Fractional integration and cointegration in financial time series
Stakenas, P.

Citation for published version (APA):
Stakenas, P. (2012). Fractional integration and cointegration in financial time series Amsterdam: Thela Thesis

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Chapter 3

Testing for no fractional cointegration: a Monte Carlo study

3.1 Introduction

Cointegration, as a concept of long-run relationship between two (non)stationary time series, was put forward in the early eighties with the seminal paper of Granger (1981). Since then cointegration analysis has become one of the main tools of empirical research in economics and finance. Although the original concept of cointegration did not exclude fractional processes\(^1\), most of studies in the area of cointegration were devoted to the so called \(I(1)/I(0)\) framework. The framework assumes that the observed multivariate time series \(X_t\) is an \(I(1)\) process, but there exists a linear combination \(\beta'X_t\), which is an \(I(0)\) process. However, an increasing number of recent studies analyze the possibility of fractional cointegration, i.e. cointegration between time series with possibly non-integer order of integration. The main question concerning fractionally cointegrated systems is the existence of fractional cointegration itself and, typically, the first step of empirical analysis is either direct estimation of the cointegration rank of an observed multivariate time series or diagnostic testing for cointegration. Tests for cointegration in the classical \(I(1)/I(0)\) framework may still retain consistency in a fractional framework, but they may also entail a considerable loss of power and/or size distortions, if observed time series are indeed fractionally (co)integrated. Hence, new methods have been considered, embedding the classical \(I(1)/I(0)\) case in a fractional framework thus allowing for more accurate inference.

In a parametric time domain setting Breitung and Hassler (2002) proposed a trace

\(^1\)For a formal definition of fractional processes, see Section 1.1.
test for cointegration rank based on a generalized eigenvalue problem, while Hassler and Breitung (2006) derived a residual-based Lagrange multiplier-type test for the null of no cointegration, both studies assuming prior knowledge on the integration order of the observed time series. Nielsen (2005) studied a multivariate Lagrange multiplier (LM) test for fractional integration and showed that the test is also an implicit test for the null of no cointegration. Avarucci and Velasco (2009) proposed a test for the rank of fractional cointegration, which has standard asymptotic properties in weakly cointegrated systems, and reports good finite sample properties. Recently, Lasak (2010) considered a likelihood ratio test for the null of no cointegration making use of a parameterization of cofractional processes due to Granger (1986), which is identified under the alternative, but not under the null. The test statistic has a non-standard asymptotic distribution depending on a nuisance parameter, which may be tabulated by means of Monte Carlo simulations. Johansen and Nielsen (2012) consider a likelihood ratio test for cointegration rank within a parametric cofractional VAR model studied in Johansen (2008), but due to a different parameterization in the general model they do not encounter an identification issue while testing for no cointegration.

Research in semiparametric methods has been equally active. A semiparametric frequency domain procedure was proposed in Nielsen and Shimotsu (2007), where they extended the Robinson and Yajima (2002) approach applying exact local Whittle analysis of Shimotsu and Phillips (2005). Their method relies on estimation of the rank of the spectral density matrix of fractionally filtered observed time series. In a similar setting Chen and Hurvich (2003a) estimated the rank of the averaged periodogram matrix of differenced and tapered observations. Marmol and Velasco (2004) proposed a Wald test for the null of spurious cointegration against weak fractional cointegration. Nielsen (2010) proposed a nonparametric variance ratio test depending on an additional parameter and reported very good finite sample properties.

However, despite considerable amount of studies on testing for fractional cointegration, there has been relatively little attention paid to finite sample comparison of tests. To the best knowledge of the author, Dittmann (2000) is the most recent extensive study and, given considerable advance of the literature in the last decade, there is a need for a study comparing finite sample properties of the newest methods and this chapter provides a Monte Carlo study for that. Since the literature in the field has grown significantly, we limit the scope of the study to testing for fractional cointegration with regression-based time domain methods and study different tests in terms of their size and power properties. We do not provide exact underlying technical assumptions behind the testing procedures, rather briefly discussing their applicability. In fact, the underlying assumptions are not
even satisfied in certain cases of the simulated DGP, but since our primary motivation of this study is to provide guidance for empirical researchers in practical situations, when underlying assumptions are not always possible to check, we do not find it inappropriate and focus on the performance of tests in finite samples.

The main idea of regression-based testing for fractional cointegration is to employ an auxiliary regressor in a correctly specified regression model under the null, and test for its significance. The power of the test naturally depends on the choice of regressor, but it can be chosen ensuring that the alternative is included in the estimated regression model, thus maximizing (asymptotic) power. The key point in the testing is that the regressors depend on the order of integration of observed time series and nearly all methods assume prior knowledge of it, which is certainly restrictive in a fractional framework. Furthermore, it can be shown that arbitrary consistent estimates may affect asymptotic properties of test statistics, see Hassler and Breitung (2006). Although certain semi- and nonparamateric methods may overcome the issue of unknown integration order of observed time series, regression-based methods typically assume knowledge of the integration order and we will also maintain this assumption, since use of an estimate rather than the true value affects all our considered asymptotic test statistics, with the exception of the test of Lasak (2010).

The chapter is structured as follows: Section 3.2 sets up the discussion and reviews tests which will be compared in this Monte Carlo study, Section 3.3 discusses finite sample simulation findings, Section 3.4 provides an empirical application, while Section 3.5 concludes. All tables and pictures are relegated to the Appendix.

### 3.2 Testing for fractional cointegration

In this section we present tests which will be compared in our Monte Carlo study. To start the discussion, consider an $n$-dimensional fractionally integrated time series $X_t$:

$$
\Phi(L) \Delta^d X_t = \epsilon_t, \quad t = 0, 1, \ldots, T,
$$

(3.1)

where $\Phi(L)$ is a matrix lag polynomial of order $k$: $\Phi(L) = I_n - \sum_{i=1}^k \Phi_i L^i$, such that: $\det(\Phi(z)) \neq 0$, $\forall |z| \leq 1$ and $\epsilon_t$ is an i.i.d. $(0, \Sigma)$ process with $\Sigma > 0$. Note, that invertibility of the matrix lag polynomial $\Phi(L)$ and positive definiteness of $\Sigma$ implies that the time series $X_t$ is not cointegrated and thus the model (3.1) represents our null hypothesis of a fractionally integrated, but not cointegrated time series $X_t$. Next consider the alternative
hypothesis, i.e. an $n$-dimensional fractionally cointegrated system:

$$\Delta(d, b)(\beta, \gamma)'X_t = u_t, \; t = 0, 1, \ldots, T,$$

where $\Delta(d, b) = \text{blockdiag}\{\Delta^d_{+}I_r, \Delta^d_{+}I_{n-r}\}$, $I_r$ is an $r \times r$ unit matrix, $b > 0$, $\beta$ is an $n \times r$ matrix, $(\beta, \gamma)$ is an $n \times n$ matrix of full rank and $u_t$ is a VAR($k$) process: $\Psi(L)u_t = \varepsilon_t$, where $\Psi(L)$ has the usual invertibility property.

The system (3.2) represents fractionally cointegrated time series, since $X_t \sim I(d)$, but $\beta'X_t \sim I(d - b)$, and although this is not the only way to parametrize fractionally cointegrated time series (in particular, see Johansen and Nielsen (2012)), we use this framework in our analysis which corresponds to the conceptual definition of fractional cointegration (1.2). Next we show how the model (3.2) relates to the model (3.1) in the framework of regression-based testing for cointegration. We rewrite (3.2) as:

$$\begin{pmatrix} \beta, \gamma \end{pmatrix}'\Delta^d_{+}X_t = \begin{pmatrix} -I_r \\ 0_{n-r \times r} \end{pmatrix} \beta' (1 - \Delta^d_{+}b) \Delta^d_{+}X_t + u_t.$$

From here:

$$\Psi(L)(\beta, \gamma)'\Delta^d_{+}X_t = \Psi(L) \begin{pmatrix} -I_r \\ 0_{n-r \times r} \end{pmatrix} \beta' (1 - \Delta^d_{+}b) \Delta^d_{+}X_t + \varepsilon_t,$$

and, after rearranging terms, we have:

$$\Delta^d_{+}X_t = \Pi_0 Z_{t-1} + \sum_{i=1}^{k} \Pi_i \Delta Z_{t-i} + \sum_{i=1}^{k} \Phi_i \Delta^d_{+}X_{t-i} + \varepsilon_t,$$

where $Z_{t-1} = (\Delta^{d-b}_{+} - \Delta^d_{+})X_t$. The matrices $\Phi_i$ and $\Pi_i$ have a multiplicative structure implied by (3.3), but in the following we consider them unrestricted thus keeping the discussion simpler. We see that the model (3.2) for cointegrated time series can be formulated within a regression framework:

$$\Delta^d_{+}X_t = \Pi W_t + \sum_{i=1}^{k} \Phi_i \Delta^d_{+}X_{t-i} + \varepsilon_t,$$

with $W_t = (Z_{t-1}', \Delta Z_{t-1}', \ldots, \Delta Z_{t-k}')'$ and $\Pi = (\Pi_0, \ldots, \Pi_k)$. Note, that with $\Pi = 0$, the model (3.5) reduces to (3.1) and represents fractionally integrated, but not cointegrated time series. The fact that the triangular fractionally cointegrated system (3.2) can be rewritten as a regression (3.5) gives motivation for the idea to test the null of fractional
3.2 Testing for fractional cointegration

integration against the alternative of fractional cointegration with the test for the null $\Pi = 0$ in (3.5) using some $\mathcal{F}_t$-adapted auxiliary regressor $W_t$, where $\mathcal{F}_t$ is a $\sigma$-field generated by a collection of random variables $\{\varepsilon_i, i \leq t\}$. However, note that the null of no cointegration (3.1) implies $b = 0$ and the regressor $Z_{t-1}$ (and thus $W_t$) vanishes in (3.5), and in order to make the regressor continuous at $b = 0$. Avarucci and Velasco (2009) proposed to use a rescaled regressor: $Z_{t-1}^* = \frac{\Delta_t b^{-1}}{b} \Delta_t^d X_t$, where indetermination 0/0 is solved using L'Hôpital’s rule: $\lim_{b \to 0} \frac{\Delta_t b^{-1}}{b} = \ln \Delta$ (see also Lobato and Velasco (2007)). Then the null hypothesis of the test for no cointegration against alternative of fractional cointegration can be formulated as a test for the null $\Pi = 0$ against the alternative $\Pi \neq 0$ in (3.5) with $W_t = (Z_{t-1}^*, \Delta Z_{t-1}^*, \ldots, \Delta Z_{t-k}^*)'$ as an auxiliary regressor. Avarucci and Velasco (2009) considered a Wald test for the subset of restrictions on $\Pi = 0$, namely, $\Pi_0 = 0$ in the regression (3.5) with statistic:

$$\Lambda_\Pi(b) = tr(S_{00}^{-1} S_{10}' S_{11}^{-1} S_{10})$$

where $S_{ij}$ are the sample moments: $S_{ij} = S_{ij}(d) = T^{-1} \sum_{t=k}^T R_d R_t'$ and $R_{0t}$, $R_{1t}$ are $\Delta_t^d X_t$, $Z_{t-1}$ regressed on $\Delta Z_{t-1}^*, \ldots, \Delta Z_{t-k}^*, \Delta_t^d X_{t-1}, \ldots, \Delta_t^d X_{t-k}$, respectively. They showed that under the the null it holds: $\Lambda_\Pi(\hat{b}) \xrightarrow{d} \chi^2_{n2}$, for $\hat{b} \xrightarrow{p} b$ and $b \in [0, 0.5)$.

Finite sample simulations of Avarucci and Velasco (2009) showed that disregarding the terms $\Delta Z_{t-i}^*, i = 1, \ldots, k$ in the regression (3.5) altogether made little difference in the performance of the test. The test statistic has standard asymptotics for “weak cointegration”, i.e. $b < 0.5$, whereas for $b > 0.5$ it has a non-pivotal test statistic depending on the parameter $b$. The case $d \geq b > 0.5$ has been considered by Lasak (2010). She considered regression (3.5) with $W_t = Z_{t-1}^*$ and showed that under the null $\Pi = 0$ it holds:

$$\sup_{b \in D_b} \Lambda_\Pi(b) \xrightarrow{d} \sup_{b \in D_b} tr \left( \int_0^1 dB B_{b-1}(s)' \left( \int_0^1 B_{b-1}(s) B_{b-1}(s)' ds \right)^{-1} \int_0^1 B_{b-1}(s) dB' \right).$$

Here $B_b(s)$ is $n$-dimensional standard type II fractional Brownian motion. We see that under the null the statistic $\sup_{b \in D_b} \Lambda_\Pi(b)$ converges to a non-standard distribution, which depends on the compact parameter set $D_b = [0.5 + \epsilon, d]$ acting as a nuisance parameter. In principle the test statistic has to be tabulated for every $d$ and it was done in Lasak (2010) for the value $d = 1$. Note, that with the value $D_b = \{1\}$, the test statistic $\sup_{b \in D_b} \Lambda_\Pi(b)$

\footnote{Note that assumption $b > 0$ is needed to identify the parameters of the model (3.2), but it is not needed in (3.5) with rescaled regressors.}

\footnote{Taking $\epsilon$ very small, such as $\epsilon = 0.01$, eliminates dependence on $\epsilon$ in practice.}
is the familiar Johansen’s trace statistic testing for cointegration in \( I(1)/I(0) \) systems.

Breitung and Hassler (2002) derived a variant of the score test with an auxiliary regressor \( W_t = \ln \Delta_+ \Delta^d_t X_t \) in the regression (3.5) and showed that the test statistic \( \Lambda_{\Pi} \) under the null has standard asymptotic properties: \( \Lambda_{\Pi} \xrightarrow{d} \chi^2_{2n} \), while under alternative of cointegration it diverges to infinity at a rate \( T \). The choice of auxiliary regressor was motivated by local optimality properties of the test for \( d \) in the univariate case (cf. Robinson (1994), Tanaka (1999)) and can be further supported by the fact that \( W_t = \ln \Delta_+ \Delta^d_t X_t \) can be seen as the limit of the regressor \( Z^*_t \) as \( b \to 0 \).

Hassler and Breitung (2006) derived a residual-based LM-test against fractional cointegration, which does not belong to the framework (3.5), but is based on a two-step procedure. Denote the observed time series as: \( X_t = (X_{1t}, X_{2t})' \), where \( X_{1t} \) is a scalar series, and suppose \( X_t \) is generated by (3.1). Then define \( \eta_t \) as \( \Delta^d_t X_{1t} \) regressed on \( \Delta^d_t X_{2t}, \Delta^d_t X_{t-i}, \) \( i = 1, \ldots, k \) and consider OLS regression:

\[
\eta_t = \phi \eta^*_t + \sum_{i=1}^{k} \Psi_i' \Delta^d_t X_{t-i} + \varepsilon_t, \tag{3.6}
\]

where \( \eta^*_t = \sum_{j=1}^{t} j^{-1} \eta_{t-j} \) and \( \Psi_i \) are \( n \times 1 \) parameter vectors. Under the null of no cointegration, \( t \)-statistic for the parameter \( \phi \) converges to the standard normal random variable: \( t_{\phi} \xrightarrow{d} N(0,1) \).

### 3.3 Finite sample performance

In this section we discuss design of Monte Carlo finite sample simulations and present simulation results.

#### 3.3.1 Monte Carlo setup

We simulate the bivariate fractionally cointegrated model (3.2) assuming a triangular cointegration structure\(^4\) with \( \beta' = (1, -\alpha) \) and \( \gamma' = (0, 1) \):

\[
X_{1t} = \alpha X_{2t} + \Delta^{-(d-b)}_+ u_{1t}, \\
X_{2t} = \Delta^d_+ u_{2t},
\]

\(^4\)Note that due to triangular structure, cointegration models with “nearly” orthogonal cointegration and adjustment vectors are not considered.
3.3 Finite sample performance

for \( t = 1, \ldots, T \). In the simulations we assume that \( u_t \) is a bivariate VAR(1) process: 
\[
\Phi(L)u_t = \varepsilon_t \quad \text{with} \quad \Phi(L) = (1 - \phi L)I_2 \quad \text{and} \quad \varepsilon_t \quad \text{is an i.i.d. Gaussian} \ (0, \Omega) \quad \text{process with} \quad \Omega = ((1, \rho)', (\rho, 1)') .
\]
We simulate this system with the values \( \alpha = 1, d = 1, T = 100, 250, 500, b = 0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.85, 1, \rho = 0, 0.5, 0.8, \phi = -0.5, 0, 0.5, 0.8 \). A number of \( R = 10000 \) Monte Carlo replications was used.

As it was mentioned above, the order of integration \( d = 1 \) is assumed to be known, while the number of lags in regressions will be selected in two different ways: infeasibly, taking \( k = 1 \) and feasibly, where the number of lags is selected using the Akaike information criterion in the regression (3.1).

Additionally, we have also simulated performance of the tests with conditionally heteroscedastic error term \( \varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})' \) with the following error generating mechanism: 
\[
\varepsilon_{it} = \sigma_{it} \eta_t, \quad \sigma_{it}^2 + 1 = 1 + 0.15 \varepsilon_{it}^2 + 0.8 \sigma_{it}^2 \quad \text{for} \quad i = 1, 2 \quad \text{and} \quad \eta_t = (\eta_{1t}, \eta_{2t})' \quad \text{is normally distributed i.i.d.} \ (0, 1) \quad \text{time series.}
\]
The parameter values of GARCH(1,1) univariate processes are typical in empirical applications.

### 3.3.2 Simulation results

In the following, we will use the following abbreviations for the test statistics: LAS for the test of Lasak (2010), AV for the test of Avarucci and Velasco (2009)
\footnote{In the test \( b \) was estimated with exact local Whittle estimator (2.18) with bandwidth \( m = [T^{0.65}] \) as integration order of \( \hat{\beta}X_t \), where \( \hat{\beta} \) is the estimate of the cointegration vector with OLS regression in levels of \( X_t \).}, HB for the test of Hassler and Breitung (2006) and BH for the test of Breitung and Hassler (2002). Firstly, we discuss simulation results concerning empirical size of the tests. Note, that under the null hypothesis of no cointegration, the test statistics BH, AV and LAS are invariant to the contemporaneous correlation parameter \( \rho \), hence the empirical size of the tests does not depend on \( \rho \), unlike power of the tests since under the alternative the tests are not invariant to \( \rho \). Since the HB test showed little sensitivity to the parameter \( \rho \), we report results for the size only with the value \( \rho = 0 \). We begin with the infeasible case \( k = 1 \) (Table 3.1). HB and BH show very good size properties even with small sample size \( T = 100 \) (although BH is slightly oversized with \( \phi = 0.5 \)). LAS also has good size properties, although in case \( \phi = 0.8 \) with \( T = 100 \) and \( T = 250 \) the test is slightly oversized, but increasing the sample size to \( T = 500 \) remedies the problem. The empirical size of AV seems to be the most distorted, being rather oversized with sample size \( T = 100 \), but increasing the sample size to \( T = 250 \) solves the problem to a certain extent, and with \( T = 500 \) the sizes are already close to the nominal levels. Table 3.2 shows empirical size for the feasible case, \( k = \hat{k}_{AIC} \). In this case, HB and BH are generally undersized and
even an increase of sample size to $T = 500$ does not seem to improve the performance. LAS again has good size properties and although the case $T = 100, \phi = 0.8$ shows some oversize, increasing the sample size corrects for that. AV shows a somewhat mixed picture: although oversized with $T = 100$ (except with $\phi = 0$), the test is undersized with $T = 500$, hence it does not seem that increasing the sample size solves the size distortion issue.

In addition, we investigate robustness of the tests against conditional heteroscedasticity simulating error terms as GARCH(1, 1) processes. In the case $k = 1$ (Table 3.3), the tests are rather oversized and increasing the sample size does not bring the empirical levels closer to the nominal ones for BH, HB and AV, thus indicating that these tests are not robust to conditional heteroscedasticity. Generally, the size of LAS is least distorted and it also improves with increasing the sample size, unlike other tests. The feasible case $k = \hat{k}_{AIC}$ (Table 3.4) shows less distortion in size for all tests and it may be explained by the fact that IC-based lag selection rules on average select more lags in the regression hence decreasing rejection levels, as it was also seen to some degree with homoscedastic errors.

Next we discuss power of the tests. We begin with the infeasible case with $k = 1$ number of lags in the test regressions and give a few remarks about the performance of tests. Generally, bigger persistence in the short-run noise $u_t$ decreases the power of tests. One of the possible explanations is that since the tests are not consistent against cointegration with the boundary value $\phi = 1$, getting closer to this value decreases the power. In addition, the power of BH, AV and LAS is affected by the contemporaneous correlation parameter $\rho$: the power of AV, LAS increases with an increase of $\rho$, while that of BH decreases. The HB test displays little sensitivity to $\rho$. Generally, the power of the tests increases significantly with increasing the sample size $T$, except the case $\rho = 0.8$ for HB test: with $T = 100$ all tests have effectively no power with $b < 0.5$, although the power improves with increase in sample size, but is nonetheless negligible for HB. It seems that in case of strong contemporaneous correlation, the power of the HB test increases (very) slowly with $T$.

Next we compare tests on a case-by-case basis. In the infeasible case with parameter values $(\rho, \phi) = \{0, 0.5\} \times \{-0.5, 0, 0.5\}$, the HB test is the all-out winner, however with $\rho = 0.8$ or $\phi = 0.8$, the picture is less clear. It is difficult to discern the best performing test in these cases, but one might say that in the cases $(\rho, \phi) = 0.8 \times \{-0.5, 0, 0.5\}$, LAS performs best, while in case $(\rho, \phi) = (0.8, 0.8)$ against alternatives of strong cointegration LAS still

---

6We have simulated both size-adjusted and unadjusted power, but in the discussion we focus on the size-adjusted power. Size-unadjusted power results are available upon request, however they are not comparatively much different.
3.3 Finite sample performance

wins, but against alternatives of weak cointegration BH is better with $(\rho, \phi) = \{0, 0.5\} \times 0.8$ and HB is better with $\rho = 0, \phi = 0.8$.

In the feasible case, the power of all tests has decreased, which is especially noticeable for the HB and BH tests against alternatives of strong cointegration. Despite that, for cases $(\rho, \phi) = \{0, 0.5\} \times \{-0.5, 0, 0.5, 0.8\}$, HB test is still the best against alternatives of weak cointegration, but against strong cointegration LAS seems to be the best. In cases $(\rho, \phi) = 0.8 \times \{-0.5, 0, 0.5, 0.8\}$ and $T = 250, 500$, BH is the best against weak cointegration, while LAS is again best against strong cointegration. In the small sample case $T = 100$ and $\rho = 0.8$, LAS and AV are performing best.

Finally, we discuss certain variants of LAS and AV tests. As was mentioned above, Johansen’s trace test can be regarded as a variant of LAS test with $D_b = 1$. It is not straightforward that LAS has more power than Johansen’s test, since in both cases the DGP under alternative (3.2) is not included in the estimated model, unless $b = 1$ or $k = 0$. Simulations showed that LAS is yet more powerful in all cases considered, although in the case $\phi = 0.8$ the difference is rather negligible. In addition, the trace test generally performed worse than HB, except against strong cointegration in certain cases. This underlines gains in efficiency when using tests for fractional cointegration in a fractional framework.

We have also compared two variants of the Avarucci and Velasco (2009) test: above-discussed where $b$ was estimated and one with fixed value of $b$ set to $b = 0.25$. AV with estimated $b$ seems to be a slightly more powerful with $T = 100, 250$ and against strong cointegration, but with $T = 500$ the difference in performance effectively disappears and becomes negligible. Hence, in large sample one might use the AV test with arbitrary chosen value of $b$ rather than estimated without much loss in efficiency as our simulations reveal.

We have also simulated power of the tests with conditional heteroscedastic error term $\varepsilon_t$. Generally, the adjusted power of tests decreased somewhat, but it showed little comparative difference between tests in terms of power performance in comparison to the i.i.d. setup, hence we do not report these results. In addition, we have also simulated i.i.d. errors $\varepsilon_t$ having student’s t-distribution with 3 d.f., which also showed little comparative difference to the results above.

Finally, we condense the simulation results into a few guidelines for empirical researchers using regression-based tests for fractional cointegration: i) LAS performs best in terms of size properties in the feasible case, but if the lag structure is known, other tests also show very good size properties; ii) LAS test works best against strong cointegration, while BH and HB are best against weak cointegration; iii) in case of conditional heteroscedasticity
Testing for no fractional cointegration: a Monte Carlo study

Table 3.1: Empirical size with $k = 1$, $\rho = 0$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$\phi$</th>
<th>BH 10%</th>
<th>BH 5%</th>
<th>BH 1%</th>
<th>LAS 10%</th>
<th>LAS 5%</th>
<th>LAS 1%</th>
<th>AV 10%</th>
<th>AV 5%</th>
<th>AV 1%</th>
<th>HB 10%</th>
<th>HB 5%</th>
<th>HB 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-0.5</td>
<td>0.104</td>
<td>0.050</td>
<td>0.008</td>
<td>0.104</td>
<td>0.052</td>
<td>0.011</td>
<td>0.134</td>
<td>0.073</td>
<td>0.015</td>
<td>0.101</td>
<td>0.050</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.109</td>
<td>0.056</td>
<td>0.009</td>
<td>0.104</td>
<td>0.054</td>
<td>0.010</td>
<td>0.119</td>
<td>0.062</td>
<td>0.012</td>
<td>0.096</td>
<td>0.048</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.126</td>
<td>0.065</td>
<td>0.014</td>
<td>0.113</td>
<td>0.061</td>
<td>0.013</td>
<td>0.122</td>
<td>0.064</td>
<td>0.013</td>
<td>0.104</td>
<td>0.048</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.102</td>
<td>0.050</td>
<td>0.009</td>
<td>0.148</td>
<td>0.077</td>
<td>0.021</td>
<td>0.140</td>
<td>0.076</td>
<td>0.016</td>
<td>0.100</td>
<td>0.049</td>
<td>0.009</td>
</tr>
<tr>
<td>250</td>
<td>-0.5</td>
<td>0.098</td>
<td>0.048</td>
<td>0.009</td>
<td>0.104</td>
<td>0.051</td>
<td>0.013</td>
<td>0.118</td>
<td>0.062</td>
<td>0.016</td>
<td>0.099</td>
<td>0.046</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.104</td>
<td>0.054</td>
<td>0.012</td>
<td>0.102</td>
<td>0.053</td>
<td>0.013</td>
<td>0.112</td>
<td>0.059</td>
<td>0.014</td>
<td>0.097</td>
<td>0.049</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.113</td>
<td>0.059</td>
<td>0.013</td>
<td>0.107</td>
<td>0.054</td>
<td>0.013</td>
<td>0.111</td>
<td>0.059</td>
<td>0.015</td>
<td>0.095</td>
<td>0.046</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.099</td>
<td>0.049</td>
<td>0.010</td>
<td>0.121</td>
<td>0.066</td>
<td>0.014</td>
<td>0.122</td>
<td>0.065</td>
<td>0.015</td>
<td>0.096</td>
<td>0.047</td>
<td>0.009</td>
</tr>
<tr>
<td>500</td>
<td>-0.5</td>
<td>0.096</td>
<td>0.049</td>
<td>0.009</td>
<td>0.105</td>
<td>0.051</td>
<td>0.011</td>
<td>0.111</td>
<td>0.057</td>
<td>0.011</td>
<td>0.095</td>
<td>0.046</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.104</td>
<td>0.054</td>
<td>0.012</td>
<td>0.102</td>
<td>0.051</td>
<td>0.012</td>
<td>0.107</td>
<td>0.053</td>
<td>0.011</td>
<td>0.092</td>
<td>0.044</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.107</td>
<td>0.053</td>
<td>0.011</td>
<td>0.104</td>
<td>0.050</td>
<td>0.012</td>
<td>0.100</td>
<td>0.049</td>
<td>0.010</td>
<td>0.091</td>
<td>0.042</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.094</td>
<td>0.046</td>
<td>0.008</td>
<td>0.108</td>
<td>0.054</td>
<td>0.012</td>
<td>0.104</td>
<td>0.052</td>
<td>0.010</td>
<td>0.100</td>
<td>0.049</td>
<td>0.008</td>
</tr>
</tbody>
</table>

Table 3.2: Empirical size with $k = k_{AIC}$, $\rho = 0$

<table>
<thead>
<tr>
<th>$T$</th>
<th>$\phi$</th>
<th>BH 10%</th>
<th>BH 5%</th>
<th>BH 1%</th>
<th>LAS 10%</th>
<th>LAS 5%</th>
<th>LAS 1%</th>
<th>AV 10%</th>
<th>AV 5%</th>
<th>AV 1%</th>
<th>HB 10%</th>
<th>HB 5%</th>
<th>HB 1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-0.5</td>
<td>0.061</td>
<td>0.022</td>
<td>0.001</td>
<td>0.094</td>
<td>0.044</td>
<td>0.007</td>
<td>0.135</td>
<td>0.073</td>
<td>0.017</td>
<td>0.078</td>
<td>0.033</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.055</td>
<td>0.018</td>
<td>0.001</td>
<td>0.090</td>
<td>0.040</td>
<td>0.006</td>
<td>0.077</td>
<td>0.034</td>
<td>0.006</td>
<td>0.077</td>
<td>0.035</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.109</td>
<td>0.054</td>
<td>0.010</td>
<td>0.102</td>
<td>0.054</td>
<td>0.010</td>
<td>0.110</td>
<td>0.055</td>
<td>0.012</td>
<td>0.096</td>
<td>0.043</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.063</td>
<td>0.024</td>
<td>0.002</td>
<td>0.136</td>
<td>0.067</td>
<td>0.016</td>
<td>0.122</td>
<td>0.064</td>
<td>0.015</td>
<td>0.079</td>
<td>0.034</td>
<td>0.005</td>
</tr>
<tr>
<td>250</td>
<td>-0.5</td>
<td>0.061</td>
<td>0.024</td>
<td>0.003</td>
<td>0.095</td>
<td>0.046</td>
<td>0.012</td>
<td>0.110</td>
<td>0.058</td>
<td>0.014</td>
<td>0.079</td>
<td>0.033</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.058</td>
<td>0.020</td>
<td>0.002</td>
<td>0.093</td>
<td>0.044</td>
<td>0.010</td>
<td>0.070</td>
<td>0.028</td>
<td>0.004</td>
<td>0.076</td>
<td>0.032</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.103</td>
<td>0.052</td>
<td>0.010</td>
<td>0.099</td>
<td>0.051</td>
<td>0.011</td>
<td>0.089</td>
<td>0.042</td>
<td>0.009</td>
<td>0.089</td>
<td>0.042</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.061</td>
<td>0.024</td>
<td>0.003</td>
<td>0.115</td>
<td>0.063</td>
<td>0.013</td>
<td>0.095</td>
<td>0.044</td>
<td>0.008</td>
<td>0.078</td>
<td>0.034</td>
<td>0.004</td>
</tr>
<tr>
<td>500</td>
<td>-0.5</td>
<td>0.061</td>
<td>0.027</td>
<td>0.003</td>
<td>0.100</td>
<td>0.048</td>
<td>0.010</td>
<td>0.103</td>
<td>0.052</td>
<td>0.008</td>
<td>0.079</td>
<td>0.035</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.057</td>
<td>0.021</td>
<td>0.001</td>
<td>0.097</td>
<td>0.049</td>
<td>0.011</td>
<td>0.067</td>
<td>0.026</td>
<td>0.003</td>
<td>0.074</td>
<td>0.033</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.102</td>
<td>0.049</td>
<td>0.010</td>
<td>0.101</td>
<td>0.048</td>
<td>0.011</td>
<td>0.083</td>
<td>0.039</td>
<td>0.007</td>
<td>0.087</td>
<td>0.040</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.059</td>
<td>0.024</td>
<td>0.002</td>
<td>0.106</td>
<td>0.052</td>
<td>0.012</td>
<td>0.081</td>
<td>0.037</td>
<td>0.005</td>
<td>0.081</td>
<td>0.036</td>
<td>0.005</td>
</tr>
</tbody>
</table>

LAS is the only test which has relatively undistorted empirical sizes.
3.4 Empirical application

We proceed with an empirical application using the dataset described in Section 1.3. In Section 2.5 we did not reject null hypothesis, that interest rate series are integrated of order $d = 1$, whereas the expectations hypothesis postulates that if individual interest rate series are $I(1)$ processes, they cointegrate. We will test this hypothesis (or more precisely, the hypothesis of no cointegration against alternative of cointegration) with tests AV, LAS, BH and HB in the bivariate systems $(M_1, M_3)$ and $(M_1, M_6)$ assuming that interest rate series are $I(1)$ processes. We select the number of lags used in regressions by inspecting the correlograms and cross-correlograms of the relevant bivariate VAR models: it turned out that including autoregressive terms of 1-st and 12-th order in bivariate VAR’s yields uncorrelated residuals in the models. Hence, in the test regressions we include two

<table>
<thead>
<tr>
<th>$T$</th>
<th>$\phi$</th>
<th>BH</th>
<th>LAS</th>
<th>AV</th>
<th>HB</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-0.5</td>
<td>0.172</td>
<td>0.097</td>
<td>0.024</td>
<td>0.156</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.174</td>
<td>0.099</td>
<td>0.024</td>
<td>0.160</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.178</td>
<td>0.101</td>
<td>0.026</td>
<td>0.174</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.176</td>
<td>0.100</td>
<td>0.027</td>
<td>0.220</td>
</tr>
<tr>
<td>250</td>
<td>-0.5</td>
<td>0.180</td>
<td>0.107</td>
<td>0.032</td>
<td>0.142</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.176</td>
<td>0.103</td>
<td>0.031</td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.159</td>
<td>0.089</td>
<td>0.024</td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.170</td>
<td>0.102</td>
<td>0.029</td>
<td>0.159</td>
</tr>
<tr>
<td>500</td>
<td>-0.5</td>
<td>0.192</td>
<td>0.121</td>
<td>0.040</td>
<td>0.135</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.180</td>
<td>0.105</td>
<td>0.033</td>
<td>0.135</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.163</td>
<td>0.095</td>
<td>0.026</td>
<td>0.134</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.184</td>
<td>0.107</td>
<td>0.032</td>
<td>0.134</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$</th>
<th>$\phi$</th>
<th>BH</th>
<th>LAS</th>
<th>AV</th>
<th>HB</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-0.5</td>
<td>0.089</td>
<td>0.037</td>
<td>0.004</td>
<td>0.134</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.078</td>
<td>0.033</td>
<td>0.004</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.141</td>
<td>0.070</td>
<td>0.013</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.099</td>
<td>0.044</td>
<td>0.007</td>
<td>0.194</td>
</tr>
<tr>
<td>250</td>
<td>-0.5</td>
<td>0.095</td>
<td>0.044</td>
<td>0.005</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.088</td>
<td>0.038</td>
<td>0.005</td>
<td>0.123</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.126</td>
<td>0.068</td>
<td>0.015</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.087</td>
<td>0.036</td>
<td>0.006</td>
<td>0.143</td>
</tr>
<tr>
<td>500</td>
<td>-0.5</td>
<td>0.100</td>
<td>0.046</td>
<td>0.007</td>
<td>0.119</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.098</td>
<td>0.042</td>
<td>0.006</td>
<td>0.120</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.136</td>
<td>0.073</td>
<td>0.017</td>
<td>0.123</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>0.092</td>
<td>0.042</td>
<td>0.006</td>
<td>0.126</td>
</tr>
</tbody>
</table>

3.4 Empirical application

We proceed with an empirical application using the dataset described in Section 1.3. In Section 2.5 we did not reject null hypothesis, that interest rate series are integrated of order $d = 1$, whereas the expectations hypothesis postulates that if individual interest rate series are $I(1)$ processes, they cointegrate. We will test this hypothesis (or more precisely, the hypothesis of no cointegration against alternative of cointegration) with tests AV, LAS, BH and HB in the bivariate systems $(M_1, M_3)$ and $(M_1, M_6)$ assuming that interest rate series are $I(1)$ processes. We select the number of lags used in regressions by inspecting the correlograms and cross-correlograms of the relevant bivariate VAR models: it turned out that including autoregressive terms of 1-st and 12-th order in bivariate VAR’s yields uncorrelated residuals in the models. Hence, in the test regressions we include two
Testing for no fractional cointegration: a Monte Carlo study

Table 3.5: Test statistics for testing the null of no cointegration

<table>
<thead>
<tr>
<th></th>
<th>LAS</th>
<th>HB</th>
<th>BH</th>
<th>AV</th>
</tr>
</thead>
<tbody>
<tr>
<td>(M1,M3)</td>
<td>52.09(&lt;0.01)</td>
<td>2.36(0.12)</td>
<td>42.88(&lt;0.01)</td>
<td>43.16(&lt;0.01)</td>
</tr>
<tr>
<td>(M1,M6)</td>
<td>39.08(&lt;0.01)</td>
<td>1.27(0.25)</td>
<td>27.14(&lt;0.01)</td>
<td>31.96(&lt;0.01)</td>
</tr>
</tbody>
</table>

Note: p-values are given in the parenthesis.

autoregressive terms of first and twelfth order. Results for all four tests are reported in the Table 3.5.

We see that all tests reject the null hypothesis with p-values less than 0.01, except the HB test, which does not reject the null hypothesis of no cointegration with p-values 0.12 and 0.25 for the systems (M1, M3), (M1, M6), respectively. To get a better picture of the strength of possible cointegration we estimate the integration orders of $\hat{\beta}X_t$ with the local Whittle estimator (cf. Robinson (1995)) with bandwidth $m = [T^{0.55}]$, where $\hat{\beta}$ is the OLS estimate of the (possible) cointegration vector. In case of cointegration, $\hat{\beta}X_t \sim I(1 - b)$, where $b$ is the strength of cointegration, whereas in the absence of cointegration: $\hat{\beta}X_t \sim I(1)$. We get ELW estimates of 0.54 and 0.77 for the systems (M1, M6), (M1, M3), respectively, which points to a strong degree of cointegration. Of course, this cannot be considered a formal procedure, but one may suppose that the low power of the HB test in strong cointegration cases, as seen in the simulations, could be the reason behind the results we get. Hence, maintaining a degree of caution, we conclude that both bivariate systems are cointegrated.

3.5 Conclusions

This chapter provided a Monte Carlo simulation study, where a few regression-based tests for fractional cointegration were compared. A number of DGP’s were considered along with two different lag selection rules and few guidelines for practitioners were given. An empirical analysis of U.S. interest rate series was also carried out and it was concluded that the bivariate time series are cointegrated.
3.5 Conclusions

(a) \( T = 100, \ k = k_{1/c} \)

(b) \( T = 100, \ k = k_{k} \)

(c) \( T = 500, \ k = k_{1/c} \)

(d) \( T = 500, \ k = k_{k} \)

Figure 3.1: Size-adjusted power curves with \( \rho = 0, \phi = -0.5 \)

(e) \( T = 500, \ k = 1 \)
Figure 3.2: Size-adjusted power curves with $\rho = 0$, $\phi = 0$. 

(a) $T = 100, k = 1$
(b) $T = 100, k = k_{ARC}$
(c) $T = 500, k = 1$
(d) $T = 500, k = k_{ARC}$
3.5 Conclusions

(a) $T = 100, k = 1$

(b) $T = 100, k = k_{1/2}$

(c) $T = 500, k = 1$

(d) $T = 500, k = k_{1/2}$

Figure 3.3: Size-adjusted power curves with $\rho = 0, \phi = 0.5$
Testing for no fractional cointegration: a Monte Carlo study

Figure 3.4: Size-adjusted power curves with $\rho = 0, \phi = 0.8$

(c) $T = 500, k = 1$

(d) $T = 500, k = k_{AIC}$

(a) $T = 100, k = 1$

(b) $T = 100, k = k_{AIC}$
3.5 Conclusions

Figure 3.5: Size-adjusted power curves with \( \rho = 0.5, \phi = -0.5 \).
Testing for no fractional cointegration: a Monte Carlo study

Figure 3.6: Size-adjusted power curves with $\rho = 0.5$, $\phi = 0$

(c) $T = 500, k = 1$

(a) $T = 100, k = 1$

(b) $T = 500, k = k_{\text{aic}}$

(d) $T = 100, k = k_{\text{aic}}$
Figure 3.7: Size-adjusted power curves with $\rho = 0.5$, $\phi = 0.5$
Figure 3.8: Size-adjusted power curves with $\rho = 0.5$, $\phi = 0.8$

(a) $T = 100$, $k = 1$

(b) $T = 500$, $k = k_{AIC}$

(c) $T = 500$, $k = 1$

(d) $T = 100$, $k = k_{AIC}$
3.5 Conclusions

(a) $T = 100$, $k = 1$

(b) $T = 100$, $k = k_{AIC}$

(c) $T = 500$, $k = 1$

(d) $T = 500$, $k = k_{AIC}$

Figure 3.9: Size-adjusted power curves with $\rho = 0.8$, $\phi = -0.5$
Testing for no fractional cointegration: a Monte Carlo study

Figure 3.10: Size-adjusted power curves with $\rho = 0.8$, $\phi = 0$

(a) $T = 100$, $k = 1$

(b) $T = 500$, $k = 1$

(c) $T = 500$, $k = 1$

(d) $T = 100$, $k = k_{ARC}$
3.5 Conclusions

Figure 3.11: Size-adjusted power curves with $\rho = 0.8, \phi = 0.5$
Figure 3.12: Size-adjusted power curves with $\rho = 0.8$, $\phi = 0.8$

(a) $T = 100$, $k = 1$
(b) $T = 100$, $k = k_{AIC}$
(c) $T = 500$, $k = 1$
(d) $T = 500$, $k = k_{AIC}$