econometrics

Article

Testing for the Equality of Integration Orders of Multiple Series

Man Wang [1] **and Ngai Hang Chan** [2,*]

[1] Department of Finance, Donghua University, Shanghai 200051, China; wangman@dhu.edu.cn
[2] Department of Statistics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong
* Correspondence: nhchan@sta.cuhk.edu.hk; Tel.: +852-3943-8519

Academic Editor: Pierre Perron
Received: 15 July 2016; Accepted: 25 November 2016; Published: 15 December 2016

Abstract: Testing for the equality of integration orders is an important topic in time series analysis because it constitutes an essential step in testing for (fractional) cointegration in the bivariate case. For the multivariate case, there are several versions of cointegration, and the version given in Robinson and Yajima (2002) has received much attention. In this definition, a time series vector is partitioned into several sub-vectors, and the elements in each sub-vector have the same integration order. Furthermore, this time series vector is said to be cointegrated if there exists a cointegration in any of the sub-vectors. Under such a circumstance, testing for the equality of integration orders constitutes an important problem. However, for multivariate fractionally integrated series, most tests focus on stationary and invertible series and become invalid under the presence of cointegration. Hualde (2013) overcomes these difficulties with a residual-based test for a bivariate time series. For the multivariate case, one possible extension of this test involves testing for an array of bivariate series, which becomes computationally challenging as the dimension of the time series increases. In this paper, a one-step residual-based test is proposed to deal with the multivariate case that overcomes the computational issue. Under certain regularity conditions, the test statistic has an asymptotic standard normal distribution under the null hypothesis of equal integration orders and diverges to infinity under the alternative. As reported in a Monte Carlo experiment, the proposed test possesses satisfactory sizes and powers.

Keywords: asymptotic normal; fractional cointegration; Monte Carlo experiment; residual-based test

JEL Classification: C12; C32

---

1. Introduction

By allowing the equilibrium error to follow a fractionally integrated process, fractional cointegration constitutes a useful extension of classical cointegration. It has received considerable attention in the statistics, finance and econometric literature. There are several notions of (fractional) cointegration for a p-dimensional time series X_t (see Engle and Granger (1987) [1], Johansen (1996) [2], Flôres and Szafarz (1996) [3] and Robinson and Yajima (2002) [4] among others). In the definition studied in Robinson and Yajima (2002) [4], a p-vector X_t is partitioned into several sub-vectors such that elements in each sub-vector have the same integration order. Furthermore, X_t is said to be (fractionally) cointegrated if a cointegration exists in any of the sub-vectors. Under this setting, partitioning X_t requires testing for the homogeneity of integration orders of multiple time series, which has attracted much interest. Current procedures usually assume stationarity and invertibility. For example, Heyde and Gay (1993) [5] and Hosoya (1997) [6] investigate this problem based on a parametric setting, and Robinson (1995) [7] and Lobato (1996 and 1999) [8,9] study the problem using the semiparametric framework. When cointegration exists or the time series becomes nonstationary, some of these tests become invalid.

Robinson and Yajima (2002) [4] construct a single-test statistic that is valid in the presence of cointegration for testing the homogeneity of the fractional integration orders of multiple (asymptotically) stationary and invertible time series. They propose estimating the fractional integration order using the local Whittle likelihood method and introduce a user-chosen number to deal with the inversion of an asymptotically singular matrix. Nielsen and Shimotsu (2007) [10] extend this test statistic to accommodate both (asymptotically) stationary and nonstationary time series by applying the exact local Whittle likelihood method of Shimotsu and Phillips (2005) [11]. The simulation results in Nielsen and Shimotsu (2007) [10] show that the test statistic is sensitive to the choice of the user chosen number, which is assumed to satisfy certain conditions. Hualde (2013) [12] proposes a residual-based test, which covers the nonstationary and noninvertible series, and is valid irrespective of whether cointegration exists. Although this test is developed for a bivariate series, extending it to the multivariate case is non-trivial because multiple comparisons are needed when high-dimensional series are involved. There are two ways to extend the Hualde (2013) [12] result. The first involves testing the equality of each pair of integration orders, which requires $p(p-1)/2$ simple tests for a p-dimensional series. When p is large, this test procedure becomes computationally intensive. The second extension is to explore the possibility of a one-step single test, which is pursued here.

In this paper, a residual-based testing procedure for the equality of integration orders of a multiple fractionally integrated process is proposed. The test encompasses both the stationary/nonstationary and invertible/noninvertible situations, and is valid even when the time series is cointegrated. The procedure is computationally feasible because it is a one-step test without inverting ill-conditioned matrices under cointegration. The test can be computed very fast even when dealing with a large p. The test statistic converges to a standard normal distribution under the null hypothesis that all integration orders are equal, and diverges when there are different integration orders.

This paper is organized as follows. In Section 2, the testing procedure and asymptotic theory are presented. Empirical sizes and powers of the proposed test are given via a Monte Carlo study in Section 3. Section 4 concludes the paper.

2. Integration Orders

Consider the following p-dimensional time series $(x_{1,t}, x_{2,t}, \ldots, x_{p,t})'$, with prime denoting transposition and $t \in \{0, \pm 1, \pm 2, \ldots\}$,

$$\begin{aligned} x_{1,t} &= \Delta^{-\delta_1}\{v_{1,t}\mathbf{1}(t>0)\}, \ \ x_{1,t} = 0, t \le 0, \\ &\vdots \\ x_{p,t} &= \Delta^{-\delta_p}\{v_{p,t}\mathbf{1}(t>0)\}, \ \ x_{p,t} = 0, t \le 0, \end{aligned} \tag{1}$$

where $\mathbf{1}(\cdot)$ is the indicator function, $\Delta = 1 - L$, L is the lag operator, and $v_t = (v_{1,t}, \ldots, v_{p,t})'$ is a vector of zero mean covariance stationary processes. Note that the series $\{x_{i,t}\}$ is nonstationary for $\delta_i > 1/2$ and "asymptotically stationary" for $\delta_i < 1/2, i = 1, \ldots, p$. By Taylor's expansion, $\Delta^\alpha = \sum_{j=0}^\infty \pi_j(-\alpha)L^j, \pi_j(\alpha) = \frac{(\alpha)_j}{j!}$, where $(\alpha)_j = (\alpha)(\alpha+1)\ldots(\alpha+j-1)$. If α is not a negative integer, then $\pi_j(\alpha) = \frac{\Gamma(j+\alpha)}{\Gamma(\alpha)\Gamma(j+1)}$. When α is a negative integer, then $\pi_j(\alpha) = 0$ for $j > -\alpha$ and $\Delta^{-\alpha}$ becomes the usual formula of differencing with integer orders. The symbol $||\cdot||$ is used to represent the Euclidean norm and $A \sim B$ means that A/B converge to a constant or converge in distribution to a random variable as n goes to ∞.

Assumption 1. *Consider the process* $v_t = A(L)\epsilon_t, t \in Z$ *with* $A(L) = \sum_{j=0}^\infty A_j L^j$. *Assume that*

1.1. $\sum_{j=1}^\infty j||A_j||^2 < \infty$;

1.2. *ϵ_t are i.i.d vectors with mean zero, positive definite covariance matrix Ω and $E||\epsilon_t||^q < \infty$ for some $q > max\{2, 1/(\bar{\delta} + 1/2)\}$, where $\bar{\delta} = min\{\delta_i\}_{i=1}^{p}$.*
1.3. *$f_{ii}(0) > 0, i = 1, 2$, where $f(\lambda)$ is the spectral density matrix of v_t and $f_{ij}(0)$ is the $(i,j) - th$ element of $f(0)$.*

Assumption 1 is mild because it is satisfied by the usual stationary and invertible autoregressive moving average (ARMA) processes. This is a common assumption for applying the functional limit theorem of Marinucci and Robinson (2000) [13], and it has appeared in a similar form as Assumptions A–C of Marmol and Velasco (2004) [14], Assumption A of Hualde (2013) [12] and Assumption 1 of Wang, Wang and Chan (2015) [15]. In Particular, the moment condition in Assumption 1.2 is discussed by Johansen and Nielsen (2012) [16]. As pointed out in Wang, Wang and Chan (2015) [15], Assumption 1.1 ensures that the limiting process of the partial sum of v_t has nondegenerated finite-dimensional distributions. Assumption 1.1 implies that $f(\lambda)$ is $Lip(\gamma), \gamma > 0$.

Under Assumption 1, model (1) means that all $x_{i,t}, i = 1, \ldots, p$ are type-II fractionally integrated processes. Furthermore, based on the fractional cointegration definition given in Robinson and Yajima (2002) [4], if the integration orders of $x_{it}, i = 1, \ldots, p$ are the same and there exists a non-zero linear combination $\beta' x_t$ that is $I(b)(b < \delta_i)$, then the p-dimensional time series x_t is said to be cointegrated. Furthermore, any multiple time series containing x_t as a sub-vector is also said to be cointegrated.

To test whether all of the $\delta_i, i = 1, \ldots, p$ are the same, we need to estimate δ_i precisely. Thus, the following assumptions are introduced.

Assumption 2. *Under both the null and alternative hypotheses,*

2.1. *there exists a positive constant $K < \infty$ and estimates $\hat{\delta}_i$ of δ_i, $i = 1, \ldots, p$, respectively, such that*

$$\sum_{i=1}^{p} |\hat{\delta}_i| \leq K, \tag{2}$$

and there exists $\kappa > 0$,

$$\hat{\delta}_i - \delta_i \sim n^{-\kappa}; \tag{3}$$

2.2. *Letting $\hat{f}(0)$ be an estimate of $f(0)$, then $\hat{f}(0) \xrightarrow{p} f(0)$, where $\xrightarrow{p}$ stands for the convergence in probability.*

Assumption 2 is very mild, as condition (2) is satisfied if $\hat{\delta}_i, i = 1, \ldots, p$ are optimizers of the corresponding functions over compact sets. $\delta_i, i = 1, \cdots, p$ can be estimated by semiparametric methods (see, for example, the log periodogram estimate of Geweke and Porter-Hudak (1983) [17] studied by Hurvich et al. (1998) [18] or the narrow-band Gaussian or Whittle estimate introduced by Künsch (1987) [19] and studied in Robinson (1995) [7] and Lobato (1999) [9]. Equation (3) is satisfied by many estimation methods, such as that used in Beran (1995) [20] and Tanaka (1999) [21]. As pointed out by Hualde and Velasco (2008) [22], Equation (3) is satisfied if δ_i is estimated from $x_{i,t}$ using the usual parametric or semiparametric methods. For example, the Whittle pseudo-maximum likelihood estimation proposed by Velasco and Robinson (2000) [23] satisfies (3). In particular, if a parametric structure is imposed on v_t, then a $\sqrt{n}$-consistent estimator results by means of a multivariate extension of Robinson (2005) [24]. Assumption 2.2 is quite common and is satisfied by many classic semiparametric or nonparametric estimates. Actually, a stricter condition on the convergence rate of $\hat{f}(0)$ ($\hat{f}(0) - f(0) = O_p(n^{-\chi})$, with χ being a positive constant) is used in many articles, such as Hualde and Robinson (2006) [25], Hualde and Robinson (2010) [26], Hualde and Velasco (2008) [22] and Wang (2008) [27], among others. In particular, Hualde and Robinson (2006) [25] discuss the convergence rate of some estimates of f, including a weighted periodogram estimate that satisfies Assumption 2.2. Hualde and Velasco (2008) [22] point out that the nonparametric estimate of $f(0)$ introduced in their paper satisfies Assumption 2.2. Once $\hat{\delta}_i$ is estimated, the nonparametric estimator of $f(0)$ can be based on the weighted averages of the periodogram of the proxy $\hat{v}_t = (x_{1,t}(\hat{\delta}_1), \ldots, x_{p,t}(\hat{\delta}_p))'$, where $x_{i,t}(\hat{\delta}_i) = \Delta^{\hat{\delta}_i}\{x_{i,t}\mathbf{1}(t > 0)\}$.

Let $h_n > 0$ be a sequence such that

$$h_n^{-1} + n^{-\kappa} h_n \to 0 \text{ as } n \to \infty. \tag{4}$$

Let $d = \sum_{i=1}^p \delta_i$, $\hat{d} = \sum_{i=1}^p \hat{\delta}_i$ and

$$\hat{a} = (I_1, I_2, \ldots, I_p)', \tag{5}$$

where $I_i = \mathbf{1}\{A_i \cap B_i\}$, $A_i = \{n^\kappa(\hat{\delta}_i - \max_{j=1,\ldots,p, j\neq i}\{\hat{\delta}_j\}) \geq -h_n\}$ and $B_i = \{n^\kappa(\hat{\delta}_i - \max_{j=1,\ldots,i-1}\{\hat{\delta}_j\}) > h_n\}$. Furthermore, for $i = 1$, let $\max_{j=1,\ldots,i-1}\{\hat{\delta}_j\} = -\infty$. Clearly, B_1 is the entire sample space with $P(B_1) = 1$.

Defining $\delta_i^* = \frac{d-\delta_i}{p-1}$ and $\hat{\delta}_i^* = \frac{\hat{d}-\hat{\delta}_i}{p-1}$, we denote:

$$\hat{F} = F(\hat{\delta}, \hat{f}(0)) = \frac{\hat{a}' \sum_t x_t(\hat{\delta}_1^*, \ldots, \hat{\delta}_p^*)}{(2n\pi)^{1/2} \hat{a}' \hat{f}(0) \hat{a}},$$

as the test for $H_0 : \delta_1 = \cdots = \delta_p$ against the alternative H_1: there exists at least a pair of (i, j) such that $\delta_i \neq \delta_j$.

Theorem 1. *Letting Assumptions 1 and 2 hold, x_t is defined in (1), and then $\hat{F} \xrightarrow{d} N(0, 1)$ under H_0 and $\hat{F} = O_p(n^{\frac{p*\max\{\delta_i\}-d}{p-1}})$ under H_1, where $\xrightarrow{d}$ stands for convergence in distribution as $n \to \infty$.*

Remark 1. *Denote the set of indices of the maxima of δ_i as $S = \{j, \delta_j = max\{\delta_i\}_{i=1}^p\}$, and let m_0 be the smallest index of the maxima, that is, $m_0 = min\{S\}$. Furthermore, let $a = \mathbf{e}_{m_0}$, where $\mathbf{e}_{m_0}$ is the unit vector that equals one at the m_0-th coordinate and zero otherwise. Then, it is shown in the proof of Theorem 1 that $\hat{a} \xrightarrow{p} a$.*

Remark 2. *The vector $\hat{a}$ can also be set as a vector of constants: $a = (a_1, \cdots, a_p)'$, which satisfies $a' f(0) a \neq 0$. As $\hat{f}(0) \to f(0)$ in probability, $a' \hat{f}(0) a > 0$ with probability 1. However, with $\{\delta_i\}_{i=1}^p$ unknown, it is not guaranteed that $\hat{F}$ diverges under H_1 at a rate as fast as that specified in Theorem 1. Wang (2008) [27] shows that different pre-determined $\hat{a}$ may lead to different divergence rates.*

Remark 3. *As pointed out in Remark 2, the choice of $\hat{a}$ has an influence on the diverging speed of $\hat{F}$. From the proof of Theorem 1, to get the theoretical diverging speed of $\hat{F}$ as in Theorem 1, define $\hat{a}$ by Equations (4) and (5). Then, $\hat{a} \xrightarrow{p} \mathbf{e}_{m_0}$ when $n \to \infty$, with m_0 being the smallest index of the maxima of $\{\delta_i\}_{i=1}^p$. Consequently, the denominator of $\hat{F}$ converges to $(2n\pi)^{1/2} f_{m_0,m_0}(0) > 0$. Similar to the analysis in Hualde (2013) [12] and Wang, Wang and Chan (2015) [15], it is natural to replace condition (4) by setting $h_n = 0$, in which case $\hat{a}$ converges to a random limit under H_0. Furthermore, the limits of the numerator and denominator of $\hat{F}$ are dependent, which complicates analysis of the asymptotic distribution of $\hat{F}$. From the definition of $\hat{a}$, it is obvious that the power of the proposed test with $h_n = 0$ is superior to that of tests with other choices of h_n. However, when the sample size $n \to \infty$, the powers of different cases will become the same. In practice, $h_n = logn^\kappa$ or $h_n = n^{\kappa/2}$ are two possible choices. In particular, if the parametric method in Hualde and Robinson (2011) [28] is used, $\kappa = 1/2$, then we can set $h_n = n^{1/4}$.*

Remark 4. *If $(x_{1,t}, x_{2,t}, \ldots, x_{p,t})$ is cointegrated with $\beta' x_t = \Delta^b u_t$, $\beta \neq 0$, $b < \delta_1 = \delta_2 = \cdots = \delta_p$, then $f(0)$ would be singular. In this situation, most of the tests in the literature involve the inverse of $f(0)$ and become invalid under H_0. However, the proposed test still works in the presence of cointegration. As $f_{m_0,m_0}(0) > 0$ by*

Assumption 1, and $\hat{a} \xrightarrow{p} a = \mathbf{e}_{m_0}$ *as mentioned in Remark 1, we have* $a' f(0) a > 0$. *Furthermore, as shown in Theorem 1,* $\hat{a}' f(0)\hat{a}$ *converges to* $a' f(0) a > 0$ *in probability. Then,* $\hat{a}' \hat{f}(0)\hat{a}$ *is positive with probability 1, and* $\hat{F}$ *remains valid under cointegration.*

3. Simulation

To assess the performance of our testing procedure, we conduct two Monte Carlo experiments. For both experiments, we generate $(x_{1,t}, x_{2,t}, x_{3,t})'$ as in (1) with v_t being a three-dimensional white noise with $\mathrm{E}(v_t) = 0$, $\mathrm{Var}(v_{i,t}) = 1$ for $i = 1, 2, 3$, $\mathrm{Cov}(v_{i,t}, v_{j,t})$ =0.5. We compute $\hat{F}$ parametrically, which means $\hat{\delta}_i, i = 1, 2, 3$ are estimated as in Hualde and Robinson (2011) [28] and $f(0)$ is estimated by $\hat{f}(0) = (2\pi n)^{-1/2} \sum_{t=1}^{n} \hat{v}_t \hat{v}_t'$.

For the first experiment, using 10,000 replications and 3 different sample sizes $n = 100, 250$ and 500, we compute the proportion of rejecting $\hat{F}$ for nominal size $\alpha = 0.01$, 0.05, and 0.1 with different combinations of $(\delta_1, \delta_2, \delta_3)$. Letting $\phi = \frac{p * \max\limits_{i=1,\ldots,p}\{\delta_i\} - d}{p-1}$, we consider $\phi = 0, 0.3, 0.6, 0.8$ and 1.0. To investigate the sensitivity of the choice of h_n, we present the result for $h_{1n} = 0, h_{2n} = \log(n^{\kappa}), h_{3n} = n^{\kappa/2}$ with $\kappa = 1/2$ in Table 1.

Table 1. Empirical sizes and powers based on different δ and α.

	n	100			250			500		
h_n	α	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
h_{1n}	ϕ = 0	0.0603	0.1572	0.2874	0.0453	0.1356	0.2317	0.0415	0.1228	0.1969
	ϕ = 0.3	0.5687	0.6726	0.7508	0.6615	0.7473	0.7881	0.7357	0.8113	0.8522
	ϕ = 0.6	0.8730	0.9116	0.9288	0.9334	0.9598	0.9657	0.9767	0.9814	0.9892
	ϕ = 0.8	0.9427	0.9562	0.9653	0.9693	0.9805	0.9833	0.9861	0.9896	0.9932
	ϕ = 1.0	0.9733	0.9750	0.9820	0.9922	0.9951	0.9964	0.9972	0.9985	0.9987
h_{2n}	ϕ = 0	0.0134	0.056	0.1127	0.0060	0.0533	0.105	0.0057	0.0523	0.1024
	ϕ = 0.3	0.4724	0.5803	0.6437	0.5360	0.6875	0.7480	0.7371	0.8158	0.8463
	ϕ = 0.6	0.8651	0.9082	0.9224	0.9392	0.9537	0.9556	0.9675	0.9804	0.9893
	ϕ = 0.8	0.9427	0.9562	0.9653	0.9693	0.9805	0.9833	0.9861	0.9896	0.9932
	ϕ = 1.0	0.9733	0.9750	0.9820	0.9922	0.9951	0.9964	0.9972	0.9985	0.9987
h_{3n}	ϕ = 0	0.0047	0.0507	0.1068	0.0046	0.0482	0.1035	0.0049	0.0484	0.1033
	ϕ = 0.3	0.4230	0.5399	0.6045	0.5168	0.6404	0.7006	0.6385	0.7334	0.7842
	ϕ = 0.6	0.8457	0.8873	0.9162	0.9384	0.9625	0.9706	0.9727	0.9748	0.9881
	ϕ = 0.8	0.9427	0.9562	0.9653	0.9693	0.9805	0.9833	0.9861	0.9896	0.9932
	ϕ = 1.0	0.9733	0.9750	0.9820	0.9922	0.9951	0.9964	0.9972	0.9985	0.9987

First, consider the sizes, that is, $\phi = 0$. We observe that for h_{1n}, $\hat{F}$ is oversized and the empirical sizes of case h_{2n} and h_{3n} are very close to the nominal sizes. As n increases, the empirical sizes under all scenarios approach the nominal sizes as expected. We also examine the power for $\phi = 0.3, 0.6, 0.8$ and 1.0. It can be seen that the empirical power increases as n and ϕ increase, and that $\hat{F}$ performs very well for all choices of $h_{in}, i = 1, 2, 3$. As expected, a smaller h_n leads to better power, so h_{1n} has the best power and h_{2n} has better power than h_{3n}. As ϕ increases, the difference decreases substantially, and it is clear that for $\phi \geq 0.6$, the powers of all $h_{in}, i = 1, 2, 3$ are almost the same. One explanation is that when ϕ is large enough, $n^{-\kappa} h_{in}, i = 1, 2, 3$ become relatively small compared with ϕ, leading to the same $\hat{a}$. As ARMA models are common in modeling stationary time series, autoregressive fractionally integrated moving averaging (ARFIMA) models constitute a reasonable approximation to x_t when the parametric method in Hualde and Robinson (2011) [28] is considered. In practice, if there is insufficient information about the true model, a general ARFIMA(p_1, δ_0, p_2) model is entertained first and a model selection procedure based on some information criteria is conducted to choose p_1 and p_2.

For the second experiment, we conduct a simulation to compare the proposed test $\hat{F}$ with the test in Nielsen and Shimotsu (2007) [10]:

$$\hat{T}_0 = m(S\hat{\delta})' \left(S\frac{1}{4}\hat{D}^{-1}\left(\hat{G}o\hat{G}\right)\hat{D}^{-1}S' + k_n^2 I_{p-1} \right)^{-1} (S\hat{d}),$$

where m is the bandwidth parameter; $\delta = (\delta_1, \delta_2, \cdots, \delta_p)'$ is the vector of integration orders of $(x_{1,t}, x_{2,t}, \ldots, x_{p,t})'$; o is the Hadamard product; I_{p-1} is the $(p-1)$-dimensional identity matrix; $S = [I_{p-1}, -\iota]$, with ι being the $(p-1)$-vector of ones; k_n is a positive sequence satisfying certain assumptions; G is the spectral density matrix of the δ'th differenced process around the origin; and D is the diagonal matrix of G. Using 5,000 replications and 3 different sample sizes n = 128, 256 and 512, we report the rejection frequencies of $\hat{F}$ with $h_{3n} = n^{\kappa}, \kappa = 1/2$, as well as $\hat{T}_0$ with bandwidth parameter $m = [n^{0.6}]$ and two choices of k_n, that is $k_{1n} = 1/\log(n)$ and $k_{2n} = 1/(\log(n))^{1/2}$ in Table 2. Here, $[z]$ denotes the largest integer smaller than or equal to z. The fractional integration order δ is estimated by the exact local Whittle likelihood for $\hat{T}_0$.

Table 2. Empirical sizes and powers of $\hat{F}$ and $\hat{T}_0$.

	n	128			256			512		
	α	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
$\hat{F}$ with h_{3n}	$\phi = 0$	0.02	0.0566	0.1148	0.016	0.0514	0.1118	0.0138	0.0514	0.1106
	$\phi = 0.3$	0.5172	0.6224	0.6842	0.5718	0.6698	0.7120	0.6398	0.7342	0.7842
	$\phi = 0.6$	0.8622	0.8976	0.9134	0.9328	0.9572	0.9680	0.9712	0.9758	0.9854
	$\phi = 0.8$	0.9592	0.9682	0.9742	0.9800	0.9850	0.9874	0.9902	0.9926	0.9938
	$\phi = 1.0$	0.9694	0.9758	0.9802	0.9858	0.9884	0.9902	0.9968	0.9976	0.9984
$\hat{T}_0$ with k_{1n}	$\phi = 0$	0.1310	0.2438	0.3278	0.1280	0.2438	0.3278	0.1010	0.2008	0.3076
	$\phi = 0.3$	0.5584	0.7144	0.7860	0.5584	0.7184	0.7860	0.5684	0.7184	0.7968
	$\phi = 0.6$	0.9722	0.9890	0.9944	0.9722	0.9890	0.9944	0.9742	0.9890	0.9974
	$\phi = 0.8$	0.9964	0.9994	0.9994	0.9968	0.9994	0.9994	0.9972	0.9996	0.9996
	$\phi = 1.0$	0.9988	0.9998	1	1	1	1	1	1	1
$\hat{T}_0$ with k_{2n}	$\phi = 0$	0.0490	0.1154	0.1808	0.0490	0.1154	0.1808	0.0498	0.1156	0.1810
	$\phi = 0.3$	0.3680	0.5662	0.6658	0.3680	0.5662	0.6658	0.3780	0.5682	0.6678
	$\phi = 0.6$	0.9352	0.9772	0.9868	0.9552	0.9782	0.9868	0.9552	0.9782	0.9868
	$\phi = 0.8$	0.9872	0.9962	0.9980	0.9892	0.9964	0.9980	0.9892	0.9964	0.9980
	$\phi = 1.0$	0.9950	0.9986	0.9994	1	1	1	1	1	1

We find that all of the three tests are oversized, and that their empirical powers increase when ϕ increases. However, the empirical powers and empirical sizes of $\hat{T}_0$ do not change much when the sample size changes from 128 to 512, while those of $\hat{F}$ improve significantly when n increases.

We first compare the simulation results of $\hat{T}_0$ with k_{1n} and k_{2n}. It is obvious that $\hat{T}_0$ is sensitive to the choice of k_n: $\hat{T}_0$ works reasonably well for $k_{2n} = 1/(\log n)^{1/2}$ and $\hat{T}_0$ over-rejects substantially for $k_{1n} = 1/\log n$. The test $\hat{T}_0$ is oversized for both k_{1n} and k_{2n}, and k_{2n} has a better empirical size and k_{1n} better empirical power. This phenomenon is also reported in Nielsen and Shimotsu (2007) [10].

We then compare $\hat{F}$ with $\hat{T}_0$ and find that, for all sample sizes n, $\hat{F}$ has much better empirical sizes than $\hat{T}_0$ for both k_{1n} and k_{2n}. The empirical power of $\hat{F}$ is not as good as that of $\hat{T}_0$ when the sample size is relatively small (128 and 256). However, as the sample size increases to 512, the empirical power of $\hat{F}$ becomes superior to that of $\hat{T}_0$.

4. Conclusions

A residual-based test for testing the equality of the integration orders of multiple fractionally integrated processes is proposed in this paper. The test is valid under cointegration and is computationally feasible. One needs only to estimate the integration order and the spectral density function of the process that generates the fractionally integrated processes. The proposed test enjoys standard asymptotics and possesses satisfactory finite sample behavior.

Acknowledgments: We thank the editor and three anonymous referees for their helpful comments and useful references, which led to an improved version of this paper. This research was supported in part by General Research Fund of HKSAR-RGC-GRF grant No. 400313, 14300514 and 14325216; HKSAR-RGCCRF:CityU8/CRG/12G; the Theme-based Research Scheme of HKSAR-RGC-TBS T32-101/15-R (Chan) and Fundamental Research Funds for the Central Universities 16D110810; and Donghua University Special Research Funds for Social Science 108-10-0108135 (Wang).

Author Contributions: Both authors contributed equally to the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Lemma A1. *Let δ_i, δ_i^* and $\hat{\delta}_i^*$ be defined as in Section 2. Then* $n^{-1/2}\sum_{t=1}^{n}\Delta^{\hat{\delta}_i^*}x_{i,t} - n^{-1/2}\sum_{t=1}^{n}\Delta^{\delta_i^*}x_{i,t} = \begin{cases} o_p(1), & \text{under } H_0, \\ o_p(n^{\delta_i-\delta_i^*}), & \text{under } H_1. \end{cases}$

Proof. Let $g(\lambda, z_t) = \Delta^{\lambda}z_t$. Then $g(\lambda, z_t) = \sum_{i=0}^{t-1}\pi_i(\lambda)z_{t-i}$ if $z_t = 0$ for $t \leq 0$, where $\pi_i(\cdot)$, $i = 1, \ldots, t-1$ are as defined in Section 2 and the derivatives $g^{(r)}(\lambda, z_t) = \sum_{i=1}^{t-1}\pi_i^{(r)}(\lambda)z_{t-i}$, where $\pi_i^{(r)}(\lambda) = d^r\pi_i(\lambda)/d\lambda^r$. Based on Taylor's expansion around δ_i, for a certain constant R to be defined subsequently, we can show that

$$\begin{aligned}
& n^{-1/2}\sum_{t=1}^{n}\Delta^{\hat{\delta}_i^*}x_{i,t} - n^{-1/2}\sum_{t=1}^{n}\Delta^{\delta_i^*}x_{i,t} \\
=& n^{-1/2}\sum_{t=1}^{n}\left(g(\delta_i - \hat{\delta}_i^*; v_{i,t}) - g(\delta_i - \delta_i^*; v_{i,t})\right) \\
=& \frac{1}{\sqrt{n}}\sum_{r=1}^{R-1}\frac{(\delta_i^* - \hat{\delta}_i^*)^r}{r!}\sum_{t=1}^{n}g^{(r)}(\delta_i - \delta_i^*; v_{i,t}) + \frac{(\delta_i^* - \hat{\delta}_i^*)^R}{R!\sqrt{n}}\sum_{t=1}^{n}g^{(R)}(\delta_i - \tilde{\delta}; v_{i,t}) \qquad \text{(A1)} \\
=& \begin{cases} o_p(1), & \text{under } H_0, \\ o_p(T^{\delta_i - \delta_i^*}), & \text{under } H_1, \end{cases} \qquad \text{(A2)}
\end{aligned}$$

where $\tilde{\delta} \in (\min(\delta_i^*, \hat{\delta}_i^*), \max(\delta_i^*, \hat{\delta}_i^*))$.

(A1) and (A2) can be derived based on reasoning similar to that of Theorem 1 of Wang, Wang and Chan (2015) [15] or Theorem 1 of Hualde (2013) [12], under Assumptions 1 and 2. In particular, to verify (A2), we apply the functional central limit theorem as in Marinucci and Robinson (2000) [13], which is guaranteed by Assumption 1. □

Proof of Theorem 1. First, we show that $\hat{a} \xrightarrow{p} a$, where $\hat{a} = (I_1, I_2, \ldots, I_p)$, with $I_i = \mathbf{1}\{A_i \cap B_i\}$, $A_i := \{n^{\kappa}(\hat{\delta}_i - \max_{j=1,\ldots,p, j\neq i}\{\hat{\delta}_j\}) \geq -h_n\}$, $B_i = \{n^{\kappa}(\hat{\delta}_i - \max_{j=1,\ldots,i-1}\{\hat{\delta}_j\}) > h_n\}$, and B_1 is as defined in Section 2.

Note that $\forall i \in \{1, \ldots, p\}$,

$$1\{A_i \cap B_i\} + 1\{A_i^c \cup B_i^c\} = 1,$$

and $\hat{a} \xrightarrow{p} a$ is immediately obtained if we show that

$$1\{n^{\kappa}(\hat{\delta}_i - \max_{j=1,\ldots,p,j\neq i}\{\hat{\delta}_j\}) \geq -h_n\} = o_p(1), \quad \text{if } \delta_i < \max_{j=1,\ldots,p}\{\delta_j\}, \tag{A3}$$

$$1\{n^{\kappa}(\hat{\delta}_i - \max_{j=1,\ldots,p,j\neq i}\{\hat{\delta}_j\}) < -h_n\} = o_p(1), \quad \text{if } \delta_i = \max_{j=1,\ldots,p}\{\delta_j\}, \tag{A4}$$

$$1\{n^{\kappa}(\hat{\delta}_i - \max_{j=1,\ldots,i-1}\{\hat{\delta}_j\}) > h_n\} = o_p(1), \quad \text{if } \delta_i \leq \max_{j=1,\ldots,i-1}\{\delta_j\}, \tag{A5}$$

$$1\{n^{\kappa}(\hat{\delta}_i - \max_{j=1,\ldots,i-1}\{\hat{\delta}_j\}) \leq h_n\} = o_p(1), \quad \text{if } \delta_i > \max_{j=1,\ldots,i-1}\{\delta_j\}. \tag{A6}$$

The reason is that, if $i = m_0$, with m_0 as defined in Remark 1, $\delta_i = \max_{j=1,\ldots,p}\{\delta_i\}$ and $\delta_i > \max_{j=1,\ldots,i-1}\{\delta_i\}$, then $1\{A_i^c \cup B_i^c\} \leq 1\{A_i^c\} + 1\{B_i^c\} = o_p(1) + o_p(1) = o_p(1)$ and $1\{A_i \cap B_i\} \xrightarrow{p} 1$.

Otherwise, if $i \neq m_0$, which means $\delta_i < \max_{j=1,\ldots,p}\{\delta_i\}$ or $\delta_i \leq \max_{j=1,\ldots,i-1}\{\delta_i\}$, then $1\{A_i \cap B_i\} \leq 1/2(1\{A_i\} + 1\{B_i\}) = o_p(1) + o_p(1) = o_p(1)$.

Therefore, $I_i \xrightarrow{p} \mathbf{1}\{i = m_0\}$, and furthermore $\hat{a} \xrightarrow{p} a$.

Then, we prove (A3)–(A6). As the definition of $1\{B_i\}$ is similar to the terms that appear in Hualde (2013) [12] and Wang, Wang and Chan (2015) [15], (A5) and (A6) can be proved with similar reasoning. We prove (A3), which means that δ_i is smaller than $\max_{k=1,\ldots,p}\{\delta_k\} = \delta_j$. Denote $Q_n = n^{\kappa}(\hat{\delta}_i - \hat{\delta}_j - (\delta_i - \delta_j))$, then $|Q_n| = O_p(1)$ based on Assumption 2. First, we show that

$$\begin{aligned} 1\{A_i\} &= 1\{n^{\kappa}(\hat{\delta}_i - \max_{k=1,\ldots,p,k\neq i}\{\hat{\delta}_k\}) \geq -h_n\} \\ &= 1\{n^{\kappa}(\hat{\delta}_i - \hat{\delta}_j) \geq -h_n\} \\ &= 1\{Q_n + n^{\kappa}(\delta_i - \delta_j) \geq -h_n\} \\ &\leq \tfrac{|Q_n|}{-h_n + n^{\kappa}(\delta_j - \delta_i)} = o_p(1), \end{aligned} \tag{A7}$$

by (4).

Similarly, for (A4), when $\delta_i = \max_{k=1,\ldots,p}\{\delta_k\} \geq \max_{k=1,\ldots,p,k\neq i}\{\delta_k\}$,

$$\begin{aligned} 1\{A_i^c\} &= 1\{n^{\kappa}(\hat{\delta}_i - \max_{k=1,\ldots,p,k\neq i}\{\hat{\delta}_k\}) < -h_n\} \\ &\leq \textstyle\sum_{k=1,k\neq i}^{p} 1\{n^{\kappa}(\hat{\delta}_i - \hat{\delta}_k) < -h_n\} \\ &= \textstyle\sum_{k=1,k\neq i}^{p} 1\{Q_n + n^{\kappa}(\delta_i - \delta_k) < -h_n\} \\ &= o_p(1), \end{aligned} \tag{A8}$$

since

$$\begin{aligned} &1\{Q_n + n^{\kappa}(\delta_i - \delta_k) < -h_n\}, \\ = &\begin{cases} 1\{-Q_n > h_n\} \leq \frac{|Q_n|}{h_n} = o_p(1), & \text{if } \delta_i = \delta_k, \\ 1\{n^{\kappa}(\delta_i - \delta_k) < -h_n - Q_n\} \leq \frac{|Q_n| + h_n}{n^{\kappa}(\delta_i - \delta_k)} = o_p(1), & \text{if } \delta_i > \delta_k. \end{cases} \end{aligned}$$

Next, we prove that

$$\begin{aligned} &F(\delta, f(0)) = \tfrac{\sum_{t=1}^{n} a' x_t(\delta_1^*,\ldots,\delta_p^*)}{(2n\pi)^{1/2} a' f(0) a}, \\ &\begin{cases} \xrightarrow{d} N(0,1), & \text{under } H_0, \\ = O_p(n^{(p*\max\{\delta_i\}-d)/(p-1)}), & \text{under } H_1. \end{cases} \end{aligned} \tag{A9}$$

Under H_0, $\delta_i^* = \delta_i$, $\frac{n^{-1/2}a' \sum_t x_t(\delta_1,\ldots,\delta_p)}{(2\pi)^{1/2}a' f(0)a}$ converges in distribution to $N(0,1)$ in view of the functional limit theorem of the $I(0)$ process. Under H_1, $\delta_i - \delta_i^* = \frac{p*\delta_i - d}{p-1}$, $\frac{a' \sum_t v_t(-(\delta_1-\delta_1^*),\ldots,-(\delta_p-\delta_p^*))}{(2n\pi)^{1/2}a' f(0)a} = O_p(n^{\frac{p*\max\{\delta_i\}-d}{p-1}})$, based on the properties of the integrated process.

Finally, we show that

$$\begin{aligned}
& n^{-1/2}\sum_{t=1}^{n}(\hat{a}' x_t(\hat{\delta}_1^*,\ldots,\hat{\delta}_p^*) - a' x_t(\delta_1^*,\ldots,\delta_p^*)) \\
=& \frac{(\hat{a}-a)'}{\sqrt{n}}\sum_{t=1}^{n} x_t(\delta_1^*,\ldots,\delta_p^*)) + \frac{\hat{a}'}{\sqrt{n}}\sum_{t=1}^{n}(x_t(\hat{\delta}_1^*,\ldots,\hat{\delta}_p^*) - x_t(\delta_1^*,\ldots,\delta_p^*)) \\
=& \begin{cases} o_p(1) & \text{under } H_0, \\ o_p(n^{(p*\max\{\delta_i\}-d)/(p-1)}), & \text{under } H_1. \end{cases}
\end{aligned} \tag{A10}$$

By Lemma A1, $||\frac{1}{\sqrt{n}}\sum_{t=1}^{n}(x_t(\hat{\delta}_1^*,\ldots,\hat{\delta}_p^*) - x_t(\delta_1^*,\ldots,\delta_p^*))||$ is $o_p(n^{\max\{\delta_i-\delta_i^*\}})$; additionally, $p*\max\{\delta_i\} - d = 0$ under H_0. Thus, it is $o_p(n^{(p*\max\{\delta_i\}-d)/(p-1)})$ under H_1, and is $o_p(1)$ under H_0.

$||\frac{1}{\sqrt{n}}\sum_{t=1}^{n} x_t(\delta_1^*,\ldots,\delta_p^*))||$ is $O_p(n^{(p*\max\{\delta_i\}-d)/(p-1)})$ and $||(\hat{a} - a)||$ is $o_p(1)$, so $\frac{(\hat{a}-a)'}{\sqrt{n}}\sum_{t=1}^{n} x_t(\delta_1^*,\ldots,\delta_p^*))$ is $o_p(n^{(p*\max\{\delta_i\}-d)/(p-1)})$.

Furthermore, based on (A9) and (A10) and given that $\hat{a}'\hat{f}(0)\hat{a} \xrightarrow{p} a'f(0)a > 0$, the proof of Theorem 1 is complete. □

References

1. Engle, R.F.; Granger, C.W.J. Cointegration and error correction: Representation, estimation and testing. *Econometrica* **1987**, *55*, 251–276.
2. Johansen, S. *Likelihood-based inference in cointegrated vector autoregressive models*, 2nd ed.; Oxford University Press: Oxford, UK, 1996.
3. Flôres, R.G., Jr.; Szafarz, A. An enlarged definition of cointegration. *Econ. Lett.* **1996**, *50*, 193–195.
4. Robinson, P.M.; Yajima, Y. Determination of cointegrating rank in fractional systems. *J. Econom.* **2002**, *106*, 217–247.
5. Heyde, C.C.; Gay, R. Smoothed periodogram asymptotics and estimation for processes and fields with possible long-range dependence. *Stoch. Proc. Appl.* **1993**, *45*, 169–182.
6. Hosoya, Y. A limit theory for long-range dependence and statistical inference on related models. *Ann. Stat.* **1997**, *25*, 105–137.
7. Robinson, P.M. Gaussian semiparametric estimation of long-range dependence. *Ann. Stat.* **1995**, *23*, 1630–1661.
8. Labato, I.G. Multivariate Analysis of Long Memory Series in the Frequency Domain. Ph.D. Thesis, University of London, London, UK, 1996.
9. Lobato, I.G. A semiparametric two-step estimation on a multivariate long-memory model. *J. Econom.* **1999**, *90*, 129–153.
10. Nielsen, M.; Shimotsu, K. Determine the cointegration rank in nonstationary fractional systems by the exact local Whittle approach. *J. Econom.* **2007**, *141*, 574–596.
11. Shimotsu, K.; Phillips, P.C.B. Exact local Whittle estimation of fractional integration. *Ann. Stat.* **2005**, *33*, 1890–1933.
12. Hualde, J. A simple test for the equality of integration orders. *Econ. Lett.* **2013**, *119*, 233–237.
13. Marinucci, D.; Robinson, P.M. Weak convergence of multivariate fractional processes. *Stoch. Proc. Appl.* **2000**, *86*, 103–120.
14. Marmol, F.; Velasco, C. Consistent testing of cointegrating relationships. *Econometrica* **2004**, *72*, 1809–1844.
15. Wang, B.; Wang, M.; Chan, N.H. Residual-based tests for fractional cointegrations. *Econ. Lett.* **2015**, *126*, 43–46.

16. Johansen, S.; Nielsen, M. A necessary moment condition for the fractional functional central limit theorem. *Econom. Theory* **2012**, *28*, 671–679.
17. Geweke, J.; Porter-Hudak, S. The estimation and application of long memory time series models. *J. Time Ser. Anal.* **1983**, *4*, 221–238.
18. Hurvich, C.M.; Deo, R.; Brodsky, J. The mean squared error of Geweke and Porter-Hudak's estimates of the memory parameter of a long memory time series. *J. Time Ser. Anal.* **1998**, *19*, 19–46.
19. Künsch, H.R. Statistical aspects of self-similar processes. In *Proceedings of the First World Congress of the Bernoulli Society*; Prokhorov, Y., Sazonov, V.V., Eds.; VNU Science Press: Utrecht, The Netherlands, 1987; Volume 1, pp. 67–74.
20. Beran, J. Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 659–672.
21. Tanaka, K. The nonstationary fractional unit root. *Econom. Theory* **1999**, *15*, 549–582.
22. Hualde, J.; Velasco, C. Distribution-free tests of fractional cointegration. *Econom. Theory* **2008**, *24*, 216–255.
23. Velasco, C.; Robinson, P.M. Whittle pseudo-maximum likelihood estimation for nonstationary time series. *J. Am. Stat. Assoc.* **2000**, *95*, 1229–1243.
24. Robinson, P.M. The distance between rival nonstationary fractionary fractional processes. *J. Econom.* **2005**, *128*, 283–300.
25. Hualde, J.; Robinson, P.M. Semiparametric Estimation of Fractional Cointegration. STICERD-Econometircs Paper Series No. EM/2006/502. 2006. Available online: http://sticerd.lse.ac.uk/dps/em/em502.pdf (accessed on 1 July 2016).
26. Hualde, J.; Robinson, P.M. Semiparametric inference in multivariate fractionally cointegrated systems. *J. Econom.* **2010**, *157*, 492–511.
27. Wang, B. Residual-based tests for fractional cointegrations. Ph.D. Thesis, The Chinese University of Hong Kong, Shatin, Hong Kong, 2008.
28. Hualde, J.; Robinson, P.M. Gaussian pseudo-maximum likelihood estimation of fractional time series models. *Ann. Stat.* **2011**, *39*, 3152–3181.