Fractional integration and cointegration in financial time series

Stakėnas, P.

Publication date
2012

Chapter 1

General introduction and overview

1.1 Fractional integration

The story of the central concept in this thesis - fractional integration - begins with Dr. Harold Edwin Hurst (1880-1978), a British hydrologist, who was commissioned by the Egyptian government to work on the projects regarding control and conservation of the water of the river Nile, who during his work collected a large amount of data (rainfall, river flow etc.) throughout the river basin. While investigating water discharge time series in relation to Aswan High Dam design, he was interested in a particular statistic, known as rescaled range statistic: R_n/S_n, where R_n is a range (the difference between the maximum and the minimum values) and S_n is a standard deviation of some time series data sample with size n. It was known at the time that for the class of processes for which the central limit theorem[1] holds, the statistic is proportional to the square root of the sample size: $R_n/S_n \sim n^{1/2}$. However, Hurst estimated that the empirical growth of the R_n/S_n statistic is proportional to n^H with estimated value $H = 0.74$, thus puzzling statisticians of the time, since the assumptions guaranteeing the functional central limit theorem were understood to be reasonably mild. This phenomenon - statistical discrepancy between the estimated value of H and 0.5 - latter was called *Hurst phenomenon.* The phenomenon later was also noticed in many other datasets related to climate, internet traffic, economics, finance, opinion polls, etc.

It turned out that constructing a stationary time series model exhibiting the Hurst phenomenon is not easy and some departures from the functional central limit theorem are required. The first idea was put forward in Moran (1964), where it was suggested to

[1]Which, without going into details, for the large class of processes could be thought of as being equivalent to the central limit theorem: $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right) \xrightarrow{d} N(0, \sigma^2)$, for a stationary process X_t with mean μ and variance σ^2.

relax the assumption of finite variance of the observed time series, implying the heavy tail phenomenon in the series. Another idea was suggested in Mandelbrot (1965), where a class of processes with finite variance, but slowly decaying autocorrelations was proposed. Temporal dependence, implied by slowly decaying autocorrelations, in such time series is too strong for the functional central limit theorem to hold. In other words, the time series could be said to have *long memory* or display *long range dependence* and a common way in the literature to define this, is absolute non-summability of its autocovariances: $\sum_h |\rho_X(h)| = \infty$, where $\rho_X(h)$ is autocovariance function of a stationary time series X_t. Naturally, time series with absolutely summable autocovariances are referred to as *weakly dependent*, *short-range dependent* or *short memory* time series.

The first parameterization of long memory time series was developed independently in Granger (1980), Granger and Joyeux (1980) and Hosking (1981):

$$X_t = \Delta^{-d}\varepsilon_t, \tag{1.1}$$

where $\Delta^{-d} = (1-L)^{-d} = \sum_{i=0}^{\infty} \frac{\Gamma(d+i)}{\Gamma(d)\Gamma(i+1)} L^i$, L is the lag operator, $\Gamma(i)$ is the Gamma function, ε_t is an i.i.d. zero-mean and finite variance time series and $-1/2 < d < 1/2$. Asymptotic properties of long memory processes, including those of type (1.1) were studied in Mandelbrot and Ness (1968), Davydov (1970), Avram and Taqqu (1987). It turned out that pre-sample observations of such time series have a non-negligible influence on the asymptotic distribution of partial sums of X_t, which stands in contrast to stationary short memory time series, where pre-sample observations have no effect on their first-order asymptotic properties. This observation led to an alternative definition of a long memory process:

$$X_t = \Delta_+^{-d}\varepsilon_t = \sum_{i=0}^{t} \frac{\Gamma(d+i)}{\Gamma(d)\Gamma(i+1)} \varepsilon_{t-i}, \tag{1.2}$$

where the subscript "$_+$" denotes truncation of the fractional filter Δ^{-d} at the time $t = 0$. Note, that for the special case $d = 1$ in (1.2), we have: $X_t = \Delta_+^{-1}\varepsilon_t = \sum_{i=1}^{t} \varepsilon_{t-i}$, i.e. X_t is a partial sum of ε_t's. This type of time series process in econometrics is well known as a *unit root* time series, a name deriving from the fact that this time series could be represented as a first-order autoregressive process whose lag polynomial has a root equal to one: $(1 - \phi L)X_t = \varepsilon_t$, $\phi = 1$. This time series X_t could be also regarded as an "integrated" series ε_t and, thinking in this way, allowing for an arbitrary value of d in (1.2) gives what was coined a *fractional unit root* or *fractionally integrated* time series, with the adjective "fractional" implying a not necessarily integer order of integration.

Marinucci and Robinson (1999) and Robinson (2005) reflected on the differences between the asymptotic properties of partial sums implied by representations (1.1) and (1.2) and introduced the terms *type I* and *type II* fractionally integrated processes for (1.1) and (1.2), respectively. The econometric literature is focused more on type II processes and in this thesis we also take this approach to fractional time series modeling (although in empirical applications some caution has to be maintained, since the choice implies different asymptotic inference procedures, see Davidson and Hashimzade (2009)) and adopt the following formal definition of fractional integration, used in Hualde (2009).

Definition 1.1. *We say that a univariate time series X_t is called integrated of order d, denoted as $X_t \sim I(d)$, if:*

$$X_t = \sum_{i=1}^{k} \Delta_+^{-d_i} u_{it}, \qquad (1.3)$$

where $d = \max_i d_i$ and the processes $u_{it}, i = 1, \dots, k$ are covariance stationary with a continuous, bounded and non-zero everywhere spectral density. We say that a multivariate time series X_t is integrated of order d if at least one of its components is $I(d)$, whereas the others have no greater integration order than d.

Admittedly, Definition 1.1 is restrictive, since it imposes $X_0 = \sum_{i=1}^{k} u_{i0}$ for the data generating process, what may not be a reasonable assumption for many observed datasets. However, later in the thesis we show that in our considered models more generally we may assume $X_0 = O_p(1)$ and also allow for inference conditional on the pre-sample values in a similar way as in Johansen and Nielsen (2012).

Note that Definition 1.1 implicitly defines $I(0)$ processes, which in some sources can be understood differently (see Davidson (2009) for a thorough discussion), but it reflects the common basic understanding that an $I(0)$ series is an (asymptotically) stationary time series for which the central limit theorem applies.

Research on fractionally integrated time series is rather voluminous and concerns a number of issues, such as modeling, representation and estimation of time series, but it is worth mentioning that the existence of long memory in data itself is not universally agreed upon: the apparent strong persistence phenomenon in time series may be explained by certain type of non-stationarities in the data, such as structural breaks, level shifts or regime-switching behaviour (see Granger and Hyung (2004), Smith (2005) and Diebold and Inoue (2001), correspondingly). However, fractional models are gaining ever more ground in empirical economic research.

In this section we have briefly discussed the concept of fractional time series, although an extensive review would take a chapter of itself and for more extensive mathematical and statistical treatment we refer to Beran (1994), Samorodnitsky (2007), Palma (2007) and Giraitis et al. (2012). An excellent, although somewhat outdated, review from an econometric perspective is Baillie (1996), while for more recent results we refer to Gil-Alana and Hualde (2009).

1.2 Fractional cointegration

We begin with a formal definition of cointegration, from Engle and Granger (1987):

Definition 1.2. *We say that a multivariate time series X_t is cointegrated, if $X_t \sim I(d)$, but there exists a non-zero vector β, such that $\beta' X_t \sim I(d-b)$ for some $b > 0$. The number r of such linearly independent vectors is called the cointegration rank of X_t.*

In the following, such a vector β is called the cointegration vector, while the time series $\beta' X_t$ is referred to as the cointegration errors. Definition 1.2 is rather general and covers all empirically relevant cases as well as some less interesting cases, such as "singular cointegration" between X_t and X_{t-1} if $X_t \sim I(1)$, or trivial cointegration in a multivariate system, where the integration order of one component is greater than the others. However, since we do not use Definition 1.2 formally and specify data generating processes in all chapters explicitly, we do not seek greater clarity (for comparison of different definitions, see Robinson and Yajima (2002)).

The cointegration mechanism could be interpreted as follows. Suppose that for a multivariate unit root time series X_t economic theory postulates that in equilibrium it holds: $\beta' X_t = 0$. Obviously, a zero deviations restriction is unreasonable and a more realistic assumption is that deviations from equilibrium form a stationary process. Hence, even though the observed time series X_t is an $I(1)$ process, there exists a stationary linear combination of X_t. Although the original definition of Engle and Granger (1987) restrict neither the observed time series nor the cointegration errors to be integrated of integer order, until the mid-90's the econometric literature was almost exclusively focused on the so called "classical cointegration" case, where the observed time series is assumed to be an $I(1)$ process, while the cointegration errors are $I(0)$ time series. One possible explanation is that, on the one hand, stationarity of time series at the time was associated with the Box-Jenkins modeling methodology with inherent short-memory properties of a series, while, on the other hand, lack of statistical evidence for existence of long memory in economic time series led researchers at the time to restrict themselves to the intuitively easily interpretable $I(1)/I(0)$ case. However, recent advances in long memory time series modeling has provided enough tools to explore the idea of fractional cointegration empirically.

Conceptually, an n-dimensional fractionally cointegrated system with cointegration rank r can be represented as follows:

$$\Delta(\psi)RX_t = u_t \sim I(0), \;\; t \geq 0, \tag{1.4}$$

where $\psi = (\delta_1, \ldots, \delta_r, d)$, $\Delta(\psi) = blockdiag\left(diag(\Delta_+^{\delta_1}, \ldots, \Delta_+^{\delta_r}), \Delta_+^d I_{n-r}\right)$ is a block-diagonal matrix, $\delta_i \leq \delta_{i+1} < d$ and R is an $n \times n$ full-rank matrix. In this framework the first r rows of R represent cointegration vectors and estimation of them is probably the most studied issue in fractional cointegration literature. One of the earliest estimation methods in the bivariate framework was the narrow-band ordinary least squares (NBLS) regression, developed in a series of papers by Robinson (1994), Robinson and Marinucci (2001a), Marinucci and Robinson (2001). The resulting estimator is an OLS estimator in a spectral regression with a degenerating frequency band around zero. The narrow-band OLS estimator reduces the bias in comparison to the OLS estimator, since the effect of contemporaneous correlation between cointegration errors is reduced. The convergence rate of the estimator depends on the values of fractional coefficients and the bandwidth used in estimation and although it is higher than that of the OLS estimator, in some part of the parameter space the NBLS estimator is inconsistent and later studies considered generalized-least-squares-type filtering of observed time series before narrow band estimation, which remedies the problem. Robinson and Hualde (2003) and Hualde and Robinson (2007) proposed estimators in the strong and the weak cointegration cases, respectively, based on narrow-band estimation of both fractionally- and GLS-filtered series. Recently, Hualde and Robinson (2010) studied a more general framework with possibly different memories of cointegration errors and proposed an estimator based on the same idea.

Another way to conduct inference in fractionally cointegrated time series is a model-based inference. The first model for fractional cointegration was proposed in Granger (1986), however the solution of the model has not been derived and its properties are unknown. Johansen (2008) proposed an alternative cofractional vector autoregression (VAR) model:

$$\Delta^d X_t = \Delta^{d-b} L_b \alpha\beta' X_t + \sum_{i=1}^{k} \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{1.5}$$

where $L_b = 1 - \Delta^b$, ε_t is an i.i.d. series with positive definite variance-covariance matrix,

$0 < b \leq d$ and α, β are $n \times r$ matrices. Under suitable conditions on the parameters, the solution of the model satisfies: $X_t \sim I(d)$, $\beta' X_t \sim I(d-b)$, thus allowing for fractional cointegration. Recently, Johansen and Nielsen (2012) developed a likelihood inference procedure in the model for all model parameters, including the cointegration rank.

Current econometric research in fractional cointegration has advanced considerably and recent empirical applications of fractional cointegration include not only economic and financial time series, but also climate, geology and political time series (for a more in-depth discussion, see Gil-Alana and Hualde (2009)).

1.3 Empirical application

A present value relation states that the market price of a financial asset is equal to the discounted expected value of its future payoffs adjusted for investment risk. Shiller (1979) showed a particular implication of the present value model to the yields of bonds which has been termed the "expectations hypothesis". Specifically, the expectations hypothesis (EH) implies that if a bivariate system of short-term and long-term interest rate series is a unit root process, then the series is cointegrated with cointegration vector $(1, -1)$. This implication has been empirically tested with cointegration methods in a large number of studies using different datasets dating back to the seminal paper Campbell and Shiller (1987). Generally, empirical evidence yielded rather mixed results as far as the validity of the expectations hypothesis is concerned, but we note that most studies relied on the classical cointegrated VAR framework, which assumes that the observed time series is an $I(1)$ series, whereas the cointegration error is an $I(0)$ series. We argue that the analysis of interest rate series has to be embedded in the fractional framework: both interest rates and spread terms may be fractionally integrated (see Lardic and Mignon (2004), Iacone (2009) among others) and in this case the analysis within $I(1)/I(0)$ framework may result in a misspecified model and wrong conclusions[2].

Throughout this thesis we use monthly U.S. Treasury yields of zero-coupon bills with maturities of 1, 3 and 6 months over the period from January 1983 to December 2000[3]. The data are obtained from the website of prof. F.X. Diebold and contains $T = 216$ monthly observations, depicted in Figure (1.1) (names of the series imply their maturities).

[2]The consequences of $I(1)/I(0)$ analysis in a fractionally cointegrated system depend on a particular analysis framework at hand, but, for example, in the framework of Chapter 4, it would result in an inefficient inference for the cointegration vector(s).

[3]Since the methods of this thesis are not robust to structural breaks in time series models, the choice of sample period was motivated by findings of a structural break in U.S. interest rate series around October 1982, see Hansen (2003) among others.

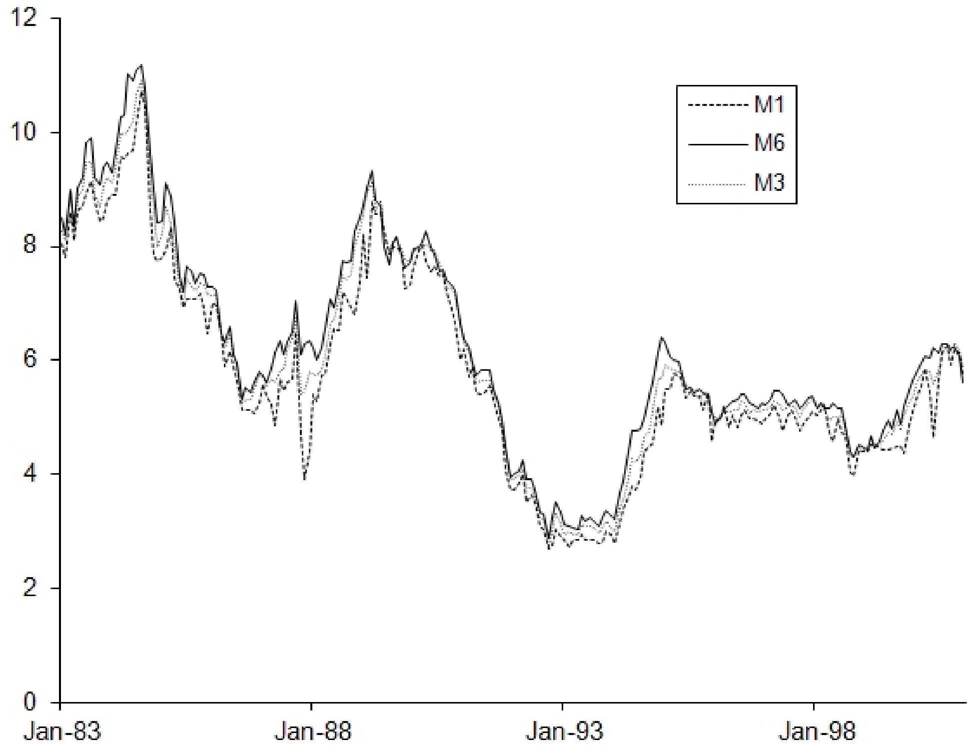

Figure 1.1: Graph of yields of T-bills

The dataset is the updated version of the Fama and Bliss (1987) dataset obtained by a piecewise-linear fitted forward rate curve (a so called unsmoothed Fama-Bliss forward rates curve). The method is based on an iterative procedure (known as “bootstrap”) in which the discount rate function is extended in each step by computing the forward rate necessary to price bonds with longer maturities given the discount rate function already fitted to price bonds with shorter maturities. The full description of filters used and estimation methods can be found on the documentation website of the Center for Research in Security Prices[4]. In the thesis we focus on the short end of the term structure, since it was observed in the literature (cf. Campbell and Shiller (1991)), that the behaviour of the short and long end of the term structure differs.

In this thesis we analyze interest rate series within the fractional framework: firstly, we test the assumption that interest rate series are $I(1)$ processes in Chapter 2, then we test for cointegration in bivariate systems of interest rates in Chapter 3, estimate the cointegration structure and test implications of EH on bivariate interest rate series in Chapter 4 and, finally, we perform model-based inference on trivariate interest rate series

[4] `http://www.crsp.com/documentation/index.html`

in Chapter 5.

1.4 Overview of the thesis

The results of this thesis are presented in four self-contained chapters, which study different aspects of fractionally (co)integrated time series models. Throughout the whole thesis the U.S. Treasury yield dataset is employed to empirically illustrate the results of the individual chapters.

Chapter 2 studies estimation and inference in fractionally integrated time series via an autoregressive approximation. It proposes a semiparametric time domain estimator for the fractional parameter d and considers tests for point null hypotheses of the type $d = d_0$, generalizing some results in previous literature.

Chapter 3 considers testing for fractional cointegration. Testing for cointegration in a classic $I(1)/I(0)$ framework is a well-developed topic and there has been a number of studies, which consider testing for fractional cointegration, however there has been relatively little attention paid so far to the comparison of their finite sample performance. We study size and power properties of regression-based tests for fractional cointegration by means of Monte Carlo simulations and provide some guidelines for empirical researchers.

In Chapter 4 we propose an optimal estimation procedure for cointegration vectors within fractional framework (1.4) based on dynamic ordinary least squares (DOLS) regression. DOLS regression was studied within the $I(1)/I(0)$ framework in Saikkonen (1991) and Stock and Watson (1993), rendering an optimal estimator for cointegration vectors and we show that a similar approach may be applied to the fractional framework.

Chapter 5 studies likelihood-based inference in the fractionally cointegrated model with known rank allowing for possibly different integration orders of cointegration errors. The parameterization of the model assumes a finite order autoregressive process for the cointegration errors, which may be restrictive in theory, but is quite common in practice. Inference in the model is based on a conditional Gaussian likelihood and it is shown that the resulting estimators are consistent and have optimal asymptotic properties.