An On-line Predictive Test for the Common Break

Cindy S.H. Wang*, Rui Fan†, Yimeng Xie‡

November 30, 2023

Abstract

This paper introduces an on-line common break test designed to identify optimal timing for portfolio adjustments especially under volatile market conditions. The key advantage of this on-line approach lies in its potential to significantly mitigate losses stemming from market and economic uncertainties. This is achieved through a robust estimation of the long memory parameter in an ARFIMA model, referred to as the FEAR estimation, coupled with a predictive test for structural breaks in the aggregated return series. Comparative analyses underscore the superior efficacy of the FEAR estimation relative to prevailing methodologies. Simulations validate the theoretical underpinnings and illustrate the promising finite sample performance of the proposed procedure. Key findings include: (i) the long memory property in the aggregation of multiple time series, with or without breaks, and (ii) the effectiveness of the on-line common break test in accurately detecting the location of common breaks in each series of the portfolio. Empirical results further underscore the practicability and effectiveness of the on-line common break test from a risk management perspective.

Key words: Long memory process, aggregated time series, FEAR-estimation, fractional differencing parameter, predictive test

MOS subject classification: 91B84; 62M10.

*Corresponding author, HSBC Business School, Peking University (UK Campus), Oxford, U.K. Email: cindywang@phbs.pku.edu.cn.

†Department of Economics, Rensselaer Polytechnic Institute, U.S.A. Email: fanr5@rpi.edu.

‡Department of Statistics and Data Science, School of Economics, and Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, China. Email: yimengxie@xmu.edu.cn.

1 Introduction

The timely adjustment of asset allocations within a portfolio has long posed a formidable challenge for both academic researchers and practitioners, including policy makers and fund managers. This challenge becomes especially pronounced during periods of significant market shocks or heightened turbulence. The ability to adeptly track market fluctuations and assess market stability emerges as a pivotal factor in successfully predicting market returns and subsequently optimizing portfolio allocations. This paper addresses this challenge by introducing an easy-to-implement and forward-looking methodology designed to efficiently detect changes in the persistence level within a portfolio. The proposed approach is particularly relevant when the aggregation of a portfolio follows a long-memory process. It aims to provide a discernible signal for portfolio adjustments, offering a practical solution for navigating dynamic market conditions.

The conventional approach to portfolio adjustment primarily revolves around predicting the returns of individual assets within the portfolio, wherein the overall portfolio is constructed by aggregating each individual return forecast. Studies such as Lustig et al. (2014) and Lyle and Wang (2015) have explored this methodology, delving into the prediction of portfolio outcomes using information from each single asset. However, this conventional strategy overlooks the correlations among diverse types of financial assets. These correlations may arise from common trends, shocks, or policy changes, see Ramchand and Susmel (1998), Solnik and Roulet (2000), Pesaran (2015) and Wang et al. (2021) for example. That induce simultaneous movements across various financial assets. In an era of increasingly integrated international financial markets, comprehending and identifying these correlations becomes pivotal for gaining insights into global market dynamics, thereby offering valuable information to adjust portfolios and enhance performance. In particular, the identification of common breaks in a multivariate system, where a change occurs in each time series at a common point (Lütkepohl (1989) and Wang and Wan (2020)), holds particular importance.

Distinct from the conventional approach, this paper specifically concentrates on evaluating the aggregation of time series, capturing the correlations among them. In practice, common shocks or events induce simultaneous movements among market assets, leading to changes in their interactions or correlations. This suggests the potential existence of common breaks or shifts in the time series of individual assets within a portfolio. Consequently, tracking the locations of breaks becomes essential as they signify the adjustment times for portfolio allocations. It is noteworthy that the aggregation of autoregressive (AR) time series often manifests statistical properties indicative of a long memory process, when examining the aggregation of asset series. This phenomenon was initially investigated by Robinson (1978) and Granger (1980). Subsequently, Zaffaroni (2004) generalized these findings by considering the aggregation of heterogeneous ARMA time series. In a related body of research, the association between structural breaks and long memory properties has also been extensively explored. For example, Diebold and Inoue (2001) show that the low-frequency periodogram of the Markov-switching process can be approximated well by a long memory process. The detailed regarding discussion could be found in Granger and Hyung (2004) and Choi et al. (2010). Thus, drawing inspiration from the findings of the aforementioned studies, our *first step* in identifying signals for portfolio adjustment entails aggregating the asset returns within the considered portfolio and subsequently estimating the long memory parameter of this aggregated return series.

It is pertinent to note that in practice, single asset returns are often modeled as autoregressive (AR) processes, providing empirical support for considering the aggregated return series for a portfolio as a long-memory process (Ferson et al. (2003) and Dai et al. (2021)). In this paper, we employ an ARFIMA(p, d, q) model to characterize the aggregated time series, with parameter d capturing the persistence level of the time series.

One contribution of this paper is the introduction of an easily implementable method, denoted as FEAR, designed for estimating the fractional integrated order d for aggregated return series. The literature contains numerous studies on parameter d estimation, encompassing methods such as exact maximum likelihood estimation (MLE) by Sowell (1992), generalized minimum distance estimation(GMD) by Mayoral (2007), local Whittle estimation (ELW) by Shimotsu and Phillips (2005), and GPH estimation by Geweke and Porter-Hudak (1983). However, Sowell (1992) points out the exact MLE of the fractional parameter d is quite time-consuming and is not very accurate in finite samples, especially when d is close to 0.5 and the sample size is small ($T = 100$ and 200). In addition, the presence of AR parameters greatly complicate the computation of the corresponding autocovariance functions. Biased estimated ds generated by the GPH are also found in small finite samples and when d is close to 0, even being less computational burden (see Crato and Ray (1996)). ELW of Shimotsu and Phillips is sensitive to bandwidth selection while GMD of Mayoral needs prior information of parametric setups and weighting matrix but is less efficient than exact MLE.

Diverging from existing approaches, the FEAR method, proposed herein, estimates d by minimizing the sum of squared forecast errors. A key advantage lies in its independence from the estimation of the p autoregressive parameters and the q moving average parameters in an ARFIMA model. Instead, we approximate the ARFIMA model using an AR(k) model, with the autoregressive order k easily selected through information criteria such as AIC. Poskitt (2007) has theoretically established the uniform convergence and asymptotic distribution of estimated coefficients in an AR(k) model when the finite-order autoregressive model is employed to approximate fractionally integrated processes. The AR approximation of the fractionally integrated process not only reduces the number of parameters requiring estimation but also mitigates forecasting errors stemming from model misspecification in estimating the values of p and q. This dual advantage enhances the practicality and robustness of the FEAR method in estimating d for aggregated return series.

This paper theoretically shows the consistency and asymptotic normality of the FEAR estimation. The range of $d \in (-0.5, 0.5)$ can be covered by the FEAR estimation rather than the limitation on $d \in (-0.5, 0.25)$ by Tieslau et al. (1996). Our simulation results also affirm the promising performance of the FEAR estimation, as evidenced by lower root mean squared errors (RMSEs) and reduced computation time compared to the widely used estimations under consideration. In addition, we conduct simulations on time series featuring both cross-sectional and time dependence, estimating the integrated order of the aggregated series using our newly designed FEAR estimation and the GPH estimation, known for its robustness to short-run dynamics. Broadly, the aggregate series of time-dependent variables consistently exhibit long-memory properties. Furthermore, as cross-sectional correlations strengthen, the aggregate series display increased persistence, contributing to the enhanced efficiency of aggregate forecasting.

The *second step* we advocate for identifying signals for portfolio adjustment involves implementing a recursive predictive t-test. This test is formulated based on the estimate of the fractional differencing parameters d using the FEAR estimation, as proposed in the first step. Additionally, we establish that the limiting distribution of this predictive test follows a standard normal distribution under the null hypothesis of no break in d. The recursive FEAR predictive test scheme is computed as follows: given a training sample of size t, we estimate the first d value, denoted as $\hat{d}_{FEAR,t}$, using the training sample with the FEAR estimation. Subsequently, we add one observation to the estimation sample and generate the second estimate of d, $\hat{d}_{FEAR,t+1}$. We then test for the difference between $\hat{d}_{FEAR,t}$ and $\hat{d}_{FEAR,t+1}$ using the predictive t-test. This process continues until the null hypothesis is rejected, i.e., a shift or break in d is detected. In this regard, this procedure inherently possesses an on-line feature. Given the strong connection between persistence in aggregate time series and correlation across individual time series, any statistically significant change in d may indicate a shift in the underlying dependent structure among assets, thereby serving as a signal for portfolio adjustment.

The Monte Carlo simulations illustrate the finite sample performance of the FEAR estimation and test. Our findings indicate that FEAR estimates are generally consistent across various data generating processes characterized by long memory. Notably, the performance of FEAR surpasses that of several classical estimation methods for fractional differencing parameters, particularly in scenarios where d is close to 0.5 and the sample size is small ($T = 100$ and 200). Additionally, we simulate multiple time series characterized by short memory processes and cross-sectional correlations. Employing the FEAR estimation method to estimate these time series, the results demonstrate that the aggregated time series displays a typical long memory pattern. Furthermore, we conduct simulations involving multiple time series with common breaks. The fractional differencing parameters of the aggregated time series are estimated using the FEAR method, and the aggregation is examined through our FEAR-predictive test. The simulation outcomes indicate that the test generally produces correct size and exhibits satisfactory power performances.

In practical terms, identifying the timing of changes in the underlying dependent structure holds paramount importance for fund managers. Timely detection allows them to respond promptly to the impacts of market shocks, thereby mitigating risks. In our empirical analysis, we establish three country-specific portfolios to validate the effectiveness of our methodology. Our predictive test demonstrates the ability to promptly capture significant market shocks, furnishing early warning signals for fund managers to adjust their portfolios. In particular, empirical evidence reveals the stability of the recursive estimates of fractional differencing parameters of three aggregated return series when there is no common event around. But the jumps or down shifts promptly happened as long as common down shocks come.

In summary, this paper makes a key contribution by introducing a two-step procedure for portfolio adjustment. The initial step involves the FEAR estimation, a novel method for efficiently gauging the fractional integrated order d and measuring the persistence of the aggregated return series, typically following a long-memory process. Without the need to estimate the p autoregressive and q moving average parameters in an ARFIMA model, we demonstrate that the FEAR estimation attains asymptotic consistency and efficiency, particularly when the ARFIMA model is well approximated by a finite-order AR(k) model. The computational advantage of the FEAR estimation lies in its simplicity—implemented by minimizing the sum of squared forecast errors—offering a practical alternative to existing, often intricate and resource-intensive estimation methods. The second step involves the FEAR-predictive test with a recursive window framework to easily detect structural changes in the underlying dependence among assets, a factor typically overlooked in conventional portfolio adjustment approaches. Simulation and empirical analysis results jointly underscore the substantial informational value gained from a nuanced understanding of asset correlations. This knowledge proves instrumental in identifying market dynamics, facilitating informed portfolio construction, and enhancing risk management strategies.

The paper is organized as follows. Section 2 presents the FEAR estimation of the fractional integrated order d and its asymptotic properties. Section 3 describes the proposed tests used to detect the change in d, proving signals for portfolio adjustment. In Section 4, we demonstrate the Monte Carlo simulations. Section 5 conducts empirical analysis by establishing three country-specific portfolios. Section 6 concludes.

2 The FEAR Estimation of the Fractional Differencing Parameter d

2.1 Model and Assumptions

The ARFIMA(p, d, q) model of a time series y_t is defined by

$$\phi(L)(1-L)^d y_t = \psi(L)\epsilon_t, \tag{2.1}$$

where $\phi(z) = 1 - \sum_{j=1}^{p} \phi_j z^j$ and $\psi(z) = 1 - \sum_{j=1}^{q} \theta_j z^j$ are polynomials with roots outside the unit circle. Without loss of generality, we assume that y_t is a linear process without a deterministic term. The lag operator L is defined by $Ly_t = y_{t-1}$. Also, the fractional difference is defined by its binomial expansion:

$$(1-L)^d = \sum_{j=0}^{\infty} \Delta_j(d) L^j,$$

where $\Delta_j(d) = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}$ with $\Gamma(\cdot)$ denoting the gamma (generalized factorial) function.

Assumption 1 *Let $\mathcal{I}_t$ denote the $\sigma-$algebra of event determined by ϵ_s with $s \le t$. Assume that ϵ_t is ergodic and that*

$$E\left[\epsilon_t | \mathcal{I}_{t-1}\right] = 0, \quad E\left[\epsilon_t^2 | \mathcal{I}_{t-1}\right] = \sigma_\epsilon^2, \quad E(\epsilon_t^4) < \infty.$$

Assumption 2 *The series y_t in (2.1) is a covariance-stationary process with a fractional differencing parameter $d \in (-1/2, 1/2)$.* [1]

Following Brockwell and Davis (2009), the long-memory process in (2.1) can be effectively represented by an infinite order autoregressive process:

$$(1-L)^d y_t = \psi(L)/\phi(L)\epsilon_t = u_t, \tag{2.2}$$

$$y_t = \sum_{j=1}^{\infty} \beta_j y_{t-j} + u_t, \tag{2.3}$$

where u_t follows an ARMA process. In practice, however, we will need a finite autoregression model to serve as an approximation to reality. Poskitt (2007) shows that y_t can be approximated by an AR(k)

[1] This assumption can be generalized to accommodate any value of d. For example, Beran (1995) demonstrates how to estimate any real $d > -1/2$ using an approximate maximum likelihood method.

model where k can be selected using the Akaike information criterion (AIC):

$$y_t = \sum_{j=1}^{k} \beta_j y_{t-j} + v_t, \tag{2.4}$$

where v_t is the prediction error. Note that the autoregressive coefficient β_j is a function of d where $\beta_j = \Delta_j(d)$. By Equation (2.4), v_t can be expressed as a function of d:

$$v_t(d) = y_t - \Delta_j(d) y_{t-j}. \tag{2.5}$$

2.2 The FEAR Estimation of d and its Asymptotic Properties

For a time series y_t following an ARFIMA process as defined in Equation (2.1), we propose an easily implementable method to estimate the fractional differencing parameter d. This method estimates d by approximating the ARFIMA model with an AR(k) model, thus termed the FEAR estimation. The estimation process involves the following steps:

1. Approximation by AR(k): First, we approximate the long-memory process y_t by an AR(k) model, with the autoregressive order k selected using AIC. Subsequently, we obtain the estimated coefficients ($\tilde{\beta}_j$) and the estimated prediction error ($\tilde{v}_t$).

2. Estimation of d: Next, we estimate the fractional differencing parameter d by minimizing the squared distance between v_t and $\tilde{v}_t$ using the following optimization problem:

$$\hat{d} = \arg\min_d S_T(d), \tag{2.6}$$

where $S_T(d) \equiv \sum_{t=1}^{T} \left[v_t(d) - \tilde{v}_t\right]^2$.

To solve (2.6), we compute the first order derivative of the objective function, $S_T(d)$, with respect to d:

$$\frac{\partial S_T(d)}{\partial d} = 2 \sum_{t=1}^{T} \left[v_t(d) - \tilde{v}_t\right] v_t'(d), \tag{2.7}$$

where $v_t'(d)$ is the first order derivative of $v_t(d)$ with respect to d.

By the expression of $v_t(d)$ in (2.5), we have

$$\begin{aligned} v_t(d) - \tilde{v}_t &= \left(y_t - \sum_{j=1}^{k} \beta_j y_{t-j}\right) - \left(y_t - \sum_{j=1}^{k} \tilde{\beta}_j y_{t-j}\right) \\ &= \sum_{j=1}^{k} \left[\Delta_j(d) - \tilde{\beta}_j\right] y_{t-j} \end{aligned} \tag{2.8}$$

Now consider $v_t'(d)$. Note that the derivative of the gamma function $\Gamma(z)$ with respect to z is:

$$\frac{d}{dz}\Gamma(a) = \Gamma(z)\psi^{di}(z), \tag{2.9}$$

where $\psi^{di}(z)$ is the digamma function with $\psi^{di}(z) = \Gamma'(z)/\Gamma(z)$. Hence, by applying the chain rule and the derivative of the gamma function in (2.9), we can show that the derivative of $\Delta_j(d)$ with respect to d, for $j = 1, 2, \ldots, k$, is:

$$\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}\psi^{di}(j+d) - \frac{\Gamma(j+d)}{(\Gamma(d))^2\Gamma(j+1)}\psi^{di}(d). \tag{2.10}$$

Based on the definition of $v_t(d)$ and summing up (2.10) over the range of j from 1 to k, we get:

$$\begin{aligned} v_t'(d) &= -\sum_{j=1}^{k}\left(\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}\psi^{di}(j+d) - \frac{\Gamma(j+d)}{(\Gamma(d))^2\Gamma(j+1)}\psi^{di}(d)\right) y_{t-j} \\ &= -\sum_{j=1}^{k}\left[\psi^{di}(j+d) - \frac{1}{\Gamma(d)}\psi^{di}(d)\right]\Delta_j(d) y_{t-j}. \end{aligned} \tag{2.11}$$

Utilizing (2.8) and (2.11), Equation (2.7) can now be expressed as:

$$\begin{aligned} \frac{\partial S_T(d)}{\partial d} &= 2\sum_{t=1}^{T}\left[\sum_{j=1}^{k}\left(\Delta_j(d) - \tilde{\beta}_j\right) y_{t-j}\right]\left[-\sum_{j=1}^{k}\left(\psi^{di}(j+d) - \frac{1}{\Gamma(d)}\psi^{di}(d)\right)\Delta_j(d) y_{t-j}\right] \\ &= -2\sum_{t=1}^{T}\left[\sum_{j=1}^{k}\left(\Delta_j(d) - \tilde{\beta}_j\right) y_{t-j}\right]\left[\sum_{j=1}^{k}\left(\psi^{di}(j+d) - \frac{1}{\Gamma(d)}\psi^{di}(d)\right)\Delta_j(d) y_{t-j}\right] \\ &= -2\sum_{t=1}^{T}\sum_{j=1}^{k}\sum_{i=1}^{k}\left(\Delta_j(d) - \tilde{\beta}_j\right)\left(\psi^{di}(i+d) - \frac{1}{\Gamma(d)}\psi^{di}(d)\right)\Delta_i(d) y_{t-j}y_{t-i} \\ &= -2\sum_{j=1}^{k}\sum_{i=1}^{k}\left(\Delta_j(d) - \tilde{\beta}_j\right)\left(\psi^{di}(i+d) - \frac{1}{\Gamma(d)}\psi^{di}(d)\right)\Delta_i(d)\left(\sum_{t=1}^{T} y_{t-j}y_{t-i}\right). \end{aligned} \tag{2.12}$$

Lastly, the estimator $\hat{d}$ can be computed by letting $\frac{\partial S_T(d)}{\partial d} = 0$, which indicates

$$\sum_{j=1}^{k}\sum_{i=1}^{k}\left(\Delta_j(\hat{d}) - \tilde{\beta}_j\right)\left(\psi^{di}(i+\hat{d}) - \frac{1}{\Gamma(\hat{d})}\psi^{di}(\hat{d})\right)\Delta_i(\hat{d})\left(\frac{1}{T}\sum_{t=1}^{T} y_{t-j}y_{t-i}\right) = 0. \tag{2.13}$$

Let $\gamma_\tau = \gamma_{-\tau} = E(y_t y_{t+\tau})$ for $\tau = 0, 1, \ldots$, denote the autocovariance function of y_t. Then the sample autocovariance function can be defined by $\hat{\gamma}_\tau = \frac{1}{T}\sum_{t=1}^{T} y_t y_{t+\tau}$. So (2.13) can be rewritten as

$$\sum_{j=1}^{k}\sum_{i=1}^{k}\left(\Delta_j(\hat{d}) - \tilde{\beta}_j\right)\left(\psi^{di}(i+\hat{d}) - \frac{1}{\Gamma(\hat{d})}\psi^{di}(\hat{d})\right)\Delta_i(\hat{d})\hat{\gamma}_{j-i} = 0.$$

Numerical methods or optimization techniques are necessary to determine $\hat{d}$, as it lacks an analytical solution.

In comparison to traditional approaches for estimating d, the preceding discussion highlights the efficacy of FEAR estimation as a viable alternative. Following Poskitt (2007), who theoretically established the uniform convergence and asymptotic distribution of estimated coefficients in an AR(k) model when a finite-order AR model is employed to approximate an ARFIMA model, the FEAR estimation approximates an ARFIMA(p, d, q) model using an AR(k) model. This approximation significantly streamlines the estimation process. Unlike conventional methods, there is no requirement to accurately specify the autoregressive order p and moving average order q. Consequently, there is no need to estimate the p autoregressive parameters and the q moving average parameters in an ARFIMA model. This not only simplifies the estimation and computation procedures but also diminishes estimation errors arising from model misspecification.

Subsequently, we lay down the theoretical groundwork for the consistency and asymptotic distribution of the estimated fractional differencing parameter d through the utilization of the FEAR method.

Theorem 2.1 (Consistency of $\hat{d}$) *If y_t is a stationary process that satisfies Assumption 1 and Assumption 2, then the FEAR estimation of d is consistent.*

Let $\hat{D}_j = \sum_{i=1}^{k}\frac{\partial\Delta_i(d)}{\partial d}\hat{\rho}_{j-i}$ and $D_j = \sum_{i=1}^{k}\frac{\partial\Delta_i(d)}{\partial d}\rho_{j-i}$ for $j = 1, .., k$. The sample autocorrelation of y_t is defined by $\hat{\rho}_{j-i} = \frac{\sum_{t=1}^{T} y_{t-j}y_{t-i}}{\sum_{t=1}^{T} y_t^2}$. Define a $1\times k$ vector $\hat{W} = (\hat{D}_1, \hat{D}_2, \cdots, \hat{D}_k)$. Since $\hat{\rho}$ is a consistent estimator for the population autocorrelation ρ (see Appendix), we have $\hat{D}_j \to_p D_j$ for $j = 1, ..., k$, and $\hat{W} \to_p W$, where $W = (D_1, D_2, \cdots, D_k)$. Also, we define a $1\times k$ vector $C = (\frac{\partial\Delta_1(d)}{\partial d}, \frac{\partial\Delta_2(d)}{\partial d}, \cdots, \frac{\partial\Delta_k(d)}{\partial d})$.

Theorem 2.2 (Asymptotic Normality of $\hat{d}$) *Under Assumption 1 and Assumption 2, the FEAR estimator $\hat{d}$ asymptotically follows a normal distribution:*

$$\sqrt{T}(\hat{d} - d) \to_d \mathcal{N}\left(0, \sigma_\epsilon^2\frac{W\Gamma_k^{-1}W'}{(CW')^2}\right), \tag{2.14}$$

where $\Gamma_k = [\gamma_{j-i}]_{j,i=1,\ldots,k}$ and γ denotes the autocovariance function of y_t.

3 The On-line Common Break Test

A common break in multivariate system of time series means that a structural break occurs in each series at a common date. Similar with "fixed τ model" considered in Joseph and Wolfson (1992) and a consequent common break model in multivariate system considered in Lütkepohl (1989) and Wang and Wan (2020), here we consider the following model

$$
\begin{aligned}
Y_{it} &= \mu_1 + e_{it}, \quad t = 1, 2, \ldots, k_0 \\
Y_{it} &= \mu_2 + e_{it}, \quad t = k_0 + 1, \ldots, T \\
i &= 1, 2, \ldots, N
\end{aligned}
$$

where N is finite and $E(e_{it}) = 0$ for all i and t. In this model, a structural break occurs at the same point k_0 in each series, which we call the common break date. Before the break, the mean of Y_{it} is μ_1, while the mean becomes μ_2 after the break. Here the mean shift $\mu_2 - \mu_1$ indicates the magnitude of break. To assure the existence of a break, we set $1 < k_0 < T$, i.e. $k_0 = [\tau_0 T]$ with $0 < \tau_0 < 1$.

Inspired by the finding of Granger (1980) illustrating the aggregation of autoregressive (AR) time series can exhibit long memory properties, thus when summing up multiple individual time series, the aggregated time series could be a long memory process. Zaffaroni (2004) subsequently generalized these findings by considering the aggregation of heterogeneous ARMA time series. Moreover, Diebold and Inoue (2001) show that a mean structural break model displays a long memory process as well[2]. Granger and Hyung (2004) show that an increase of the number of mean breaks makes the memory of the process seemingly more persistent. Choi et al. (2010) support this phenomenon by examining the breaks in a realized volatility with long memory property. That implies when common breaks occur in each series of a multivariate system, there exists a sudden change in the value of the long memory parameter of the aggregated series[3]. To be more specific, when common break in mean comes, a sudden change in the fractional differencing parameter happens.

We then introduce a predictive test leveraging the FEAR estimation to identify common breaks in a multivariate system. The detailed implementation of this predictive test is outlined below.

- Step 1. Based on the aforementioned argument, we first sum up N time series to generate an aggregated series S_t, i.e. $S_t = \sum_{i=1}^{N} Y_{it}$, which follows a long memory process.

- Step 2. We estimate the fractional differencing parameter of S_t by the new proposed FEAR estimation with the recursive window scheme. That means when one new observation arrives, we reestimate the fractional differencing parameter of the new aggregated series S_t by the FEAR

[2]In fact, Diebold and Inoue (2001) also show that the low-frequency periodogram of the Markov-switching process can be approximated well by a long memory process.

[3]This phenomenon is also supported by our Monte Carlo simulations in Section 4.2.

estimation until the end of observation T, given a training sample of size $W = \delta T$, $\delta \in (0,1)$. We can thus obtain a series of estimated fractional differencing parameters.

- Step 3. Intuitively, in order to detect the sudden change in ds, we construct the following predictive-t test called the on-line common break test,

$$\hat{t} = \frac{\hat{d}_{FEAR,t+1} - \hat{d}_{FEAR,t}}{\hat{\sigma}_{t+1}}$$

where $\hat{d}_{FEAR,t}$ denotes FEAR estimation of d value in the recursive scheme with the t observations, while $\hat{d}_{FEAR,t+1}$ denotes d estimator with $t+1$ observations. $\hat{\sigma}_{t+1}$ is the standard deviation of $\hat{d}_{FEAR,t+1} - \hat{d}_{FEAR,t}$. That is to say, this recursive FEAR predictive test scheme is computed as follows: given a training sample of size t, we estimate the first d value, say $\hat{d}_{FEAR,t}$, using the training sample with the FEAR estimation. Then we add one observation to the estimation sample, and generate the second estimate of d, $\hat{d}_{FEAR,t+1}$. We then test for difference between $\hat{d}_{FEAR,t}$ and $\hat{d}_{FEAR,t+1}$ by this new test. We continue like this until the null hypothesis of no break is rejected, i.e. we can find a shift or break in d.

The resulting limiting distribution of this $\hat{t}$ is summarized in Theorem 3.1.

Theorem 3.1 *When a data generating process (DGP) displays a long memory process and satisfies (2.1), under the null hypothesis of no change in* d*, the on-line common break test built up on the FEAR estimation asymptotically follows a standard normal distribution,*

$$\hat{t} \to_d N(0,1).$$

4 Monte Carlo Simulations

In this section, we examine the finite sample performances of FEAR estimation and the on-line common break test. Each Monte Carlo simulation experiment in this section is conducted 1000 times.

4.1 Simulations for Estimations of Fractional Differencing parameters

In order to present comprehensive comparison of FEAR and other methods, we consider DGPs of three different long memory processes including fractional white noise process ARFIMA $(0,d,0)$, special one ARFIMA $(1,d,0)$, general case ARFIMA $(1,d,1)$. In particular, these DGPs are designed as following,

$$(1-L)^d (y_t - \mu) = \varepsilon_t, \tag{4.1}$$

$$(1 - \phi L)(1 - L)^d (y_t - \mu) = \varepsilon_t, \tag{4.2}$$

$$(1 - \phi L)(1 - L)^d (y_t - \mu) = (1 + \theta L)\varepsilon_t, \tag{4.3}$$

where $\mu = 2$ and $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$ with $\sigma^2 = 1.5$. ϕ and θ are both set to be 0.4, while the fractional differencing parameter d is set as $-0.2, 0.2, 0.3, 0.4, 0.49$, such that a larger range of d is covered compared to Tieslau et al. (1996), who only consider the range of d from -0.5 to 0.25. The sample sizes considered here are $100, 200, 300$, and 500.

For each of the above DGPs, we firstly obtain the estimates of d by GPH, MLE, ELW, GMD and FEAR respectively, in all of which AIC is used to select the orders of models. We compute and compare the estimation bias (BIAS) and the root mean squared error (RMSE) of these estimators. The simulation results are reported in Tables 1-3. Generally it is shown in all these tables that the RMSEs produced by the FEAR estimation decrease with the increase of the sample size T regardless of the value of d, which indicates the FEAR estimation achieves consistency for all of the three different long memory processes. It is noteworthy that the consistency of FEAR estimation is even achieved for high level of fractional differencing parameter, therefore this outcome completes the shortage of Tieslau et al. (1996) whose method can only cover the range of $d \in (-0.5, 0.25)$.

Table 1 reports the simulation results for estimating fractional white noise. It can be seen in most of cases, FEAR estimator has relatively small RMSEs, which highlights the estimator is relatively efficient. For instance, when $d = -0.2$ and $d = 0.2$, RMSEs of GPH, ELW and GMD are all larger than that of FEAR for all choices of T. Compared to that, although GPH estimator is better at generating smaller biases for different combinations of d and T, its RMSEs are too large to be sufficiently efficient which indicates the estimator's variance is of high level. Table 2 presents the simulation results for estimating special ARFIMA model. It shows FEAR estimation performs best when considering both bias and RMSE, especially when d ranges from 0.2 to 0.4. Again it is observed that GPH estimation generally produces the smaller biases, but it is dominated by FEAR estimation in terms of small RMSE. The other alternative methods generate larger absolute values of bias and larger RMSE than FEAR. Table 3 reports the simulation results for estimating general ARFIMA model. It can be seen that when d ranges from -0.2 to 0.3, FEAR estimation performs relatively well in terms of both bias and RMSE. It is also observed that for other values of d, FEAR estimation still has comparatively advantages over GPH, MLE and GMD in producing smaller level of bias and RMSE.

To summarize, in all cases GPH estimation generally has small bias, but meanwhile causes large RMSEs. MLE estimation has relatively good finite sample performance when estimating ARFIMA $(0, d, 0)$, but it produces large bias and RMSEs when estimating ARFIMA $(1, d, 0)$ and ARFIMA $(1, d, 1)$. Similarly, ELW method has bad performance when estimating ARFIMA $(1, d, 0)$. GMD method also produces large RMSEs when estimating ARFIMA $(1, d, 0)$ and ARFIMA $(1, d, 1)$. Compared to that, simulation outcomes of FEAR estimation keep relatively small RMSEs in all cases, es-

Table 1: Performance of various estimation methods for ARFIMA (0, d, 0)

		GPH		MLE		ELW		GMD		FEAR	
d	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
-0.2	100	0.0116	0.2953	-0.0371	0.0998	0.0255	0.1339	-0.0321	0.1158	0.0051	0.0912
	200	-0.0090	0.2241	-0.0238	0.0601	0.0065	0.0970	-0.0195	0.0679	0.0053	0.0637
	300	-0.0141	0.2020	-0.0150	0.0510	0.0032	0.0814	-0.0120	0.0557	0.0087	0.0548
	500	0.0021	0.1703	-0.0120	0.0380	0.0006	0.0641	-0.0104	0.0402	0.0056	0.0416
0.2	100	0.0030	0.2776	-0.0428	0.1017	0.0236	0.1313	-0.0270	0.1098	-0.0595	0.0921
	200	0.0036	0.2336	-0.0220	0.0649	0.0131	0.0945	-0.0166	0.0702	-0.0394	0.0626
	300	0.0010	0.2102	-0.0144	0.0500	0.0101	0.0786	-0.0111	0.0522	-0.0322	0.0505
	500	-0.0039	0.1775	-0.0094	0.0389	0.0071	0.0637	-0.0073	0.0409	-0.0251	0.0417
0.3	100	0.0148	0.2925	-0.0409	0.1011	0.0353	0.1337	-0.0289	0.1081	-0.0928	0.1130
	200	0.0067	0.2346	-0.0204	0.0649	0.0104	0.0970	-0.0146	0.0719	-0.0713	0.0843
	300	0.0167	0.1992	-0.0124	0.0517	0.0099	0.0813	-0.0106	0.0554	-0.0601	0.0707
	500	0.0060	0.1727	-0.0109	0.0382	0.0043	0.0648	-0.0095	0.0403	-0.0536	0.0608
0.4	100	0.0121	0.2973	-0.0420	0.1012	0.0287	0.1287	-0.0301	0.1100	-0.1405	0.1511
	200	-0.0059	0.2358	-0.0204	0.0663	0.0141	0.0982	-0.0155	0.0727	-0.1155	0.1218
	300	0.0250	0.2033	-0.0118	0.0486	0.0136	0.0790	-0.0096	0.0506	-0.1019	0.1069
	500	0.0041	0.1751	-0.0098	0.0367	0.0053	0.0659	-0.0079	0.0390	-0.0942	0.0977
0.49	100	0.0148	0.2975	-0.0355	0.1027	0.0304	0.1335	-0.0256	0.1232	-0.1912	0.1980
	200	0.0211	0.2433	-0.0197	0.0641	0.0131	0.0924	-0.0159	0.0699	-0.1648	0.1684
	300	0.0315	0.2069	-0.0115	0.0497	0.0117	0.0772	-0.0099	0.0534	-0.1523	0.1549
	500	0.0154	0.1732	-0.0053	0.0374	0.0099	0.0664	-0.0038	0.0403	-0.1417	0.1434

Note: The simulation results are for DGP (4.1).

pecially when ARFIMA $(1, d, 1)$ is estimated, which most possibly reflect the realistic economic data among the three long memory processes considered in the simulations. It also indicates the variance of FEAR estimation is relatively small, which justifies the stability of the method. Overall comprehensively considering all the cases, we find that FEAR estimation is generally competitive to the four alternatives estimation methods.

4.2 Simulations for Aggregations of Cross-sectional and Time Dependent Series

Firstly, we generate a multivariate system of time series as following which will be used for aggregation,

$$y_{i,t} = \varepsilon_{it}, \tag{4.4}$$

$$(1 - 0.5L)\, y_{it} = (1 + 0.3L)\, \varepsilon_{i,t}, \tag{4.5}$$

Table 2: Performance of various estimation methods for ARFIMA (1, d, 0)

		GPH		MLE		ELW		GMD		FEAR	
d	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
-0.2	100	0.0711	0.3010	-0.2212	0.3276	0.2601	0.2909	-0.0399	0.2851	0.2891	0.2952
	200	0.0403	0.2352	-0.1597	0.2658	0.1714	0.1972	-0.0721	0.2301	0.2934	0.2961
	300	0.0342	0.1994	-0.1179	0.2162	0.1439	0.1639	-0.0576	0.1840	0.2968	0.2988
	500	0.0174	0.1667	-0.0719	0.1549	0.1078	0.1264	-0.0476	0.1365	0.2982	0.2993
0.2	100	0.0804	0.2914	-0.2641	0.3680	0.2620	0.2952	-0.0577	0.2782	0.0726	0.0835
	200	0.0336	0.2309	-0.1861	0.2899	0.1766	0.2023	-0.0703	0.2175	0.0862	0.0913
	300	0.0271	0.2074	-0.1428	0.2521	0.1473	0.1686	-0.0725	0.2014	0.0919	0.0950
	500	0.0236	0.1731	-0.0883	0.1836	0.1089	0.1268	-0.0496	0.1464	0.0957	0.0975
0.3	100	0.0656	0.3100	-0.2523	0.3664	0.2607	0.2910	-0.0591	0.2790	0.0060	0.0390
	200	0.0353	0.2305	-0.1927	0.2967	0.1741	0.1989	-0.0683	0.2132	0.0207	0.0332
	300	0.0270	0.2026	-0.1429	0.2488	0.1449	0.1651	-0.0630	0.1966	0.0272	0.0345
	500	0.0245	0.1711	-0.0859	0.1763	0.1054	0.1250	-0.0521	0.1504	0.0317	0.0364
0.4	100	0.0808	0.3048	-0.2513	0.3638	0.2606	0.2915	-0.0719	0.2663	-0.0660	0.0741
	200	0.0389	0.2422	-0.1900	0.3024	0.1756	0.1994	-0.0844	0.2232	-0.0497	0.0551
	300	0.0329	0.1995	-0.1403	0.2438	0.1409	0.1605	-0.0729	0.1876	-0.0434	0.0474
	500	0.0236	0.1715	-0.0836	0.1812	0.1122	0.1304	-0.0503	0.1523	-0.0363	0.0394
0.49	100	0.0903	0.3059	-0.2412	0.3538	0.2632	0.2960	-0.0647	0.2571	-0.1346	0.1380
	200	0.0497	0.2334	-0.1740	0.2887	0.1802	0.2046	-0.0691	0.2156	-0.1179	0.1197
	300	0.0475	0.2080	-0.1294	0.2325	0.1450	0.1627	-0.0746	0.1894	-0.1105	0.1117
	500	0.0377	0.1722	-0.0787	0.1746	0.1101	0.1274	-0.0458	0.1435	-0.1043	0.1051

Note: The simulation results are for DGP (4.2).

Table 3: Performance of various estimation methods for ARFIMA (1, d, 1)

		GPH		MLE		ELW		GMD		FEAR	
d	T	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
-0.2	100	0.0131	0.2841	-0.2204	0.4159	0.0162	0.1293	0.0463	0.4632	0.0015	0.0890
	200	0.0163	0.2262	-0.1191	0.3228	0.0137	0.0945	0.0782	0.3878	0.0054	0.0671
	300	0.0068	0.2049	-0.0963	0.2587	0.0038	0.0788	0.0838	0.3731	0.0072	0.0528
	500	-0.0003	0.1693	-0.0483	0.1806	-0.0006	0.0675	0.0657	0.3190	0.0074	0.0430
0.2	100	-0.0002	0.2877	-0.3504	0.5552	0.0320	0.1373	-0.0893	0.3927	-0.0583	0.0925
	200	0.0143	0.2313	-0.2132	0.4181	0.0131	0.0945	-0.0488	0.2989	-0.0373	0.0603
	300	0.0027	0.2083	-0.1343	0.3179	0.0102	0.0819	-0.0194	0.2319	-0.0312	0.0516
	500	-0.0006	0.1669	-0.0883	0.2235	0.0045	0.0645	-0.0176	0.1729	-0.0259	0.0408
0.3	100	0.0130	0.2959	-0.3548	0.5687	0.0225	0.1337	-0.1109	0.3919	-0.0958	0.1159
	200	0.0050	0.2378	-0.2244	0.4223	0.0112	0.0986	-0.0505	0.2735	-0.0704	0.0831
	300	0.0070	0.2008	-0.1542	0.3474	0.0108	0.0790	-0.0265	0.2391	-0.0603	0.0707
	500	0.0033	0.1741	-0.0923	0.2331	0.0055	0.0643	-0.0243	0.1684	-0.0529	0.0606
0.4	100	0.0167	0.2892	-0.3701	0.5863	0.0282	0.1325	-0.0976	0.3669	-0.1398	0.1511
	200	0.0130	0.2329	-0.2274	0.4370	0.0194	0.0970	-0.0416	0.2835	-0.1128	0.1193
	300	0.0156	0.2061	-0.1362	0.3210	0.0111	0.0801	-0.0373	0.2119	-0.1042	0.1094
	500	0.0090	0.1791	-0.0903	0.2432	0.0050	0.0634	-0.0227	0.1655	-0.0932	0.0969
0.49	100	0.0129	0.3036	-0.3948	0.6065	0.0351	0.1400	-0.1258	0.3905	-0.1901	0.1972
	200	0.0243	0.2391	-0.2261	0.4369	0.0173	0.0956	-0.0642	0.2795	-0.1635	0.1671
	300	0.0178	0.2025	-0.1543	0.3453	0.0112	0.0789	-0.0310	0.2192	-0.1533	0.1560
	500	0.0195	0.1719	-0.0778	0.2186	0.0077	0.0634	-0.0222	0.1509	-0.1415	0.1433

Note: The simulation results are for DGP (4.3).

for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$, where the error term $\varepsilon_{i,t}$ needs to be serially and cross-sectionally correlated. To that end, we generate

$$e = (e_1, e_2, \ldots, e_N) = \begin{pmatrix} e_{1,1} & e_{2,1} & \cdots & e_{N,1} \\ e_{2,2} & e_{2,2} & \cdots & e_{N,2} \\ \vdots & \vdots & \vdots & \vdots \\ e_{1,N} & e_{2,N} & \cdots & e_{N,N} \\ e_{1,N+1} & e_{2,N+1} & \cdots & e_{N,N+1} \\ \vdots & \vdots & \vdots & \vdots \\ e_{1,N+T-1} & e_{2,N+T-1} & \cdots & e_{N,N+T-1} \end{pmatrix}_{(N+T-1)\times N}$$

where

$$e_t = (e_{1,t}, e_{2,t}, \ldots, e_{N,t})' \sim \mathcal{N}(0, \Sigma)$$

with

$$\Sigma = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}.$$

Then the $T \times N$ matrix of error terms in (4.4) and (4.5) is generated as

$$\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N) = \begin{pmatrix} \varepsilon_{11} & \varepsilon_{21} & \cdots & \varepsilon_{N1} \\ \varepsilon_{12} & \varepsilon_{22} & \cdots & \varepsilon_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{1T} & \varepsilon_{2T} & \cdots & \varepsilon_{NT} \end{pmatrix}_{T\times N} = \begin{pmatrix} e_{1,N} & e_{2,N-1} & \cdots & e_{N,1} \\ e_{1,N+1} & e_{2,N} & \cdots & e_{N,2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{1,N+T-1} & e_{2,N+T-2} & \cdots & e_{N,T} \end{pmatrix}_{T\times N}$$

where $\varepsilon_i = (e_{i,N+1-i}, , \ldots, e_{i,N+T-i})'$. By this design, it follows

$$E(\varepsilon_{i,t}\varepsilon_{i+j,t+j}) = \rho$$

for $i = 1, 2, \ldots, N, j = 1, 2, \ldots, N - i$ and (4.4) and (4.5) imply y_{it} is also serially and cross-sectionally correlated. We consider $\rho = 0.1, 0.5, 0.8$ which correspond to weak, semi-strong and strong serial and cross-sectional correlation. The sample size is set as combinations of $N = 10, 15, 20, 30$ and $T = 100, 300, 500$.

Now we examine potential long memory properties of the aggregated series $S_t = \sum_{i=1}^{N} y_{it}$. In particular, we apply the commonly used GPH and our newly built FEAR estimator to estimate the fractional differencing parameter of S_t for each iteration. Then we calculate the average of the estimated fractional differencing parameter over iterations, which are reported in Tables 4 and 5. It is shown that as N and T are sufficiently large, the larger the value ρ is, the higher is the average of the estimated integrated order. This finding is consistent regardless of whether GPH or FEAR is used for estimation. It therefore provides strong evidences that the aggregation procedure generates long memory process, of which the persistence can be strengthened as the cross-sectional and serial correlations between the constitutes of the aggregated time series are increased. The simulation results also shed light on portfolio analysis. As assets forming portfolio are often cross-sectionally and serially correlated to certain degrees, it is reasonable to analyze the aggregated time series as a long memory process.

Table 4: Estimated fractional differencing parameter using GPH

ARMA(0,0) model

	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.8$		
N\T	100	300	500	100	300	500	100	300	500
10	0.286	0.174	0.147	0.543	0.217	0.156	0.724	0.225	0.161
15	0.303	0.219	0.173	0.616	0.411	0.245	0.789	0.482	0.265
20	0.323	0.270	0.226	0.645	0.614	0.431	0.849	0.784	0.493
30	0.313	0.317	0.307	0.652	0.646	0.655	0.868	0.765	0.791

ARMA(1,1) model

	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.8$		
N\T	100	300	500	100	300	500	100	300	500
10	0.358	0.199	0.149	0.636	0.248	0.168	0.830	0.263	0.170
15	0.385	0.262	0.180	0.723	0.448	0.270	0.894	0.516	0.286
20	0.408	0.308	0.244	0.767	0.674	0.449	0.970	0.809	0.522
30	0.395	0.357	0.341	0.762	0.687	0.673	0.963	0.802	0.814

Note: The simulation results in the top panel are for DGP (4.4) while that in the bottom panel are for DGP (4.5). ρ is the correlation coefficient used to generate errors in the two DGPs.

4.3 Simulations for the On-line Common Break Test

The DGP considered for the simulations of the on-line common break test is as following,

$$\begin{aligned} y_{it} &= e_{it}, \quad t = 1, 2, \ldots, k_0 \\ y_{it} &= \mu + e_{it}, \quad t = k_0 + 1, \ldots, T \\ i &= 1, 2, \ldots, N \end{aligned}$$

Table 5: Estimated fractional differencing parameter using FEAR

ARMA(0,0) model

	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.8$		
N\T	100	300	500	100	300	500	100	300	500
10	0.157	0.123	0.123	0.346	0.338	0.326	0.300	0.117	0.084
15	0.167	0.155	0.158	0.386	0.396	0.397	0.368	0.212	0.177
20	0.179	0.185	0.190	0.399	0.431	0.435	0.403	0.291	0.257
30	0.201	0.239	0.258	0.405	0.461	0.461	0.414	0.391	0.393

ARMA(1,1) model

	$\rho = 0.1$			$\rho = 0.5$			$\rho = 0.8$		
N\T	100	300	500	100	300	500	100	300	500
10	0.256	0.217	0.188	0.251	0.139	0.097	0.292	0.213	0.201
15	0.276	0.249	0.221	0.340	0.265	0.242	0.333	0.159	0.124
20	0.285	0.257	0.246	0.393	0.378	0.371	0.358	0.202	0.130
30	0.299	0.288	0.277	0.431	0.466	0.475	0.376	0.327	0.297

Note: The simulation results in the top panel are for DGP (4.4) while that in the bottom panel are for DGP (4.5). ρ is the correlation coefficient used to generate errors in the two DGPs.

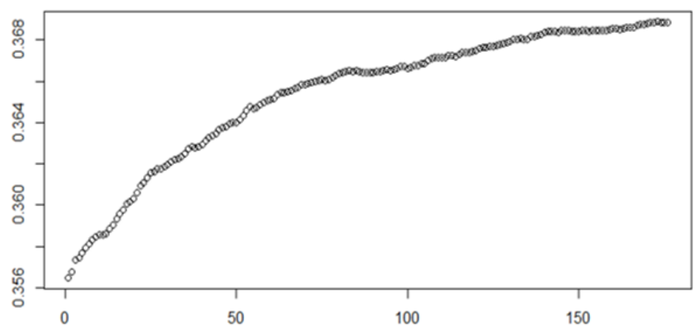
where the error are generated as

$$e_{it} = 0.8e_{it-1} + \varepsilon_{it} + 0.7\varepsilon_{it-1}$$

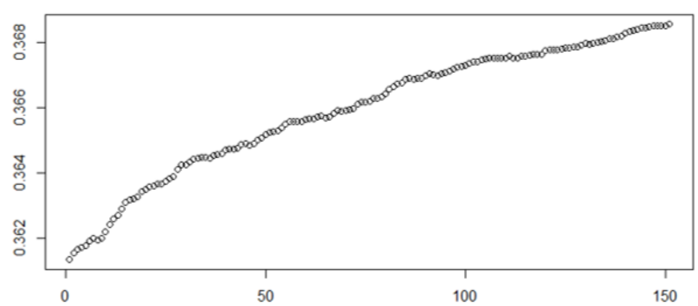
with $\varepsilon_{it} \sim i.i.d.N(0,1)$ for all i and t. As shown before, the on-line common break test is then based on the aggregated time series $S_t = \sum_{i=1}^{n} y_{it}$, where a recursive window scheme is applied by setting $W = 100$.

Firstly, we examine the case where there is no break by setting $\mu = 0$. For each iteration, we apply the FEAR estimator to S_t via recursive window scheme and obtain 151 estimates of fractional differencing parameter. Then we compute the average of the values of these estimates over iterations. The results for $N = 30, 50$ and $T = 250$ are presented in Figure 1, where the horizontal axis is the number of additional observations compared to the first 100 observations and each plotted point is based the average of 1000 replications of estimated fractional differencing parameter. It is shown in Figure 1 that the estimated fractional differencing parameter rises as recursive window size increases. Also the estimated value is larger when $N = 50$ than that when $N = 30$. It therefore indicates when there are more observations used to aggregate, the fractional parameter d value becomes larger. This finding is consistent with Granger (1980).

Now we consider the case where there exists one break by setting $\mu = 0.5$, $N = 30, 50$, $T = 250$ and $k_0 = 151$. Similar to the case with no break, Figure 2 presents the graph of average estimated fractional differencing parameters over iterations, which are computed via recursive window scheme. It is noteworthy that the horizontal axis of true break location is 51 since the beginning window size is 100. There are several interesting findings from these plots in Figure 2. Firstly, a sudden change of the estimated parameter can be observed at the true break location $k_0 = 151$ regardless of the value N. Secondly, comparing the two subplots, we can intuitively find that the change magnitude of estimates becomes larger when N is larger. It therefore implies that the detection of common break in multivariate system might be more precise and more obvious when N is larger, which could be explained by the reason that the summation of structural break in univariate series can magnify the common break in multivariate system.



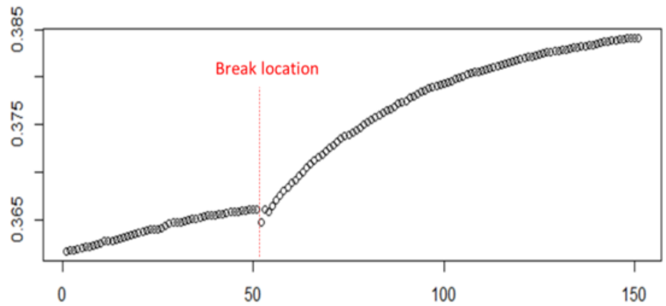
(a) Recursive d estimation (N = 30)



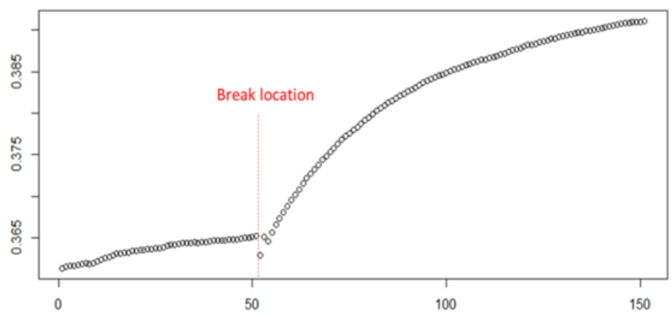
(b) Recursive d estimation (N = 50)

Figure 1: Aggregation without break

Then we consider applying the on-line common break test to S_t to detect the break point. To that end, the combinations of $N = 5, 10, 20, 30, 50$ and $T = 100, 250$ are considered for the sample size and the break location k_0 is set to 75 and 150 for different T respectively. The size and power of the test are shown in Table 6, which reports rejection rates of the on-line common break test at nominal 5% significance level, and the values in size row are computed by letting $\mu = 0$ while that in the power row are computed by letting $\mu = 0.5$. In particular, the rejection rates in size rows are all smaller than 5% no matter what values N and T are. It indicates the test has small Type I error for different sample



(a) Recursive d estimation (N = 30)



(b) Recursive d estimation (N = 50)

Figure 2: Aggregation with break

sizes. In addition to that, the rejection rates in power rows are sufficiently large, implying the test has satisfactory power. Moreover, it is observed that the rejection rates rise significantly when the value of N is increased. This finding implies the on-line common break test is more powerful when larger number of time series are aggregated.

Table 6: Rejection rates of the on-line common break test at 5% significance level

N		5	10	20	30	50
$T = 100$	size	3.06%	3.16%	3.09%	3.14%	3.01%
	power	44.90%	47.70%	56.20%	64.60%	70.60%
$T = 250$	size	2.24%	2.25%	2.30%	2.35%	2.35%
	power	41.90%	47.40%	56.50%	70.00%	94.30%

5 Empirical Application

Understanding the optimal timing for adjusting a portfolio comprising financial assets is crucial for practitioners, especially in the presence of unexpected common shocks or systemic risks. Therefore, a common break test for detecting changes in the time series properties of portfolio returns becomes essential, as it can mitigate risks stemming from market fluctuations. Financial assets are often serially and cross-sectionally dependent, attributed to common shock effects, sequential technological innovations, increasing globalization, and spillover effects among other factors. Based on the theoretical discussions and Monte Carlo simulation studies outlined earlier, it is evident that the return aggregation of assets in a portfolio is likely to follow a long memory process. Consequently, the FEAR-predictive test, designed to identify structural breaks in the fractional differencing parameter, emerges as a promising and effective on-line tool for signaling the timing of portfolio adjustments. To underscore the utility and effectiveness of this FEAR-predictive test, we present an empirical application that tests for impending collapses and crises by detecting changes in the long memory parameters of the aggregation of various stock market indices. In essence, our aim is to identify opportune times to adjust portfolio allocations using the FEAR-predictive test.

The data under consideration is sourced from Bloomberg and encompasses the daily returns of stock market indices for 19 countries, spanning from the beginning of 2000 to May 2022. These countries comprise a mix of developed nations (US, UK, Japan, Switzerland, Europe, Norway, Sweden, South Korea) and developing nations (South Africa, China, Turkey, Thailand, Saudi Arabia, Philippines, India, Mexico, Chile, Brazil, Malaysia). We analyze three scenarios of aggregated series. Scenario 1 involves aggregating the stock returns of 8 developed countries over the entire duration, while Scenario 2 entails summing up the stock returns of 11 developing countries. These two time series allow us to discern distinct patterns between developed and developing countries. To provide a more comprehensive analysis of globalization, we further construct Scenario 3, aggregating stock returns for all 19 countries across the entire sample period. In brief, these three scenarios can be interpreted as three distinct portfolio settings. The first two scenarios focus on stock indices of developed and developing countries, respectively, while the third setting encompasses all 19 countries, facilitating a holistic perspective on global market dynamics.

For each aggregated time series, we employ the FEAR-predictive test to recursively identify changing patterns in the long memory parameter d. This process initiates with a training sample size of 100. Without the loss of generality, we anticipate detecting breaks when the values of the estimated d shift, even if the shift is subtle. The Figures 3-5 present the recursively FEAR-estimated d values and the identified break dates for the three scenarios. On these figures, the horizontal axis represents time, and the vertical axis denotes the estimated d values. Additionally, we mark the detected dates and provide indications of their corresponding influential economic events.

Notably, significant crises or crashes in the market align closely with the detected dates, underscoring the practical feasibility of the FEAR-predictive test. Several compelling findings emerge from the analysis. Figure 3 illustrates six detected breaks for Scenario 3, with the line representing the recursive FEAR-estimates. The first break in May 2006 anticipates the subsequent eruption of the 2007 subprime crisis. The second break in January 2008 precedes the bankruptcy of Lehman Brothers and the onset of the global financial crisis. The third break in May 2010 aligns with the outbreak of the Greek debt crisis, marking the inception of the European debt crisis. The fourth break around August 2011 foreshadows the diffusion of the 2012 European debt crisis. The fifth break in August 2015 coincides with the Chinese stock crash and subsequent stock market crashes. The most recent break detected in early March 2020 correlates with the global spread of the 2020 Covid-19 pandemic. The identified patterns strongly imply that changes in d values are not only informative but also closely linked to substantial fluctuations in the fundamental economic system. This underscores the practical utility and effectiveness of the FEAR procedure in capturing and signaling these important shifts in market dynamics.

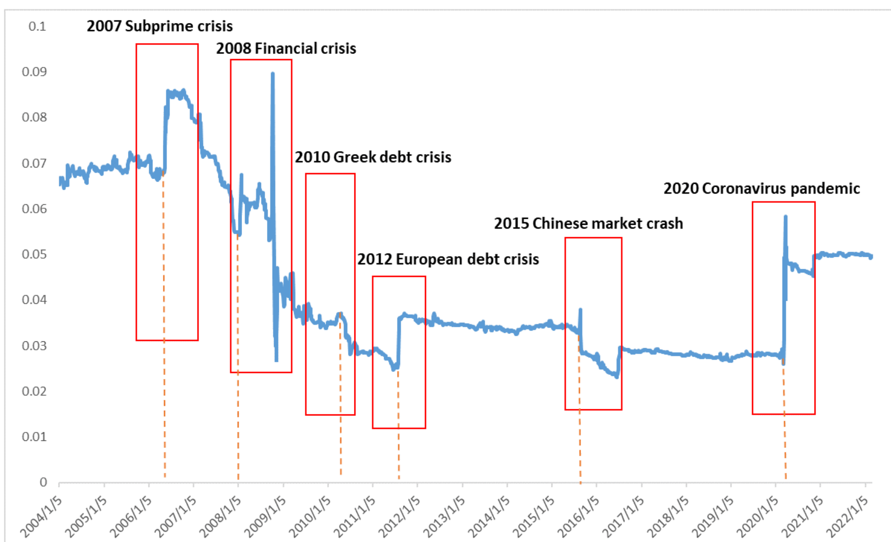


Figure 3: Recursive FEAR-d estimates for aggregated stock market index return from both developed and developing countries

Figures 4 and 5 depict the recursive estimates for Scenarios 1 and 2, respectively. We find that the breaks detected by the FEAR-predictive test for these two scenarios closely mirror those found for Scenario 3 in Figure 3, with the exception of the 2010 Greek debt crisis and the 2012 European debt crisis in Figure 5. An explanation for this phenomenon is that the Greek debt crisis and the European debt crisis are comparatively regional, and most of the considered developing countries are not located in Europe. Specifically, six breaks are detected for Scenario 1, with break dates in May 2006, September 2008, May 2010, September 2011, September 2015, and March 2020. Meanwhile, four detected dates for Scenario 2 are in May 2006, January 2008, August 2015, and March 2020, but two of them are earlier compared to those for Scenario 1. This suggests that stock markets of developing countries may be more sensitive to these crises. Additionally, it is worth highlighting that the magnitudes of most breaks for Scenario 2 are often larger than those for Scenario 1.

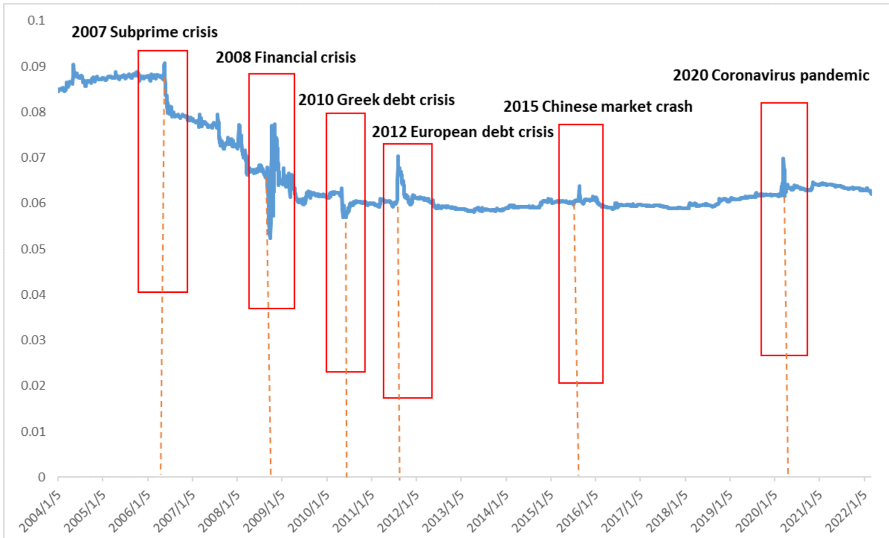


Figure 4: Recursive FEAR-d estimates for aggregated stock market index return from developed countries

In summary, the majority of detected breaks for Scenarios 1-3 precede the occurrence or widespread impact of significant global events. The recursive FEAR-estimates of d exhibit remarkable stability in periods devoid of common events. However, notable jumps or downward shifts occur promptly when the market becomes more volatile or experiences common shocks. This finding underscores the stability and accuracy of the FEAR estimation. Collectively, these results affirm the efficacy of the FEAR-predictive test as a reliable early warning measure for signaling impending financial crises and systemic risks. Such insights can be instrumental in guiding portfolio adjustments and minimizing risks.

6 Conclusion

In this study, we introduce the FEAR estimation designed for estimating the fractional differencing parameter d in a stationary long-memory process. This is particularly valuable for portfolio adjustment, given that most aggregated return series exhibit characteristics of long-memory time series. Theoretical foundations are provided to demonstrate the consistency and asymptotic distribution of the FEAR estimator. This enables the construction of a FEAR-predictive test, aiding in the detection of changes in the time series properties of portfolio returns and offering valuable insights for adjusting a portfolio

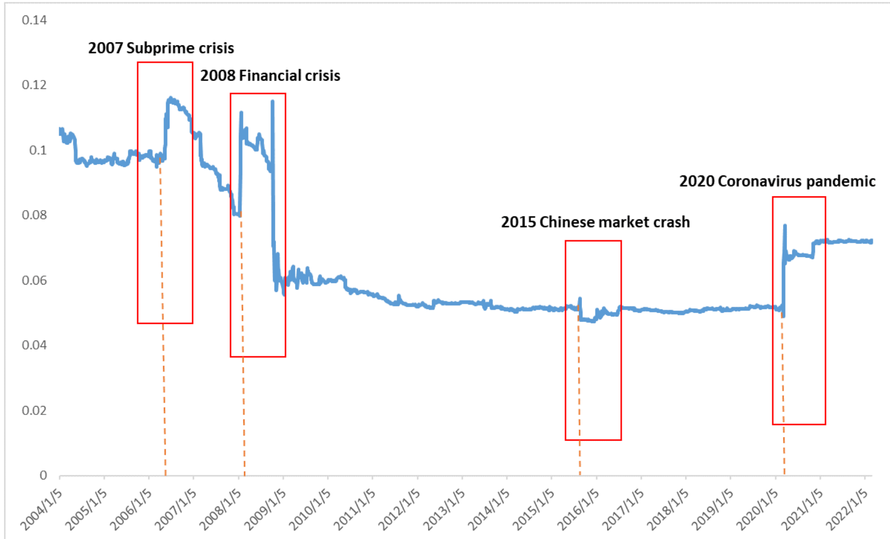


Figure 5: Recursive FEAR-d estimates for aggregated stock market index return from developing countries

based on market dynamics. Both simulation experiments and empirical studies affirm the practicality and robustness of the FEAR estimation and the proposed on-line test.

A Technical Appendix

Proof of Theorem 2.1. The FEAR estimator solves the first-order condition $\partial S_T(d)/\partial \hat{d} = 0$. By the Taylor expansion, we write:

$$\frac{\partial S_T(d)}{\partial \hat{d}} = \frac{\partial S_T(d)}{\partial d} + \frac{\partial^2 S_T(d)}{\partial \bar{d}^2}(\hat{d} - d) = 0,$$

where $\bar{d}$ lies between d and $\hat{d}$. Then we solve for $(\hat{d} - d)$ and get:

$$(\hat{d} - d) = -\Big[\frac{\partial^2 S_T(d)}{\partial \bar{d}^2}\Big]^{-1}\frac{\partial S_T(d)}{\partial d} = -\Big[\frac{1}{2T\hat{\rho}_0}\frac{\partial^2 S_T(d)}{\partial \bar{d}^2}\Big]^{-1}\frac{1}{2T\hat{\rho}_0}\frac{\partial S_T(d)}{\partial d}, \tag{A.1}$$

where $\hat{\rho}_0 \equiv (1/T)\sum_{t=1}^T y_t^2$ denotes the sample variance of y_t.

By Equation (2.12),

$$\begin{aligned} \frac{1}{2T\hat{\rho}_0}\frac{\partial S_T(d)}{\partial d} &= \sum_{j=1}^k\sum_{i=1}^k\Big(\tilde{\beta}_j - \Delta_j(d)\Big)\frac{\partial \Delta_i(d)}{\partial d}\frac{\sum_{t=1}^T y_{t-j}y_{t-i}}{\sum_{t=1}^T y_t^2} \\ &= \sum_{j=1}^k\sum_{i=1}^k\Big(\tilde{\beta}_j - \Delta_j(d)\Big)\frac{\partial \Delta_i(d)}{\partial d}\hat{\rho}_{j-i}, \end{aligned} \tag{A.2}$$

where $\hat{\rho}_{j-i}$ is the sample autocorrelation of y_t. The derivative of $\Delta_j(d)$ with respect to d is:

$$\frac{\partial \Delta_i(d)}{\partial d} = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}\psi^{di}(j+d) - \frac{\Gamma(j+d)}{(\Gamma(d))^2\Gamma(j+1)}\psi^{di}(d).$$

and $\psi^{di}(z)$ is the digamma function with $\psi^{di}(z) = \Gamma'(z)/\Gamma(z)$.

As is shown in Theorem 5 of Poskitt (2007), if y_t is a stationary process that satisfies Assumption 1 and Assumption 2, $\tilde{\beta}_j$ is a consistent estimate for β_j, we know that $\left(\tilde{\beta}_j - \Delta_j(d)\right)$ is $o_p(1)$. In addition, because the sample autocorrelation $\hat{\rho}$ is consistent and asymptotically follows a normal distribution as is provided by Hosking (1984, 1996), we obtain $\hat{\rho} = O_p(1/\sqrt{T})$, and thus

$$\frac{1}{2T\hat{\rho}_0}\frac{\partial S_T(d)}{\partial d} = o_p(1). \tag{A.3}$$

Differentiating the objective function $S_T(d)$ twice yields:

$$\begin{aligned}
\frac{\partial^2 S_T(d)}{\partial d^2} &= \frac{\partial}{\partial d}\left[2\sum_{t=1}^{T}\left(v_t(d) - \tilde{v}_t\right)v_t'(d)\right] \\
&= 2\sum_{t=1}^{T}\left[v_t'(d)^2 + v_t(d)v_t''(d) - \tilde{v}_t v_t''(d)\right] \\
&= 2\sum_{t=1}^{T} v_t'(d)^2 + 2\sum_{t=1}^{T}\left[v_t(d) - \tilde{v}_t\right]v_t''(d)
\end{aligned}$$

Since $v_t(d) = y_t - \sum_{j=1}^{k}\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}y_{t-j} = y_t - \sum_{j=1}^{k}\Delta_j(d)y_{t-j}$, we have

$$v_t'(d) = -\sum_{j=1}^{k}\frac{\partial \Delta_j(d)}{\partial d}y_{t-j}, \quad v_t''(d) = -\sum_{j=1}^{k}\frac{\partial^2 \Delta_j(d)}{\partial d^2}y_{t-j}.$$

The second derivative of $S_T(d)$ multiplied by $\frac{1}{2T\hat{\rho}_0}$ can be rewritten as:

$$\begin{aligned}
\frac{1}{2T\hat{\rho}_0}\frac{\partial^2 S_T(d)}{\partial d^2} &= \frac{1}{T\hat{\rho}_0}\sum_{t=1}^{T}\left[-\sum_{j=1}^{k}\frac{\partial \Delta_j(d)}{\partial d}y_{t-j}\right]^2 - \frac{1}{T\hat{\rho}_0}\sum_{t=1}^{T}\left[v_t(d) - \tilde{v}_t\right]\left[\sum_{j=1}^{k}\frac{\partial^2 \Delta_j(d)}{\partial d^2}y_{t-j}\right] \\
&= \sum_{j=1}^{k}\sum_{i=1}^{k}\left[\frac{\partial \Delta_j(d)}{\partial d}\frac{\partial \Delta_i(d)}{\partial d}\left(\frac{\sum_{t=1}^{T} y_{t-j}y_{t-i}}{\sum_{t=1}^{T} y_t^2}\right)\right] \\
&\quad -\frac{1}{T\hat{\rho}_0}\sum_{t=1}^{T}\left[\sum_{j=1}^{k}\left(\Delta_j(d) - \tilde{\beta}_j\right)y_{t-j}\right]\left[\sum_{j=1}^{k}\frac{\partial^2 \Delta_j(d)}{\partial d^2}y_{t-j}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{k}\sum_{i=1}^{k}\left[\frac{\partial\Delta_j(d)}{\partial d}\frac{\partial\Delta_i(d)}{\partial d}\left(\frac{\sum_{t=1}^{T} y_{t-j}y_{t-i}}{\sum_{t=1}^{T} y_t^2}\right)\right] \\
&\quad -\sum_{j=1}^{k}\sum_{i=1}^{k}\left[\left(\Delta_j(d)-\tilde{\beta}_j\right)\frac{\partial^2\Delta_i(d)}{\partial d^2}\left(\frac{\sum_{t=1}^{T} y_{t-j}y_{t-i}}{\sum_{t=1}^{T} y_t^2}\right)\right] \\
&= \sum_{j=1}^{k}\sum_{i=1}^{k}\left[\frac{\partial\Delta_j(d)}{\partial d}\frac{\partial\Delta_i(d)}{\partial d}\hat{\rho}_{j-i}\right] - \sum_{j=1}^{k}\sum_{i=1}^{k}\left[\left(\Delta_j(d)-\tilde{\beta}_j\right)\frac{\partial^2\Delta_i(d)}{\partial d^2}\hat{\rho}_{j-i}\right] \\
&= \sum_{j=1}^{k}\sum_{i=1}^{k}\left[\frac{\partial\Delta_j(d)}{\partial d}\frac{\partial\Delta_i(d)}{\partial d}\hat{\rho}_{j-i}\right] + o_p(1). \qquad \text{(A.4)}
\end{aligned}
$$

The second term in (A.4) is $o_p(1)$ because $\hat{\rho}$ is a consistent estimator of ρ as is shown by Hosking (1984, 1996). In addition, $\left(\tilde{\beta}_j - \Delta_j(d)\right)$ is $o_p(1)$ under Assumption 1 and Assumption 2 by Poskitt (2007). From Equation (A.1), the results in (A.3) and (A.4) imply that $(\hat{d} - d) = o_p(1)$.

■

Proof of Theorem 2.2. Under Assumption 1 and Assumption 2, the FEAR estimator $\hat{d}$ is consistent by Theorem 2.1. It follows that $\bar{d}$ in (A.1) converges to d in probability. Thus, by (A.4) we have

$$
\text{plim}\left(\frac{1}{2T\hat{\rho}_0}\frac{\partial^2 S_T(d)}{\partial \bar{d}^2}\right) = \sum_{j=1}^{k}\left[\frac{\partial\Delta_j(d)}{\partial d}\left(\sum_{i=1}^{k}\frac{\partial\Delta_i(d)}{\partial d}\rho_{j-i}\right)\right] = CW'. \qquad \text{(A.5)}
$$

Let $\beta = (\beta_1, \beta_2, ..., \beta_k)'$. Poskitt (2007) shows the asymptotic distribution of $\tilde{\beta}$:

$$
\sqrt{T}(\tilde{\beta} - \beta) \to_d \mathcal{N}(0, \sigma_\epsilon^2 \Gamma_k^{-1}), \qquad \text{(A.6)}
$$

where $\Gamma_k = [\gamma_{j-i}]_{j,i=1,...,k}$ and γ denotes the autocovariance function of y_t. Therefore, we have

$$
\sqrt{T}(\tilde{\beta}_j - \Delta_j(d)) \to_d \mathcal{N}(0, \sigma_\epsilon^2 [\Gamma_k^{-1}]_{jj}),
$$

where $[\Gamma_k^{-1}]_{jj}$ is the j-th diagonal term of Γ_k^{-1}. In (A.2),

$$
\sqrt{T}\left(\frac{1}{2T\hat{\rho}_0}\frac{\partial S_T(d)}{\partial d}\right) = \sum_{j=1}^{k}\left(\left[\sqrt{T}\left(\tilde{\beta}_j - \Delta_j(d)\right)\right]\left(\sum_{i=1}^{k}\frac{\partial\Delta_i(d)}{\partial d}\hat{\rho}_{j-i}\right)\right). \qquad \text{(A.7)}
$$

By the asymptotic distribution of $\tilde{\beta}$ in (A.6), the right-hand side of this equation thus is a linear function of random variables that are jointly normally distributed. Equation (A.7) can be expressed in matrix form as:

$$\sqrt{T}\left(\frac{1}{2T\hat{\rho}_0}\frac{\partial S_T(d)}{\partial d}\right)=\hat{W}\left[\sqrt{T}\left(\tilde{\beta}-\Delta(d)\right)\right], \tag{A.8}$$

where $\Delta(d)=(\Delta_1(d),\Delta_2(d),\cdots,\Delta_k(d))'$.

Given (A.6) and (A.8), it follows that

$$\sqrt{T}\left(\frac{1}{2T\hat{\rho}_0}\frac{\partial S_T(d)}{\partial d}\right)\to_d \mathcal{N}(0,\sigma_\epsilon^2 W\Gamma_k^{-1}W'). \tag{A.9}$$

Finally, using (A.1), (A.5) and (A.9), we obtain the desired result:

$$\sqrt{T}(\hat{d}-d)\to_d \mathcal{N}\left(0,\sigma_\epsilon^2\frac{W\Gamma_k^{-1}W'}{(CW')^2}\right). \tag{A.10}$$

■

Proof of Theorem 3.1. According to Theorems 2.1 and 2.2, Theorem 3.1 sustains immediately. ■

References

Beran, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(4):659–672.

Brockwell, P. J. and Davis, R. A. (2009). *Time series: theory and methods*. Springer science & business media.

Choi, K., Yu, W.-C., and Zivot, E. (2010). Long memory versus structural breaks in modeling and forecasting realized volatility. *Journal of International Money and Finance*, 29(5):857–875.

Crato, N. and Ray, B. K. (1996). Model selection and forecasting for long-range dependent processes. *Journal of Forecasting*, 15(2):107–125.

Dai, Z., Zhu, H., and Kang, J. (2021). New technical indicators and stock returns predictability. *International Review of Economics & Finance*, 71:127–142.

Diebold, F. X. and Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics*, 105(1):131–159.

Ferson, W. E., Sarkissian, S., and Simin, T. T. (2003). Spurious regressions in financial economics? *The Journal of Finance*, 58(4):1393–1413.

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4):221–238.

Granger, C. W. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14(2):227–238.

Granger, C. W. and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the s&p 500 absolute stock returns. *Journal of Empirical Finance*, 11(3):399–421.

Hosking, J. R. (1984). Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research*, 20(12):1898–1908.

Hosking, J. R. (1996). Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series. *Journal of Econometrics*, 73(1):261–284.

Joseph, L. and Wolfson, D. B. (1992). Estimation in multi-path change-point problems. *Communications in Statistics-Theory and Methods*, 21(4):897–913.

Lustig, H., Roussanov, N., and Verdelhan, A. (2014). Countercyclical currency risk premia. *Journal of Financial Economics*, 111(3):527–553.

Lütkepohl, H. (1989). Prediction tests for structural stability of multiple time series. *Journal of Business & Economic Statistics*, 7(1):129–135.

Lyle, M. R. and Wang, C. C. (2015). The cross section of expected holding period returns and their dynamics: A present value approach. *Journal of Financial Economics*, 116(3):505–525.

Mayoral, L. (2007). Minimum distance estimation of stationary and non-stationary arfima processes. *The Econometrics Journal*, 10(1):124–148.

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. *Econometric Reviews*, 34(6-10):1089–1117.

Poskitt, D. S. (2007). Autoregressive approximation in nonstandard situations: The fractionally integrated and non-invertible cases. *Annals of the Institute of Statistical Mathematics*, 59(4):697–725.

Ramchand, L. and Susmel, R. (1998). Volatility and cross correlation across major stock markets. *Journal of Empirical Finance*, 5(4):397–416.

Robinson, P. M. (1978). Statistical inference for a random coefficient autoregressive model. *Scandinavian Journal of Statistics*, pages 163–168.

Shimotsu, K. and Phillips, P. C. (2005). Exact local whittle estimation of fractional integration. *Annals of Statistics*, 33(4):1890–1933.

Solnik, B. and Roulet, J. (2000). Dispersion as cross-sectional correlation. *Financial Analysts Journal*, 56(1):54–61.

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53(1-3):165–188.

Tieslau, M. A., Schmidt, P., and Baillie, R. T. (1996). A minimum distance estimator for long-memory processes. *Journal of Econometrics*, 71(1-2):249–264.

Wang, C. S., Hsiao, C., and Yang, H.-H. (2021). Market integration, systemic risk and diagnostic tests in large mixed panels. *Econometric Reviews*, 40(8):750–795.

Wang, C. S. and Wan, S. K. (2020). A var approach to forecasting multivariate long memory processes subject to structural breaks. In *Essays in Honor of Cheng Hsiao, Advanced Econometrics*. Emerald Publishing Limited.

Zaffaroni, P. (2004). Contemporaneous aggregation of linear dynamic models in large economies. *Journal of Econometrics*, 120(1):75–102.