In contrast to the mean reverting model, the mean averting model implies that volatility *increases* (and autocorrelation gets more *positive*) with increases in the time horizon as returns wander *away* from a constant mean. This model for stock returns was proposed by Lo and MacKinlay (1988) and Kim et al. (1991, 1998).

The mean reverting model for stock prices was first made popular by the research of Summers (1986), Fama and French (1988), Poterba and Summers (1988), and DeBondt and Thaler (1989); and later extended by Campbell and Shiller (1998) and Daniel (2001). Caporale and Gil-Alana (2002) provided an analysis using fractional differencing methodology. A representative sample of articles broadly challenging the mean reversion hypothesis include Lo and MacKinlay (1988), Cecchetti et al. (1990), Lo (1991), Kim et al. (1991), Kim et al. (1998), Kim and Nelson (1998), Richardson (1993), Chambers (1998), Lobato and Savin (1998), and Perron and Vodounou (2004).

In this study we will reexamine the mean reversion hypothesis for stock prices. In the process of doing so we discuss and illustrate the use of some recently developed econometric models of fractional integration with overlapping data specifically suited to this purpose. Since fractional integration, temporal aggregation, and overlapping data are frequently and increasingly used in financial research, we suggest that our results will also be useful to others using these methodologies.

The remainder of our study is organized as follows. Section 2 develops the relationship between mean reversion, fractional integration, temporal aggregation, and overlapping data; and it sets the stage for our specific examination of overlapping stock returns. Section 3 describes the S&P 500 stock return data. Section 4 details the tests of fractional integration and their implementation. Section 5 presents our results. Section 6 summarizes and concludes our study.

2 Mean Reversion, fractional integration, and overlapping data

2.1 Three definitions of mean reversion

In spite of its important implications and popularity as a topic of research and debate, there is a surprising lack of unanimity about precisely what is meant by mean reversion in stock prices or, for that matter, by mean reversion in time series generally. This circumstance is especially apparent in the literature on fractional integration when describing the behavior of $I(d)$ series with $d \in [1/2, 1)$. The reasons for this can perhaps best be understood in the context of two models of fractional integration, Type I and Type II. Each of them implies a slightly different, but important, definition of mean reversion as we will explain below.

The difference between Type I and Type II models of fractional integration is essentially in how the fractional differencing operator $(1-L)^d$ is defined. A detailed description and comparison is given by Marinucci and Robinson (1999). A univariate time series y_t is said to be integrated of order $d(y_t \sim I(d))$, for any real number d, if y_t admits a representation:

$$(1-L)^d y_t = u_t 1\{t \geq 1\}, \quad t = 0, \pm 1, \ldots \tag{1}$$

where $1\{\cdot\}$ is the indicator function, L is the lag operator, u_t is a covariance stationary process with $E[u_t] = 0$ for all t, and the spectral density of u_t is positive and continuous at the zero frequency. More explicitly, by expanding $(1-L)^d$, Eq. 1 can be rewritten as:

$$\sum_{k=0}^{t-1} A_k y_{t-k} = u_t, \tag{2}$$

with $A_0 = 1$ and, for $k \geq 1$, $A_k = ((k-d-1)/k)A_{k-1}$. If $y_t \sim I(d)$, for a noninteger d, then y_t is said to be fractionally integrated. The series y_t is covariance stationary if $d < 1/2$, and nonstationary if $d \geq 1/2$. This model of fractional integration is called Type II by Shimotsu and Phillips (2006), and it is also discussed in Robinson (1994) and Tanaka (1999). Note that in Type II fractional integration, the process is initialized at time $t = 0$ with all past innovations set to zero and the fractional differencing operator expressed as a finite order moving average. An advantage of the Type II model is that it is valid for all real number values of d.

In the Type I definition, the fractional differencing operator is inverted and represented as an infinite order moving average:

$$y_t = (1-L)^{-d}u_t = \sum_{k=0}^{\infty} B_k u_{t-k}, \tag{3}$$

with $B_0 = 1$ and, for $k \geq 1$, $B_k = ((d+k-1)/k)B_{k-1}$. Such a representation is possible provided that $d < 1/2$.

In the context of fractional integration, there is considerable disagreement in the literature about the definition of *mean reversion*. For example, Robinson (2003, p. 20) writes, "The region d in [1/2, 1) is referred to as mean reverting, MA coefficients of x_t decaying, albeit more slowly than under stationarity, $d < 1/2$." Thus Robinson suggests that if shocks *eventually dissipate*, then the resulting behavior of the series should be called mean reverting *regardless* of whether the data are stationary. This is consistent with the Type I model of mean reversion and implies a specific model of mean reversion, *decaying moving average coefficients associated with the lags of* y_t.

Phillips and Xiao (1999, p. 34) argue *against* this characterization of mean reversion writing, "A process with $d \geq 0.5$ has nonstationary long memory and a variance that explodes as t goes to infinity. Such processes are, in fact, not mean reverting, although their impulse responses…decay to zero provided $d < 1$, and so shocks are not persistent in this case." Thus Phillips and Xiao argue for a definition of mean reversion as a time series property that arises *only* in the case of stationarity. This is consistent with the Type II model and implies a second definition of mean reversion, *covariance stationarity*.

It should also be noted that in the Type I definition of mean reversion described in Robinson (2003) above, due to the lack of ergodicity in nonstationary series (i.e., $1/2 \leq d < 1$), the sample mean has *no* asymptotic limit. Therefore, technically speaking, *there is no mean value to which the series reverts.* See Phillips (2001) for a discussion and descriptive analysis of the limiting behavior of the sample moments of nonstationary time series; and see Davidson and Hashimzade (2009) for an expanded discussion of Type I and Type II fractional (long memory) processes.

A third definition of mean reversion associated with autocorrelation coefficients was *specifically developed for stock prices* by Fama and French (1988). As noted in the Introduction, the mean reverting model is derived from the temporary (mean reverting) component of stock *prices* that induces the negative (*and significant*) autocorrelation in *returns*. Fama and French (1988) and subsequent articles used a regression-based testing methodology to show that stock returns are significantly negatively autocorrelated at long horizons. Specifically, in their model the slowly decaying temporary (stationary) component in prices over long time horizons induces the negative autocorrelation in returns, resulting in a ∪-shaped pattern for autocorrelations over time.

2.2 Fractional integration

Starting with the path-breaking work of Dickey and Fuller (1979) and Said and Dickey (1984), attention and subsequent formal statistical analyses in the mainstream time series econometrics literature have focused on $I(1)$ unit root time series. That is, time series that need to be first-differenced to become $I(0)$ stationary. At this point, as discussed in Caporale and Gil-Alana (2002), it is important to note that $I(0)$ and stationarity are *not* identical concepts. Specifically, $I(0)$ implies covariance stationarity, but covariance stationarity does not strictly imply $I(0)$. As noted above, *a fractionally integrated $I(d)$ process is also covariance stationary if $d < 1/2$*. For excellent comprehensive summaries of this literature, see Stock (1994) and Maddala and Kim (1999). The emphasis on unit roots was due to the belief that many economic time series follow a random walk, a subcategory of unit root time series.

However, Stock (1991) and DeJong et al. (1992) formally showed that the unit root test suffers from a "knife's edge" problem. That is, the classic DF unit root test pits a null of a nonstationary process where the first-order autoregressive root is *exactly* 1.0, against the alternative that the first-order autoregressive root is less than one in absolute value. But, as shown in a number of previous studies, the classic DF-ADF unit root test is not consistent against *fractional* alternatives, except in special cases.[1] Fractional integration models allow one to test for the presence of noninteger orders of integration *greater or less than 1.0*. This in turn opened the door to an explicit recognition and focus on the development of fractionally integrated time series models. See Baillie (1996), Gil-Alana and Robinson (1997), Henry and Zaffaroni (2003), and Gil-Alana and Hualde (2009) for detailed summaries of applications in finance and economics.

In this study we will focus on the well-known GPH (Geweke and Porter-Hudak 1983) method for estimating the fractional integrating coefficient, which we will discuss below. Even if one chooses a statistical test for fractional integration, the issue of structural breaks should also be considered before one can begin to have confidence in the results of the test. For example, Diebold and Inoue (2001) and Granger and Hyung (2004), among others show that structural breaks and $I(d)$ models can be easily confused. Thus we will also include a model of fractional integration that explicitly allows structural breaks.

2.3 Temporal aggregation, bias reduction, and the Souza result

It is now well established that, in theory, semiparametric (frequency domain) estimators of the fractional differencing parameter (d) are unbiased by so-called "short-term ARMA dynamics." This results from the fact that these d-estimates focus on the behavior of the spectrum at or near the zero frequency. That is, long memory is a low

[1] See for example Diebold and Rudebusch (1991), Hassler and Wolters (1994), Lee and Schmidt (1996), and Kramer (1998).

frequency phenomenon, while short memory manifests itself at higher frequencies. See Boudoukh et al. (1994) and Robinson (2003) for detailed discussions.

It is equally well known that this theoretical result is not so easily realized in practice. See Agiakloglou et al. (1993) for a discussion with regard to the semiparametric, log-periodogram GPH regression model. This is due to the fact that in order to estimate the semiparametric d, one must choose a bandwidth parameter $m = n^{\exp}$ (where m is the frequency or number of periodogram ordinates used, n is the total number of time series observations, and exp is the exponent used to calculate m). In choosing a value for m, the classic "bias versus variance" trade-off appears. Specifically, the smaller the value for m, the smaller the bias due to short-term effects, but the greater the sampling error in the estimate of d. See Smith et al. (1997) and Souza and Smith (2002) for a discussion. Several methods have been suggested for choosing the "optimal" bandwidth that minimizes the bias versus variance trade-off. So far, none of them are without problems and limitations. See Moulines and Soulier (2003) for a discussion.

One method of reducing the bias due to short-term ARMA dynamics is to *temporally aggregate* the data. Temporal aggregation of time series data has a long history in applied statistics and econometrics dating at least from the seminal articles of Amemiya and Wu (1972), Robinson (1978), and Granger (1980), and more recently Marcellino (1999). For an excellent review of this literature with respect to univariate ARIMA-GARCH models, see Silvestrini and Veredas (2008). There are two types of temporal aggregation: *stock* and *flow*. Stock variables are systematically sampled processes aggregated by "skip-sampling." Flow variables are temporally aggregated by summing. This is the typical form of aggregation used in financial economics. In flow aggregation, a time series of length n is divided into n/k consecutive segments, where k is a positive integer defining the desired number of segments. Each segment is then further defined as the sum of the observations in that segment. Souza and Smith (2004) show that temporal aggregation of flow variables can reduce the bias in the d-estimate caused by short-term dynamics, while increasing its standard error due to the shortening of the series by aggregation.

A number of other studies have shown that the degree of temporal aggregation does not change the value of d. See, e.g., Chambers (1998) and Smith (2005); Souza (2007). This result suggests that one can estimate the true long memory parameter of any temporally aggregated process by focusing on frequencies sufficiently close to zero. However, in his study of quarterly and annual UK macroeconomic data, Chambers (1998) shows that this theoretical result can become problematic in practice.

In a recent study, Souza (2007) builds on the findings in Chambers (1998) and further investigates the relationship between temporal aggregation and bandwidth selection. Souza shows that, under conditions defined in Robinson (1995a,b) and Hurvich et al. (1998) plus some mild additional conditions, the estimate of d derived from a temporally aggregated sample is asymptotically equivalent to the estimate obtained from the original sample with the same bandwidth. Hence, in Souza's words, "… there is no need to aggregate the series just to diminish the [short-term] bias, it is enough to use a narrower bandwidth in the estimation" (Souza 2007, p. 716). Specifically, Souza (2004, 2007) shows that, for positive values of d, as the sample size increases the bias tends to disappear for both stock and flow variables as a smaller band of frequencies is used for estimation. He then shows that this result is valid for the GPH estimator and for the Gaussian semiparametric estimator of Robinson (1995b). Specifically Souza shows that if a time series is stationary, then its temporally aggregated sums should also be stationary, holding *constant* the bandwidth (i.e., number of harmonic frequencies). We shall call this the *Souza result*.

This basic result can now be briefly summarized. Given some relatively mild assumptions, if one calculates the (frequency domain) d-value for a time series variable, then the d-estimate for temporal aggregations of that variable *using the same* m*-value* should be equivalent.[2] Hence, while one may want to use temporal aggregation for some *other* reason (noted below), it is *not* necessary in order to calculate d-values. In fact as we will show, due precisely to the effects of short-term dynamics, the d-value for temporal aggregations may be spuriously inflated. At this point, it is important to make clear that the Souza result is derived for a constant bandwidth (m), *not* a constant exponent used to calculate the bandwidth.

2.4 Overlapping data

One of the most popular forms of temporal aggregation in financial economics is *overlapping data*. Among the first published studies in financial economics to use overlapping observations is the article by Hansen and Hodrick (1980). In their landmark study, Hansen and Hodrick used overlapping data to study the market efficiency of foreign exchange rates. Their rationale for using overlapping data was twofold: one, to increase the nominal sample size and two, to deal with the fact that the sampling interval for spot rates occurred more frequently than the interval for forward rates. That is, the forecasting interval was longer than the sampling interval. Hence the use of overlapping data allows a forecast model to "update" its forecasts more frequently.

Another reason sometimes given for the use of overlapping data is to obtain greater statistical efficiency [i.e., lower error variance, see Muller (1993)]. Richardson and Smith (1991) demonstrated the increase in the power of statistical tests due to the use of overlapping observations. In a study focusing on financial economics, Harri and Brorsen (2002) found that a surprising 34% (weighted-average) of empirical articles published in just three journals they examined for the year 1996 (*The Journal of Finance*, *The American Economic Review*, and *The Journal of Futures Markets*) used overlapping data. With regard to the topic of this study, a number of previous studies have used overlapping ("rolling period") data to study the existence of long memory in stock returns, including Fama and French (1988), Lo and MacKinlay (1988), Lo (1991), and Coggin (1998).

3 Data

The common stock return data used in this study are taken from the monthly S&P 500 stock index total returns (price change plus dividend) data file maintained by

[2] As a Reviewer noted, it would be helpful if Souza provided a statistical test for the equivalence of the temporally aggregated and nonaggregated d-estimates. To our knowledge, such a formal test has not been derived for both overlapping and nonoverlapping data. See Ohanissian et al. (2008) for an example using nonoverlapping data.

Professor Robert Shiller of Yale University. The S&P 500 is a market capitalization-weighted, large capitalization stock index currently maintained by the Standard & Poors Corporation. The monthly returns are expressed in December 2003 dollars using the CPI-U (Consumer Price Index-All Urban Consumers) published by the U.S. Bureau of Labor Statistics, converted to natural logs to get continuously compounded real returns, and include the period January 1871 through December 2003 ($n = 1{,}596$ months).

To our knowledge, the S&P 500 data maintained by Professor Shiller is the longest continuous data file of U.S. stock returns in existence. Shiller's original data and the complete methodology for maintaining the data file are available at his web site: http://www.econ.yale.edu/~shiller.[3]

As noted above, our primary data are monthly total returns which we convert to continuously compounded real returns. Consistent with a number of previous studies that use monthly data [see, e.g., Fama and French (1988), Lo and MacKinlay (1988), and Lo (1991)], we also aggregate the monthly returns into "rolling period returns." By this we mean overlapping, continuously compounded returns. The overlapping periods chosen are: 1 year (12 months), 2 years (24 months), and 3 years (36 months). Previous studies have also suggested a structural break in the pattern of mean reversion in U.S. stock returns at or near the end of World War II. Fama and French (1988) noted evidence of a break, but suggested that it was largely due to sampling error. Kim et al. (1991) and Kim and Nelson (1998) also found a structural break at the end of 1945, after which the evidence for mean reversion disappeared. For this reason, in addition to examining the full series, we also include a test for spurious long memory due to the presence of structural breaks.

4 Methodology

4.1 The GPH method

Shortly after the initial work by Dickey and Fuller (hereafter denoted DF) on unit root models, Granger and Joyeux (1980) and Hosking (1981) suggested the potential usefulness of *fractional* values of d in Eq. 1. Geweke and Porter-Hudak (1983) then developed a semiparametric frequency domain, log-periodogram regression model for estimating d commonly referred to as the GPH method. As noted above, fractional integration models allow one to test for the presence of noninteger orders of integration greater or less than 1.0. The statistical methodology used in this study involves models of fractional integration. Specifically, we will use the well-known semiparametric GPH method of estimating the fractional d-value.

[3] As more fully explained by Shiller on his web site, the Standard and Poor's Monthly Composite Stock Index was extended back from year-end 1925 to 1871 as the Cowles Commission Common Stock Index developed by Alfred Cowles and Associates. Dividend and earnings data before 1926 are also from Alfred Cowles and Associates, interpolated from annual data. Starting with 1926, all these data are available from the Standard and Poor's Statistical Service in daily and monthly formats. The CPI-U published by the U.S. Bureau of Labor Statistics begins in 1913. For years before 1913, Shiller spliced to the CPI Warren and Pearson's price index, by multiplying it by the ratio of the indexes in January 1913.

If we let $I(\omega_j)$ denote the sample periodogram at the jth Fourier frequency, evaluated at $\omega_j = 2\pi j/n$, $j = 1, 2, \ldots, m$ (the bandwidth parameter), $m \prec n$, and n = total number of observations; then the original Geweke and Porter-Hudak (1983) GPH estimator of d is based on the OLS regression of the log-periodogram on the log frequency:

$$\log[I(\omega_j)] = \beta_0 + \beta_1 \log[4 \sin^2(\omega_j/2)] + \varepsilon_j, \tag{4}$$

where $d_{\text{GPH}} = -\beta_1$. Robinson (1995a,b) shows that the GPH estimate is consistent and asymptotically normally distributed for the covariance stationary range $-1/2 < d < 1/2$. We chose the GPH method due to its widespread availability, ease of calculation and the fact that it is one of the methods used in Souza (2007). It is also the method used in Smith (2005), which we will discuss below in the context of structural breaks. We note that there are several recently developed methods of estimating the fractional difference parameter that have more desirable statistical properties [see, e.g., Robinson (1994), Shimotsu and Phillips (2005), and Shimotsu (2009) and the summary in Doukhan et al. (2003)]. We will show below that our main results are not sensitive to our choice of the GPH method.

4.2 Structural change and "spurious" long memory

Some recent efforts to model the long memory (fractionally integrated) process in the presence of structural breaks include Gil-Alana (2004), Gil-Alana (2008), Smith (2005), Shimotsu (2006), and Ohanissian et al. (2008). A recent comprehensive summary of this literature is provided by Banerjee and Urga (2005). We chose the model of Smith (2005) to test for the existence of spurious long memory in our stock returns data. Smith presents a GPH methodology that is valid in the presence of slowly varying level shifts.

4.3 Short-term dynamics and bandwidth

Since our use of overlapping data creates significant short-term dependence in our data, we *initially* follow the suggestion of Geweke and Porter-Hudak (1983) and use $m = n^{0.50}$ as the bandwidth in our GPH log-periodogram regression model. In addition, Robinson and Henry (1999) have shown that smaller values for m act to diminish the effect of non-normality on the log-periodogram regression errors.

5 Results

5.1 Unit root tests

As a preliminary statistical test, we applied the ADF unit root test of Dickey and Fuller (1979), the reverse-regression $\text{ADF}_{\max}$ test of Leybourne (1995), and the min-LM unit root test of Lee and Strazicich (2003, 2004) that allows a single endogenous structural break under both the null and alternative unit root hypothesis. We use their Model A,

Table 1 Summary of unit root tests

Returns	n	nlag	DF-ADF t-test	$\text{ADF}_{\max}$ t-test	LS-ADF min LM-test
One-month	1,596	4	−15.87	−15.84	−15.00
12-month	1,585	11	−11.79	−11.79	−11.93
24-month	1,573	10	−8.46	−8.46	−8.23
36-month	1,561	9	−5.85	−5.85	−5.50

All test regressions include a constant term. n = number of months and nlag = number of lags chosen by AIC criterion. $\text{ADF}_{\max}$ is the unit root test of Leybourne (1995), and LS-ADF is the min LM-test of Lee and Strazicich (2003, 2004). All tests reject the unit root null at the 0.05 level or less

which allows a single break in the intercept of the test regression. The $\text{ADF}_{\max}$ test uses the maximum ADF t-test value from the standard (forward) and reverse ADF test regressions (where "reverse" means the time series values are temporally reversed). It is well known that the classic Dickey–Fuller test suffers from low power, and a number of more recent tests have been proposed to address this problem. Leybourne et al. (2005) compared the test of Leybourne (1995) to a number of more recent tests and found that it has relatively superior size and power characteristics. The results are presented in Table 1. We see that in each case the unit root null is clearly rejected at the 0.05 level or less.

5.2 GPH fractional d-estimates

Table 2 presents the results of the GPH method estimates of d along with the asymptotic standard errors, using $m = n^{0.35}$ to $m = n^{0.90}$ in increments of 0.05, for the 1-, 12-, and 36-month returns data. We also calculated the log-periodogram regression standard error (SE) of the GPH estimates of d using the asymptotic variance $(\pi^2/6)$derived by Geweke and Porter-Hudak (1983). We see in Table 2 that the results using $m = n^{0.50}$ as recommended by Geweke and Porter-Hudak (highlighted in bold) show in each case that the point estimate of d is less than 0.50. The 95% confidence intervals include zero and do not equal or exceed 0.50, except in the case of the 36-month returns where the 95% lower/upper bounds are 0.148/0.608. Hence our results broadly suggest that real 1-month U.S. stock returns are covariance stationary, as are the 12-, 24- and 36-month overlapping compound sums using $m = n^{0.50}$. The results in Table 2 also show that the short-term bias due to ARMA effects steadily increases as we increase the value of the m-exponent *and* the level of aggregation. This *generally* supports the use of $m = n^{0.50}$, but *also* illustrates the problem of increasing (short-term) ARMA bias as the level of aggregation increases. That is, Table 2 also suggests that the GPH estimates may spuriously increase with m (the frequency) *and* the level of overlapping temporal aggregation.[4] All this may require the use of an even *smaller* m-value. We will further discuss this complication below.

[4] An Appendix is available from the authors that recalculates Table 2 using the FELW method of Shimotsu (2009) and shows that the basic results are not affected by substituting FELW for the GPH method.

Table 2 Summary of GPH d-estimates

($n=1{,}596$) m-exponent	Frequency	1-month GPH	Asymptotic SE	($n=1{,}585$) m-exponent	Frequency	12-month GPH	Asymptotic SE
0.35	13	−0.220	0.2419	0.35	13	−0.292	0.2419
0.40	19	−0.284	0.1872	0.40	19	−0.300	0.1872
0.45	28	−0.222	0.1462	0.45	28	−0.221	0.1462
0.50	**40**	**−0.206**	**0.1176**	**0.50**	**40**	**−0.160**	**0.1176**
0.55	58	−0.163	0.0946	0.55	58	−0.049	0.0946
0.60	84	0.057	0.0766	0.60	83	0.203	0.0771
0.65	121	0.067	0.0627	0.65	120	0.511	0.0629
0.70	175	0.072	0.0514	0.70	174	0.907	0.0516
0.75	253	0.048	0.0424	0.75	251	0.803	0.0426
0.80	365	0.043	0.0354	0.80	363	0.863	0.0355
0.85	528	0.100	0.0298	0.85	525	1.004	0.0299
0.90	763	0.163	0.0259	0.90	759	1.081	0.0260
($n=1{,}573$) m-exponent	**Frequency**	**24-month GPH**	**Asymptotic SE**	**($n=1{,}561$) m-exponent**	**Frequency**	**36-month GPH**	**Asymptotic SE**
0.35	13	−0.240	0.2419	0.35	13	−0.142	0.2419
0.40	19	−0.284	0.1872	0.40	19	−0.168	0.1872
0.45	27	−0.218	0.1495	0.45	27	0.010	0.1495
0.50	**40**	**−0.151**	**0.1176**	**0.50**	**40**	**0.378**	**0.1176**
0.55	57	−0.085	0.0955	0.55	57	0.703	0.0955
0.60	83	0.236	0.0771	0.60	82	0.814	0.0777
0.65	120	0.524	0.0630	0.65	119	0.906	0.0632
0.70	173	0.773	0.0517	0.70	172	0.984	0.0519
0.75	250	0.762	0.0427	0.75	248	0.975	0.0429
0.80	361	0.851	0.0356	0.80	359	1.005	0.0357
0.85	521	0.985	0.0300	0.85	518	1.051	0.0301
0.90	753	1.070	0.0260	0.90	748	1.098	0.0261

5.3 Testing the Souza result

As shown by Chambers (1998) and further developed in Souza (2007), if a time series is stationary, then its temporally aggregated sums should also be stationary, holding constant the bandwidth (i.e., number of harmonic frequencies). A test of this proposal, which we have termed the Souza result, is contained in Table 3.

As noted above there have been some attempts to calculate the optimal bandwidth parameter m, though none of them are entirely satisfactory. To our knowledge, there have been no attempts to formally calculate optimal bandwidth consistent with the Souza result. As discussed above, the basic rule is just to use the *same* bandwidth for

Table 3 Confirming the Souza result

n	Temporal aggregation	Frequency	GPH	FELW
1596	One-month	22	−0.2088 (−0.5424, 0.1248)	−0.2113 (−0.4202, −0.0024)
1585	12-month	22	−0.2460 (−0.5796, 0.0876)	−0.2117 (−0.4206, −0.0028)
1573	24-month	22	−0.2136 (−0.5472, 0.1200)	−0.1624 (−0.3713, 0.0465)
1561	36-month	22	−0.1034 (−0.4370, 0.2302)	−0.0568 (−0.2657, 0.1521)

Note: 95% CI in parentheses. *GPH* is the Geweke Porter-Hudak estimate, and *FELW* is the feasible exact local Whittle estimate of Shimotsu (2009). The standard error of the GPH estimate is the log-periodogram regression SE, and the standard error of the FELW estimate is $(1\sqrt{4}\,\mathrm{m})$ or 0.1066 here

each temporally aggregated d-estimate. Using trial and error, in Table 3 we show that a constant frequency of $m = 22$ (i.e., $1596^{0.41911}$) does a good job of confirming the Souza result (to within sampling error) in our data.

Since, as we have noted, the GPH estimate of the fractional differencing parameter is not the most efficient and has been updated with more recent semiparametric frequency domain estimates, we also present the results using the feasible exact local Whittle estimate (FELW) of Shimotsu (2009) in Table 3. The FELW estimator is a modification of the exact local Whittle estimator of Shimotsu and Phillips (2005) to allow an unknown mean and polynomial time trend. It is consistent for $d \succ -\frac{1}{2}$, and has a $N(0, \frac{1}{4})$ limit distribution for $d \in (-\frac{1}{2}, 2)$ [$d \in (-\frac{1}{2}, \frac{7}{4})$ with a polynomial trend]. The derivation of this model is rather complex and we refer the reader to Shimotsu (2009) for complete details. The FELW method has both a larger range of consistent and normally distributed values and a smaller standard error than the GPH method. Table 3 shows that the FELW estimates basically confirm the GPH estimates, and *both* confirm the Souza result in our data.

5.4 The Smith GPH estimate for spurious long memory

Smith (2005) presents a GPH estimate of long memory that allows slowly varying level shifts. Specifically, Smith shows that the standard GPH estimate is biased in the presence of level shifts. He then shows how adding an additional regressor to the GPH log-periodogram regression in Eq. 4, $-\log(p^2 + \omega_j^2)$ [where, keeping our notation consistent with Eq. 4, p is estimated as $p = kj/n$ for some constant $k \succ 0$], would reduce the bias caused by level shifts. Smith recommends setting $k = 3$. He then shows that this specification of the GPH regression model significantly reduces the bias, although with somewhat *less* precision. Smith argues that the loss in precision is offset by the reduction in bias. He calls this new model the *Modified GPH* estimator. Table 4 presents the results of estimating the Modified GPH using $m = 22$. In Table 4 we see that, while the Modified GPH estimates are uniformly *higher* than the GPH and FELW estimates in Table 3 (and have a larger standard error as expected), the 95% confidence intervals *still* include zero suggesting $I(0)$ stationarity.

Table 4 The modified GPH estimates with slowly varying level shifts

n	Temporal aggregation	Frequency	Modified GPH
1,596	One-month	22	−0.4310 (−1.2248, 0.3628)
1,585	12-month	22	−0.4834 (−1.2772, 0.3104)
1,573	24-month	22	−0.4058 (−1.1996, 0.3880)
1,561	36-month	22	−0.2894 (−1.0832, 0.5044)

95% CI in parentheses. Modified GPH is the Smith (2005) estimate

Table 5 Structural break at end of 1945

n	Time Period	Temporal aggregation	Frequency	GPH	FELW
900	1/1871–12/1945	One-month	22	−0.1969 (−0.5307, 0.1369)	−0.3083 (−0.5172, −0.0994)
696	1/1946–12/2003	One-month	22	0.1927 (−0.1411, 0.5265)	0.0680 (−0.1409, 0.2769)

95% CI in parentheses. *GPH* is the Geweke Porter-Hudak estimate, and *FELW* is the feasible exact local Whittle estimate of Shimotsu (2009)

We noted earlier that previous research has suggested and investigated the existence of a structural break in the time series of U.S. stock returns at the end of 1945. For this reason, we also looked at the results of such a break in our data. Table 5 presents the results of the GPH and FELW fractional d estimates for the time periods 1871–1945 and 1946–2003. In order to make the results consistent with the results in Table 3 which confirms the Souza result, we used the *same* frequency from Table 3 ($m = 22$). Table 5 shows that the 95% confidence intervals for 1/1871–12/1945 tend slightly to the negative, but still contain zero (GPH) or a value close to zero (FELW). For the 1/1946–12/2003 period, both the GPH and FELW estimates are positive and the confidence intervals include zero.

5.5 Mean reversion and negative autocorrelation

Fama and French (1988) and subsequent articles use a regression-based methodology to show that real long-horizon stock returns are negatively autocorrelated. Our results suggest a *different* interpretation. Figure 1 presents the autocorrelation function (ACF) for the 1-, 12-, 24-, and 36-month returns. We have shown (in Table 3) that the point estimate for the fractional differencing parameter for monthly real U.S. stock returns is approximately -0.21 with a 95% confidence interval that includes zero. In Fig. 1 we see that the ACF for 1-month returns displays the classic $d = 0.0$, stationary pattern of rapid decay to zero. The ACFs for the overlapping temporal aggregations are slightly more nuanced. That is, as we move to the 12-, 24- and 36-month returns, we can clearly see the impact of short-term ARMA dynamics on ACFs induced by overlapping data.

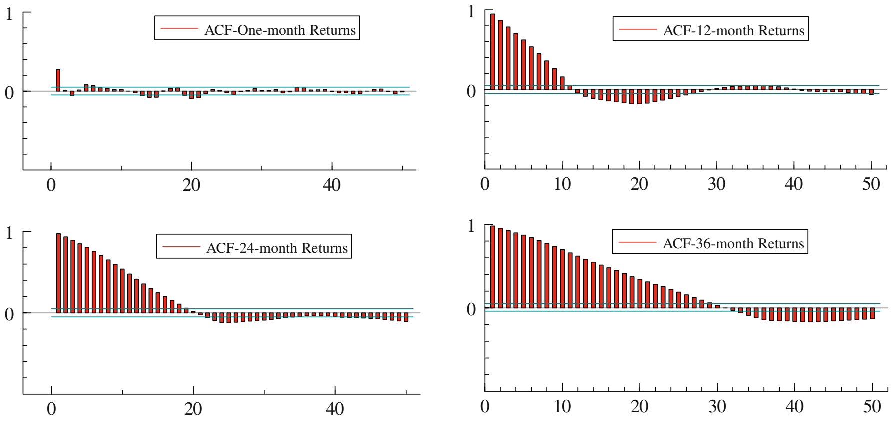

Fig. 1 Autocorrelation functions for log of real one-month and overlapping returns

Specifically, Fig. 1 shows that in each case the autocorrelations gradually decay to zero for the first N lags (where N = the number of cumulated months) and then oscillate (negative to positive) at longer horizons (through 50 lags here). The oscillating pattern becomes more well defined as the level of aggregation increases. For the 1-month returns, the ACF crosses zero at 1 month. For the 12-month overlapping returns, it crosses zero at about 12 months, and similarly for the 36-month overlapping returns. Many of these autocorrelations lie outside the standard 95% confidence interval. This increasing negativity was also found by Fama and French (up to return horizons 3–4 years). Thus the pattern in Fig. 1 is broadly consistent with the ∪-*shaped* behavior of regression coefficients described by Fama and French, who also used overlapping real data.

The key point here is the Souza result for the fractional differencing parameter with overlapping data implies that the behavior of the ACF function and the Fama and French overlapping regression model results will *both* be consistent with $d = 0.0$, stationary monthly stock returns (to within sampling error). That is, using the Souza result in Table 3, we find that although our *point estimates* of d are negative, *the 95% confidence intervals include zero*. Thus we *reject* the hypothesis of mean reversion in stock prices spuriously induced by negative autocorrelation in (overlapping) returns. In fact, our results imply that the permanent (random walk) component of stock prices *overwhelms* any temporary (mean reverting) predictable component, producing a fractional d-value for returns indistinguishable from zero.

Our results suggest that, *despite* previous efforts by Fama and French and others to address the problems caused by short-term ARMA dynamics, the observed pattern in their data of an increasingly negative (regression-based) ACF *remains an artifact* of short-term ARMA dynamics induced by (increasingly) overlapping returns. Therefore the usual regression model autocorrelation and heteroskedasticity correction procedures [e.g., Newey and West (1987) and Andrews (1991)] may *not* be useful with overlapping data. This also suggests the superiority of frequency domain fractional integration models to standard ACF models in the presence of significantly large short-term ARMA dynamics.

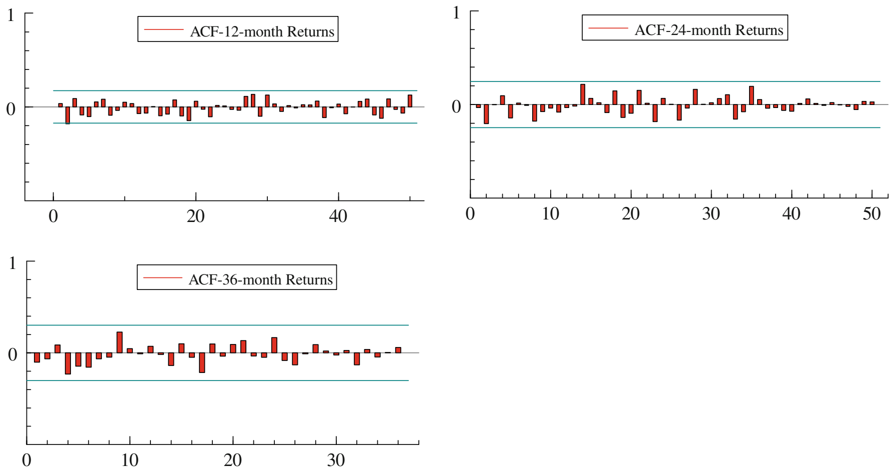

Fig. 2 Autocorrelation functions for log of real non-overlapping returns

To help clarify and illustrate our findings graphically, in Fig. 2 we present the ACFs for 12-, 24-, and 36-month *nonoverlapping* returns. If it is true that the pattern of increasingly negative ACFs as we lengthen the return horizon in Fig. 1 is largely the (artifactual) result of the increasing influence of short-term ARMA dynamics induced by overlapping returns, then we should *not* observe the same pattern in Fig. 2 as that presented in Fig. 1. An inspection of Fig. 2 supports our argument stated above. That is, the ACFs in Fig. 2 (for nonoverlapping returns) are *not* increasingly negative and stay within the 95% confidence bands around zero. This is consistent with our finding that the 1-month *and* overlapping stock returns in our data fit a $d = 0$ model and conform to the Souza result. We note that our findings are also consistent with the results presented in Boudoukh et al. (1994, 2008) who present detailed arguments based on market efficiency and persistent regressors, respectively, to show that prior research has significantly *overstated* the magnitude of autocorrelations of short- *and* long-horizon stock returns.

6 Summary and conclusion

Fractional integration, temporal aggregation, and overlapping data are frequently and increasingly used in financial research. Using recently developed econometric models of fractional integration with overlapping data, this study examined the time series properties of real monthly U.S. stock returns over the period 1871–2003, allowing the possibility of structural breaks during this period. We find that real one-month and rolling (overlapping) 12-, 24-, and 36-month returns are covariance stationary, with values for the memory parameter well approximated by $d = 0$ to within sampling error. Thus, we *reject* the hypothesis of mean reversion in stock *prices* induced by significant negative autocorrelation in *returns*. Our results imply that the permanent (random walk) component of stock prices *overwhelms* any temporary (mean reverting) component, producing a fractional d-value for stock *returns* indistinguishable from zero. We also confirm the Souza result. That is, holding the bandwidth constant, overlapping (and nonoverlapping) temporal aggregation should not affect frequency domain, semiparametric d-estimates such as the GPH and feasible exact local Whittle for a time series of interest.

Our results challenge the finding of Fama and French (1988) and others that has now become a *stylized fact* in financial economics: long-horizon stock returns are negatively autocorrelated and display a ∪-*shaped* pattern. Our results suggest that this finding is largely an artifact increasing short-term ARMA dynamics impacting the observed ACF. Our results also illustrate the limitations of standard ACF models of overlapping returns, and suggest that this problem can be minimized by the use of frequency domain fractional integration models and applying the Souza result. Thus the usual regression model autocorrelation and heteroskedasticity correction procedures [e.g., Newey and West (1987) and Andrews (1991)] may *not* be useful with overlapping data. We suggest that these are useful results for financial researchers who use fractional differencing, temporal aggregation, and overlapping data in their research.

The S&P 500 used here has the longest constructed history of any U.S. stock index and is often considered to "represent" the U.S. stock market. This is the main reason it has been the focus of many empirical studies of mean reversion. However, as noted in Siegel (2005) and elsewhere, the S&P 500 is primarily a large capitalization index with a growth stock bias. That is, it tends to slightly favor stocks with higher than average price-to-earnings and price-to-book ratios. While the mean reversion hypothesis is typically stated *without* reference to a specific market capitalization, Fama and French (1988) found that mean reversion is more evident in portfolios of small firms. For these reasons, it may be interesting and instructive for a future study to include small capitalization, and value and growth stocks in an analysis such as that presented here.

Acknowledgments The authors thank two referees for helpful and constructive comments. They also thank Leonardo Souza and Jeremy Smith for clarifying comments on an earlier draft, and Katsumi Shimotsu and Aaron Smith for graciously providing the Matlab and Gauss code to estimate their fractional differencing models. The standard disclaimer applies.

References

Agiakloglou C, Newbold P, Wohar M (1993) Bias in an estimator of the fractional difference parameter. J Time Ser Anal 14:235–246

Amemiya T, Wu RY (1972) The effect of aggregation on prediction in the autoregressive model. J Am Stat Assoc 67:628–632

Andrews DWK (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59:817–858

Baillie RT (1996) Long memory processes and fractional integration in econometrics. J Econom 73:5–59

Banerjee A, Urga G (2005) Modelling structural breaks, long memory and stock market volatility. J Econom 129:1–34

Beran J (1994) Statistics for long memory processes. Chapman & Hall, London

Bogle JC, Malkiel BG (2006) Turn on a paradigm? Wall Str J: A14.

Boudoukh J, Richardson MP, Whitelaw RF (1994) A tale of three schools: insights on autocorrelations of short-horizon stock returns. Rev Financ Stud 7:539–573

Boudoukh J, Richardson MP, Whitelaw RF (2008) The myth of long horizon predictability. Rev Financ Stud 21:577–1605

Campbell JY, Shiller RJ (1998) Valuation ratios and the long-run stock market outlook. J Portf Manag 24: 11–26
Caporale GM, Gil-Alana LA (2002) Fractional integration and mean reversion in stock prices. Q Rev Econ Financ 42:599–609
Cecchetti SG, Lam PS, Mark NC (1990) Mean reversion in equilibrium asset prices. Am Econ Rev 80:398–418
Chambers MJ (1998) Long memory and aggregation in macroeconomic time series. Int Econ Rev 39:1053–1072
Coggin TD (1998) Long-term memory in equity style indexes. J Portf Manag 24: 37–46
Daniel K (2001) The power and size of mean reversion tests. J Empir Financ 8:493–535
Davidson J, Hashimzade N (2009) Type I and type II fractional Brownian motions. Comput Stat Data Anal 53:2089–2106
DeBondt W, Thaler R (1989) A mean-reverting walk down wall street. J Econ Perspect 3: 189–202
DeJong DN, Nankervis JC, Savin NE, Whiteman CH (1992) Integration versus trend-stationarity in time series. Econometrica 60:423–433
Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 74:427–431
Diebold FX, Inoue A (2001) Long memory and regime switching. J Econom 105:131–159
Diebold F, Rudebusch G (1991) On the power of Dickey–Fuller tests against fractional alternatives. Econ Lett 35:155–160
Doukhan P, Oppenheim G, Taqqu MS (eds) (2003) Theory and applications of long-range dependence. Birkhauser, Boston
Fama EF, French KR (1988) Permanent and temporary components of stock prices. J Political Econ 96:246–273
Geweke J, Porter-Hudak S (1983) The estimation and application of long memory time series models. J Time Ser Anal 4:221–238
Gil-Alana LA (2004) A joint test of fractional integration and structural breaks at known periods of time. J Time Ser Anal 25:691–700
Gil-Alana LA (2008) Fractional integration and structural breaks at unknown periods of time. J Time Ser Anal 29:163–185
Gil-Alana LA, Hualde J (2009) Fractional integration and cointegration. An overview and an empirical application. In: Mills TC, Patterson K (eds) Palgrave handbook of econometrics, vol 2. Palgrave Macmillan, New York
Gil-Alana LA, Robinson PM (1997) Testing of unit root and other nonstationary hypotheses in macroeconomic time series. J Econom 80:241–268
Granger CWJ (1980) Long memory relationships and the aggregation of dynamic models. J Econom 14:227–238
Granger CWJ, Hyung N (2004) Occasional structural breaks and long memory with an application to the S&P 500 absolute returns. J Empir Financ 11:399–421
Granger CWJ, Joyeux R (1980) An introduction to long memory time series and fractional differencing. J Time Ser Anal 1:15–30
Hansen LP, Hodrick RJ (1980) Forward exchange rates as optimal predictors of future spot rates. J Political Econ 88:829–853
Harri A, Brorsen BW (2002) The overlapping data problem. Working paper, Department of Agricultural Economics, Oklahoma State University. http://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=72728
Hassler U, Wolters J (1994) On the power of unit root tests against fractional alternatives. Econ Lett 45:1–5
Henry M, Zaffaroni P (2003) The long range dependence paradigm for macroeconomics and finance. In: Doukhan P, Oppenheim G, Taqqu MS (eds) Theory and applications of long-range dependence. Birkhauser, Boston
Hosking JRM (1981) Fractional differencing. Biometrika 60:165–176
Hurvich CM, Deo R, Brodsky J (1998) The mean square error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. J Time Ser Anal 19:19–46
Kim JK, Nelson CR (1998) Testing for mean reversion in heteroskedastic data II. J Empir Financ 5:385–396
Kim JK, Nelson CR, Startz R (1991) Mean reversion in stock prices. Rev Econ Stud 58:515–528
Kim JK, Nelson CR, Startz R (1998) Testing for mean reversion in heteroskedastic data based on Gibbs-sampling-augmented randomization. J Empir Financ 5:131–154

Kramer W (1998) Fractional integration and the augmented Dickey–Fuller test. Econ Lett 61:269–272
Lee D, Schmidt P (1996) On the power of the KPSS test of stationarity against fractionally integrated alternatives. J Econom 73:285–302
Lee J, Strazicich MC (2003) Minimum Lagrange multiplier unit root test with two structural breaks. Rev Econ Stat 85:1082–1089
Lee J, Strazicich MC (2004) Minimum LM unit root test with one structural break. Working paper, Department of Economics, Appalachian State University. http://econ.appstate.edu/RePEc/pdf/wp0417.pdf
Leybourne SJ (1995) Testing for unit roots using forward and reverse Dickey–Fuller regressions. Oxf Bull Econ Stat 57:559–571
Leybourne S, Kim TH, Newbold P (2005) Examination of some more powerful modifications of the Dickey–Fuller test. J Time Ser Anal 26:355–369
Lo AW (1991) Long-term memory in stock market prices. Econometrica 59:1279–1313
Lo AW, MacKinlay AC (1988) Stock market prices do not follow random walks. Rev Financ Stud 1: 41–66
Lobato IN, Savin NE (1998) Real and spurious long-memory properties of stock-market data. J Bus Econ Stat 16:261–268
Maddala GS, Kim IM (1999) Unit roots, cointegration, and structural change. Cambridge University Press, Cambridge
Marcellino M (1999) Some consequences of temporal aggregation in empirical analysis. J Bus Econ Stat 17:129–136
Marinucci D, Robinson PM (1999) Alternative forms of fractional Brownian motion. J Stat Inference Plan 80:111–122
Moulines E, Soulier P (2003) Semiparametric spectral estimation for fractional processes. In: Doukhan P, Oppenheim G, Taqqu MS (eds) Theory and applications of long-range dependence. Birkhauser, Boston
Muller UA (1993) Statistics of variables observed over overlapping intervals. Discussion paper, O&A Research Group, Olsen, Ltd., Zurich. http://www.olsen.ch/fileadmin/Publications/Working_Papers/931130-ntervalOverlap.pdf
Newey WK, West KD (1987) A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55:703–708
Ohanissian A, Russell JR, Tsay RS (2008) True or spurious long memory? A new test. J Bus Econ Stat 26:161–175
Perron P, Vodounou C (2004) Tests of return predictability. J Empir Financ 11:202–230
Phillips PCB (2001) Descriptive econometrics for non-stationary time series with empirical illustrations. J Appl Econom 16:389–413
Phillips PCB, Xiao Z (1999) A primer on unit root testing. In: McAleer M, Oxley L (eds) Practical issues in cointegration analysis. Blackwell Publishers, Oxford. Also appears in J Econ Surv 12:423–470 (1998)
Poterba JM, Summers LH (1988) Mean reversion in stock prices. J Financ Econ 22:27–59
Richardson M (1993) Temporary components of stock prices: a skeptic's view. J Bus Econ Stat 11:199–207
Richardson M, Smith T (1991) Tests of financial models in the presence of overlapping observations. Rev Financ Stud 4:227–254
Robinson PM (1978) Statistical inference for a random coefficient autoregressive model. Scand J Stat 5:163–168
Robinson PM (1994) Efficient tests of nonstationary hypotheses. J Am Stat Assoc 89:1420–1437
Robinson PM (1995a) Log-periodogram regression of time series with long range dependence. Ann Stat 23:1048–1072
Robinson PM (1995b) Gaussian semiparametric estimation of long range dependence. Ann Stat 23:1630–1661
Robinson PM (ed) (2003) Time series with long memory. Oxford University Press, Oxford
Robinson PM, Henry M (1999) Long and short memory conditional heteroskedasticity in estimating the memory parameters of levels. Econom Theory 15:299–336
Said SE, Dickey DA (1984) Testing for unit roots in autoregressive moving average models of unknown order. Biometrika 71:599–607
Shimotsu K (2006) Simple (but effective) tests of long memory versus structural breaks. Working paper no. 1101, Department of Economics, Queens University (Canada). http://qed.econ.queensu.ca/pub/faculty/shimotsu/

Shimotsu K (2009) Exact local Whittle estimation of fractional integration with unknown mean and time trend. Econom Theory, published online (First View) by Cambridge University Press 07 Oct 2009. doi:10.1017/S0266466609990065
Shimotsu K, Phillips PCB (2005) Exact local Whittle estimation of fractional integration. Ann Stat 33:1890–1933
Shimotsu K, Phillips PCB (2006) Local Whittle estimation of fractional integration and some of its variants. J Econom 130:209–233
Siegel JJ (2005) The future for investors. Crown Business Press, New York
Silvestrini A, Veredas D (2008) Temporal aggregation of univariate and multivariate time series models, a survey. J Econ Surv 22:458–497
Smith A (2005) Level shifts and the illusion of long memory in economic time series. J Bus Econ Stat 23:321–335
Smith J, Taylor N, Yadav S (1997) Comparing the bias and misspecification in AFRIMA models. J Time Ser Anal 18:507–528
Souza LR (2004) Spectral properties of temporally aggregated long memory processes. Working paper, The Graduate School of Economics (EPGE), Rio de Janeiro. http://www.econ.puc-rio.br/pdf/seminario/2004/a%20note%20on%20aggreg.pdf
Souza LR (2005) A note on Chambers' "Long memory and aggregation in macroeconomic time series". Int Econ Rev 46:1059–1062
Souza LR (2007) Temporal aggregation and bandwidth selection in estimating long memory. J Time Ser Anal 28:701–722
Souza LR, Smith J (2002) Bias in the memory parameter for different sampling rates. Int J Forecast 18:299–313
Souza LR, Smith J (2004) Effects of temporal aggregation on estimates and forecasts of fractionally integrated processes. Int J Forecast 20:487–502
Stock JH (1991) Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series. J Monet Econ 28:435–459
Stock JH (1994) Unit roots, structural breaks, and trends. In: Engle R, McFadden D (eds) Handbook of econometrics, vol IV. Elsevier, Amsterdam
Summers LH (1986) Does the stock market rationally reflect fundamental values? J Financ 41:591–600
Tanaka K (1999) The nonstationary fractional unit root. Econom Theory 15:549–582