Monitoring Breaks in Fractional Cointegration

Maik Dierkes, Krischan Fitter*and Philipp Sibbertsen

November 7, 2024

Abstract

We extend the monitoring of structural breaks in classic cointegration proposed by Wagner and Wied (2017) to explicitly allow for fractional cointegration and breaks in these fractional relations with possible deterministic trends. To estimate the parameters we use a fully modified OLS estimator and we estimate the integration order by the exact local whittle. In order to build the test statistic we establish a CUSUM test for a break in parameters or a break in the order of integration and derive the limiting distribution of the cumulative sum of the modified OLS residuals by using representations by Davidson and Hashimzade (2009) and Fox and Taqqu (1987). Using these limiting results we propose a detector and its limiting distribution as a function of fractional Brownian motions and prove the consistency of our procedure against fixed and local alternatives. The critical values for the monitoring are derived by bootstrap. In a Monte-Carlo study we show the finite sample behavior of our test and compare it to the one by Wagner and Wied (2017) in different scenarios of fractional cointegration. To conclude we show the applicability of the test by presenting the results of applying the test in the context of momentum investing.

Keywords: long-memory time series; fractional cointegration; structural change; monitoring

JEL Classification: C32; C12; C52

1 Introduction

This paper presents a modified residual based approach to monitor fractional cointegration relationships between fractionally integrated time series and hereby extends the monitoring procedure presented by Wagner and Wied (2017) to a more general framework. Similar to them we propose a fractional cointegration regression, which allows for deterministic trends. By doing so we can distinguish two distinct types of structural breaks. First, there may be changes in the fractional cointegration relationship. This could mean that the reduction of the order of integration changes or even vanishes so that the relation turns spurious. Second, there may be a break in the deterministic components, which induces a change in trend or a break in the linear coefficients of the regression.

*Corresponding author: **Krischan Fitter**, Leibniz University Hannover, Institute of Statistics, Königsworther Platz 1 30167 Hannover, Germany, E-Mail: fitter@statistik.uni-hannover.de

Maik Dierkes, Leibniz University Hannover, Institute of Banking and Finance, Königsworther Platz 1 30167 Hannover, Germany, E-Mail: maik.dierkes@finance.uni-hannover.de

Philipp Sibbertsen, Leibniz University Hannover, Institute of Statistics, Königsworther Platz 1 30167 Hannover, Germany, E-Mail: sibbertsen@statistik.uni-hannover.de

Monitoring is a type of sequential test for detecting structural breaks in time series and regressions that can be used repeatedly and consistently every time new data is available. The need for monitoring procedures in comparison to usual structural break tests that are applied to a given historical data set is pointed out e.g. by Chu et al. (1996). They show in a simulation study that if one applies a structural break test repeatedly after a few new data points are available, the test will reject the null of stability with probability converging to one even in case of stability. In applied work it is often the case that one faces new data at some frequency. One example could be daily financial time series where a new data point is available each trading day. In theses cases monitoring methods show their high value in terms of governing structural breaks on a daily basis.

The most general principles of monitoring procedures are presented by Chu et al. (1996). Here, the calculation of the parameters in the regression happens in a so-called *calibration* period at the beginning of the sample, rather than estimating the parameters based on the whole available sample. The calibration period is assumed to be free of structural breaks. This needs to hold true to obtain a set of parameters and residuals as a starting input for the detector. The detector can then either be based on a CUSUM-type test statistic, that monitors the cumulative sum of residuals or on a fluctuation-type test statistic, that monitors the fluctuations in re-estimating the parameters of the regression with new data. Our procedure is based on a CUSUM-type detector as presented in the sections to come.

Our contribution to the literature therefore consists of a consistent monitoring procedure for structural breaks in fractional cointegration regressions. It is stated by Wagner and Wied (2017) that their proposed method for monitoring classical cointegration also is consistent against fractional alternatives. For this reason a comparison between the two methods for fractional alternatives is also presented in this paper to analyze if our method is better suited for the fractional cointegration framework it is derived for. We find that our detector has substantially better size properties for fractional cointegration than their method. Also, we present an application of our detector in the context of momentum investing and locate a structural break in a fractional cointegration regression during the momentum crash in April 2009.

The rest of the paper is organized as follows: Section 2 briefly discusses the basic theory and idea behind the concept of monitoring structural change. Based on that Section 3 presents our Detector for monitoring structural breaks in fractional cointegration, derives the limiting distribution and shows the behavior under fixed and local alternatives. The corresponding proofs are gathered in the appendix. In chapter 4 we present the results of the Monte-Carlo simulation and compare the size and power properties of our detector with the detector of Wagner and Wied (2017). Chapter 5 shows the results of the application of our detector in momentum investment. Finally, section 6 concludes the paper.

2 Monitoring Stationarity and Principles

This section provides the theoretical base for monitoring methods with focus on the monitoring procedure provided in the working paper by Wagner and Wied (2015) that can be used to monitor stationarity of a time series and thus also to monitor cointegration in its classical sense. The

principle behind this kind of monitoring dates back to Chu et al. (1996) and their proposals for monitoring structural breaks in linear regressions. For this consider the cointegration regression

$$Y_t = D_t^{'}\theta_D + X_t^{'}\theta_X + u_t, \quad t = 1, ..., T \tag{1}$$

where X_t is $I(1)$ and u_t is $I(0)$. Now assume we estimated this regression using e.g. the FM-OLS technique by Phillips and Hansen (1990) on the calibration period of length $[mT]$, with $0 < m < 1$ to get our set of parameters and we want to monitor for stationarity of u_t for the remaining sample $[mT + 1], ..., T$. Usually the null of stability, so in this case stationarity, and the alternative of a structural break are considered. This means under the alternative there is some point in the sample $[rT]$ from which the process (1) behaves like a random walk, with per assumption $0 < m < r < 1$.

To find the point $[rT]$ Wagner and Wied (2015) use an idealized scenario of a dataset of length T and a known long-run variance ω to show the intuition of using a detector for monitoring. For this define $S_i = \sum_{t=1}^{i} u_t$. Then the detector, which is based on the KPSS test for stationarity proposed by Kwiatkowski et al. (1992), for stationarity of u_t is given by

$$H^m(s) := \frac{1}{\omega^2}\left(\frac{1}{T}\sum_{i=[mT]+1}^{[sT]}(\frac{1}{\sqrt{T}}S_i)^2 - \frac{1}{T}\sum_{i=1}^{[mT]}(\frac{1}{\sqrt{T}}S_i)^2\right). \tag{2}$$

Here, the first sum displays the sum of the error terms over the monitoring period, while the second sum represents the cumulative sum of errors in the calibration period up until observation [mT]. So, the test statistic is based on a comparison of cumulative sums, which becomes larger under the alternative of a non-stationary process.

However, such a detector can not be applied directly on the set of time series. The goal is to use this detector in a way such that it finds potential structural breaks as fast as possible while also having minimal size distortions. To address this problem define the detection time $\tau_m(H^m, g, c)$, which depends on the chosen detector function H^m, a weighting function $g()$ and the chosen critical value c, which also depends on the function g. The detection time is given by the first point in time, for which the null is rejected if τ_m is finite and has the form

$$\tau_m := \min_{s:[mT]+1\leq[sT]\leq T}\left\{\left|\frac{H^m(s)}{g(s)}\right| > c\right\}.$$

This means that the null is rejected once the weighted detector exceeds some critical value for the first time in absolute value. So, in order to use the detector proposed by Wagner and Wied (2017) or the one proposed in this article one also needs to choose a weighting function to standardize the detector. Determining this function $g()$ is the result of a trade-off between the aforementioned minimal detection delay and minimal size distortions, as both properties can not be achieved simultaneously in an analytically tractable way.

If $\left|\frac{H^m(s)}{g(s)}\right| < c$ for all t the rejection time is interpreted as ∞ meaning that the null is not rejected at any point in the monitoring period. Similar to classical testing methods we want to set some

significance level to the monitoring procedures. This translates in this case to

$$\begin{aligned}\lim_{T\to\infty} P(\tau_m < \infty) &= \lim_{T\to\infty} P\left(\min_{s:[mT]+1\leq[sT]\leq T}\left\{\left|\frac{H^m(s)}{g(s)}\right| > c\right\} < \infty\right) \\ &= \lim_{T\to\infty} P\left(\sup_{s:[mT]+1\leq[sT]\leq T}\left\{\left|\frac{H^m(s)}{g(s)}\right| > c\right\}\right. \\ &= P\left(\sup_{m\leq s\leq 1}\left|\frac{H^m(s)}{g(s)}\right| > c\right) = \alpha,\end{aligned} \tag{3}$$

so that we reject the null with probability α, which is our chosen significance level. In our case of fractional cointegration the weighting function is

$$g(s) = s^{2d_x+1},$$

which is the generalized form of the weighting function used by Wagner and Wied (2017) in the case of no trend term allowing for more flexibility in the degree of integration. In the case of classical cointegration $d_x = 1$ and the weighting becomes s^3 as in Wagner and Wied (2017). A heuristic explanation for the weighting function $g(s)$ in general is given by Trapani and Whitehouse (2020). The detector function should grow in absence of breaks, however, it grows slower than $g(s)$ such that after a break occurs the residuals are substantially larger and $H(s)$ grows faster than $g(s)$.

3 Monitoring Fractional Cointegration

We consider monitoring breaks in a fractional cointegration regression allowing for possible deterministic trends which can also be subject to breaks. Our model is given as

$$Y_t = D_t'\theta_D + X_t'\theta_X + u_t, \tag{4}$$

$$X_t = \sum_{s=1}^{t} v_s \tag{5}$$

The following set of assumptions has to hold for equations (4) and (5):

Assumption 1. *For the process $X_t = \sum_{s=1}^{t} v_s$, v_s is $I(d_v)$ and thus X_t is $I(1+d_v)$.*

Assumption 2. *The trend component follows $D_t = (1, t, t^2, \ldots, t^{p-1})$ with normalization $G_D = diag(T^{1/2}, T^{3/2}, \ldots, T^{p-1/2})$.*

Assumption 3. *The residuals u_t are $I(d_u = d_x - b)$ with $d_u < d_x$. Together with assumption 1 this results in two possible scenarios: 1: $d_x + d_u > 1$ and 2: $d_x + d_u \leq 1$. In both cases it must hold that $0 < d_x + d_u$.*

Assumption 4. *It holds that*

$$T^{2d_u-1/2}\sum_{t=1}^{[sT]}\begin{pmatrix} u_t \\ v_t \end{pmatrix} \quad \Rightarrow \quad \Omega^{1/2}\mathbf{B_d}(s),$$

where $\mathbf{B_d}$ *is a vector fractional Brownian motion with memory parameter* d_u *and* d_x *and* $\Omega^{1/2}$ *is defined as in (8).*

The parameter of interest is $\theta = (\theta_D`, \theta_X`)$ in combination with the integration order d_u resulting in our hypotheses:

$$H_0 : \begin{cases} \theta_1 = \theta, t = 1, \dots, T \\ u_t \sim I(d_X - b), b \geq 0, t = 1, \dots, T \end{cases} \tag{6}$$

versus

$$H_1 : \begin{cases} \theta_1 \neq \theta \quad \text{for some} \quad m \leq r < 1 \quad \text{or} \\ u_t \sim I(d_X - b), b \geq 0, t = 1, \dots, [rT] \quad \text{and} \\ u_t \sim I(d_X - b), b = 0, t = [rT] + 1, \dots, T. \end{cases} \tag{7}$$

As an estimator $\hat{\theta}_D$ and $\hat{\theta}_X$ for the parameters θ_D and θ_X we use the fully modified OLS as in Phillips and Hansen (1990).

Denote by $\hat{\Omega}$ a long-run variance estimator such as the HAC by Andrews (1991) for the matrix

$$\Omega = \begin{pmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{uv} & \Omega_{vv} \end{pmatrix}$$

with u referring to the residuals from equation (4) and v to the residuals from equation (5). It should be kept in mind that a MAC-estimator such as the one proposed by Robinson (2005) or Abadir et al. (2009), which are designed to deal with fractionally integrated processes, might also be applicable in special cases like purely stationary cointegration to get estimates for the long-run variance. However, in general they do not allow for non-stationary series. This means they are not applicable for the general fractional cointegration case, which usually presents a combination of non-stationary time series to obtain a stationary set of residuals. Further define

$$\Omega^{1/2} = \begin{pmatrix} \omega_{u \cdot v} & \lambda_{uv} \\ 0 & \Omega_{vv}^{1/2} \end{pmatrix}. \tag{8}$$

where $\omega_{u \cdot v}^2 = \Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}$ and $\lambda_{uv} = \Omega_{uv}(\Omega_{vv}^{1/2})^{-1}$ analogously to Wagner and Wied (2017). Based on this we obtain the adjusted FM-OLS residuals

$$\begin{aligned} \hat{u}_{t,m}^+ =& y_t^+ - D_t`\theta_{D,m} - X_t`\hat{\theta}_{x,m} \\ & u_t - v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{uv}^{-1} - D_t'(\hat{\theta}_{D,m} - \theta_D) - X_t'(\hat{\theta}_{X,m} - \theta_X), \end{aligned} \tag{9}$$

where the subscript m denotes that the corresponding estimate is based on the calibration period and $y_t^+ = y_t - \Delta X_t`\hat{\Omega}_{vv,m}^{-1}\hat{\Omega}_{uv,m}$.

Before deriving the limiting distribution of the residuals we need to define the process

$$\Xi_{Qu} = \frac{1}{\Gamma(d_u)} \int_0^1 (1-s)^{d_u-1} \int_0^s Q dW ds,$$

based on the work by Davidson and Hashimzade (2009). Where W denotes a standard Brownian motion and Q is a fractional Brownian motion with parameter $1 - d_x$. For the case $d_x + d_u > 1$, which is the situation given in Davidson and Hashimzade (2009), we obtain the limiting result

$$T^{d_x - d_u} \sum_{t=1}^{T} x_t u_t \xrightarrow{d} \Xi_{Qu}.$$

For $d_x + d_u \leq 1$ we have

$$T^{2d_x - 1} \sum_{t=1}^{T} x_t u_t \xrightarrow{d} \Xi_{Qu}.$$

The convergence rate for the case $d_x + d_u \leq 1$ can be found in Leschinski (2017). Following the previous assumptions we obtain Lemma 1 concerning the limiting behavior of the adjusted residuals:

Lemma 1. *a) Let there be a fractional cointegration model as equations (4) and (5) with assumptions 1 to 3 in place. Then the adjusted FM-OLS residuals have the limiting distribution for $m \leq s \leq 1$ if $d_x + d_u > 1$:*

$$\begin{aligned} T^{d_u - d_x} \sum_{t=1}^{[sT]} \hat{u}^+_{t,m} &\Rightarrow \hat{\omega}_{u \cdot v} \left(B_{d_u}(s) - \int_0^s J(z)' dz \left(\int_0^m J(z) J(z)' dz \right)^{-1} \int_0^m J(z) d\Xi_{uv}(z) \right) \\ &=: \hat{\omega}_{u \cdot v} W_{d_u}(s), \end{aligned} \tag{10}$$

where B_d denotes fractional Brownian motion with parameter d and $J(s) = (D(s)', B_{d_X}(s)')'$. Alternatively, if $d_x + d_u \leq 1$ and $0.5 < d_x$ the rate of convergence changes to:

$$\begin{aligned} T^{2d_x - 1} \sum_{t=1}^{[sT]} \hat{u}^+_{t,m} &\Rightarrow \hat{\omega}_{u \cdot v} \left(B_{d_u}(s) - \int_0^s J(z)' dz \left(\int_0^m J(z) J(z)' dz \right)^{-1} \int_0^m J(z) d\Xi_{uv}(z) \right) \\ &=: \hat{\omega}_{u \cdot v} W_{d_u}(s). \end{aligned} \tag{11}$$

Yet the limiting distribution remains equal.

b) In the stationary case $0 < d_x < 1/2$ and $0 \leq d_u < d_x$ we obtain for the adjusted FM-OLS residuals the limiting distribution for $m \leq s \leq 1$:

$$\begin{aligned} T^{1 - \frac{d_x + d_u}{2}} \sum_{t=1}^{[sT]} \hat{u}^+_{t,m} &\Rightarrow \hat{\omega}_{u \cdot v} \left(B_{d_u}(s) - \int_0^s J(z)' dz \left(\int_0^m J(z) J(z)' dz \right)^{-1} \int_0^m J(z) dZ(z) \right) \\ &=: \hat{\omega}_{u \cdot v} \mathcal{W}_{d_u}(s), \end{aligned} \tag{12}$$

where $Z(t) = \frac{\sigma_x \sigma_u}{\sqrt{a_1 a_2}} \int_{R^2} \int_0^t \left[\prod_{i=1}^2 (s - x_i)^{-\frac{d_i + 1}{2}} 1_{(x_i < s)} \right] ds dB_1(x_1) dB_2(x_2)$ with B_1, B_2 Brownian motions, $a_1 = A(d_1, d_1), a_2 = A(d_2, d_2)$, $A(d_1, d_2) = \int_0^\infty x^{-\frac{d_1+1}{2}} (x+1)^{-\frac{d_2+1}{2}}$ and σ_1 and σ_2 are real numbers.

Again B_d denotes fractional Brownian motion with parameter d and Ξ_{uv} and $J(s)$ are defined as before.

Now define as a crucial part of our test statistic the cumulative sum $\hat{S}_i := \sum_{t=1}^{i} \hat{u}_{t,m}^{+}$. The test statistic is given as

$$\hat{H}_m(s) = \frac{1}{\hat{\omega}^2_{u\cdot v,m}} \left(T^{2d_u-1} \sum_{i=[mT]+1}^{[sT]} (T^{d_u-1/2}\hat{S}_i)^2 \right). \tag{13}$$

This test statistic can then be used in equation (3) in order to get an α-level monitoring procedure for breaks in fractional cointegration. It can be seen that the test statistic is based on the test statistic proposed by Shin (1994) and Kwiatkowski et al. (1992) for the null of cointegration (stationarity) against an alternative of no cointegration (generalized to allow for fractional cointegration). Other forms for the test statistic could also be possibly based on the different forms presented in the context of polynomial cointegration regressions by Knorre et al. (2021). These include a self-normalized version of the test statistic, a moving-window approach and a version based on the difference between the partial sums in the calibration and monitoring period. These statistics might also be generalized into the fractional cointegration framework, however, we will only focus on the one given in equation (13).

For our test statistic we get the limiting distribution

$$\hat{H}_m(s) \Rightarrow \int_m^s W^2_{d_u}(z)dz \tag{14}$$

in the cases described in part a) of Lemma 1. For stationary fractional cointegration as described in part b) of Lemma 1 we obtain the limiting distribution

$$\hat{H}_m(s) \Rightarrow \int_m^s \mathcal{W}^2_{d_u}(z)dz, \tag{15}$$

with $\mathcal{W}^2_{d_u}$ as in (12).

In a next step we establish consistency for our detector under fixed and local alternatives.

Proposition 1. *Let the data be generated by (4) and (5) and consider the fixed alternative. For this u_t is assumed to be $I(d_u = d_x - b)$ up to $[rT]$ and from that on b becomes zero, resulting in no fractional cointegration between X and Y as u_t is then $I(d_x)$. Alternatively, let $\theta_1 \neq \theta$ from $[rT]$ on. Then for any $0 < c < \infty$ and continuous $g()$ for which $0 < g() < \infty$ it holds that:*

$$\lim_{T\to\infty} P(\tau_m(\hat{H^m}, g, c(\alpha, g) < \infty) = 1,$$

meaning that the monitoring procedure is consistent.

Proposition 2. *Let the data be generated by (4) and (5) and consider the local alternative. Here u_t is assumed to be $I(d_u = d_x - b)$ up to $[rT]$ again and from that on $u_t = u_t^0 + T^{d_x-1/2}\delta \sum_{i=[rT]+1}^{t} \xi_i$, where $u_t^0 \sim I(d_u)$ and independent of this $\xi_t \sim I(d_X)$ and $\delta > 0$. Then it holds that*

$$\lim_{T\to\infty} P(\tau_m(\hat{H^m}, g, c(\alpha, g) < \infty) \geq 1 - \epsilon,$$

with $0 < \epsilon < 1 - \alpha$ meaning that the detector has non-trivial local power.

The proofs of the previous results are gathered in appendix A.

4 Finite Sample Performance

In this section we investigate the performance of our proposed monitoring test statistic in finite samples. We estimate the critical values by simulation for the 5% significance level using the previously defined weighting function. We analyze the properties of the monitoring procedure for a fixed sample size of 1000 observations and different starting points of the monitoring period $m = (0.25, 0.5, 0.75)$. The monitoring method by Wagner and Wied (2017) is also consistent against breaks in the degree of integration, if the break means that the residual u_t was stationary before and then is non-stationary fractionally integrated after the break. As this is also the main field of application for our proposal a comparison of detection times and rates seems appropriate. The critical values for the application of our detector are derived by bootstrap. They depend on a combination of three parameters m, d_x and d_u. A brief introduction on how the critical values behave with respect to changes in the three parameters can be seen in Figure 1:

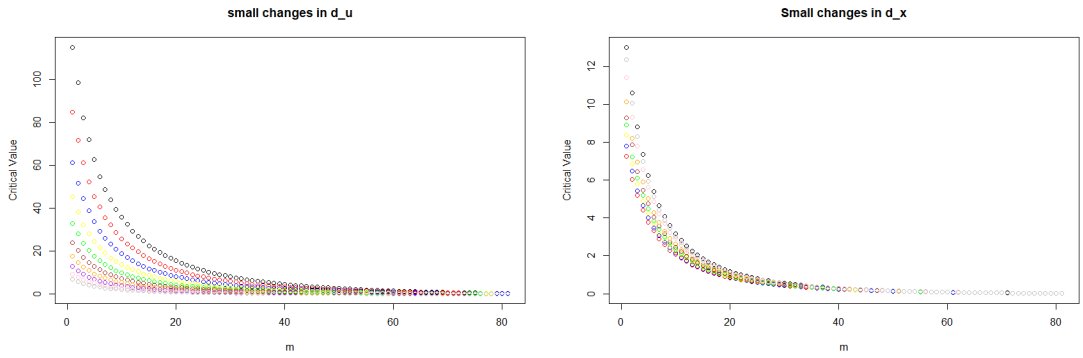

Figure 1: Changes in Critical Values

The different colors in the plots refer to different values of d_u in the left panel and different values of d_x on the right panel. On the left the critical values for $d_x = 0.8$ and $d_u = 0.1, 0.2$ in increments of 0.01 are displayed and a large dispersion depending on the value of d_u can be seen especially for smaller values of m. On the right the critical values for the combinations of $d_u = 0.1$ and $d_x = 0.8, 0.9$ in 0.01 increments can be seen. Here, the changes are not that large, however, for smaller m they could influence the test decision.

Now we analyze how the detector for fractional cointegration breaks behaves in finite samples by using the following model:

$$y_t = \mu + \gamma t + x_t \beta + u_t \tag{16}$$

$$x_t = \sum_{s=1}^{t} v_s \tag{17}$$

where u_t is $I(d_u)$ with $d_u < 0.5$ and x_t is $I(d_x)$ with $0.5 < d_x < 1$ and v_t is fractionally integrated white noise. Here, we focus on the simple case, where $\mu = 3$ and $\gamma = 0$ without a loss of generality. For estimation we use FM-OLS to get parameters and the Bartlett kernel with

automatic bandwidth selection by Andrews (1991) for long-run variance estimation. Note that size and power properties change slightly, if for example the quadratic spectral and/or bandwidth selection by Newey and West (1994) are used but they remain at the same order.

To analyze the detection times we set a break at the 850th observation, where u_t becomes $\mathrm{I}(d_x)$ for the remainder of the sample. This means that the break is in comparison to the simulation study by Wagner and Wied (2017) not in the beginning of the monitoring period but in the middle of it. Due to this there is the possibility for the detector to detect the break too early.

Table 1: Size for m=0.75 with adjustment, $\alpha = 0.05$

d	*Our Detector*						*Wagner and Wied (2017)*					
$d_u =$	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
$d_x = 1$	0.054	0.043	0.040	0.035	0.035	0.038	0.048	0.197	0.383	0.503	0.585	0.609
$d_x = 0.9$	*0.054*	0.048	0.040	0.037	0.039	0.039	0.054	0.216	0.392	0.529	0.614	0.638
$d_x = 0.8$	*0.055*	*0.050*	0.047	0.037	0.038	0.036	0.061	0.225	0.414	0.551	0.636	0.666
$d_x = 0.7$	*0.058*	*0.046*	*0.039*	0.038	0.036	0.035	0.070	0.240	0.439	0.584	0.656	0.689
$d_x = 0.6$	*0.056*	*0.049*	*0.046*	*0.037*	0.035	0.032	0.078	0.250	0.458	0.604	0.689	0.716
$d_x = 0.4$	0.062	0.052	0.044	0.036			0.095	0.278	0.495	0.655		

Table 2: Size for $m = 0.5$ with adjustment, $\alpha = 0.05$

d	*Our Detector*						*Wagner and Wied (2017)*					
$d_u =$	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
$d_x = 1$	0.056	0.049	0.045	0.043	0.039	0.042	0.053	0.132	0.239	0.348	0.415	0.454
$d_x = 0.9$	*0.053*	0.051	0.046	0.043	0.042	0.040	0.035	0.106	0.211	0.311	0.395	0.442
$d_x = 0.8$	*0.058*	*0.052*	0.046	0.045	0.040	0.041	0.023	0.080	0.187	0.295	0.375	0.413
$d_x = 0.7$	*0.063*	*0.060*	*0.048*	0.044	0.043	0.039	0.015	0.059	0.159	0.274	0.359	0.399
$d_x = 0.6$	*0.065*	*0.060*	*0.054*	*0.045*	0.039	0.035	0.007	0.040	0.141	0.267	0.349	0.390
$d_x = 0.4$	0.067	0.064	0.054	0.044			0.002	0.029	0.012	0.253		

The size of the fractional cointegration monitoring approach and the detector proposed by Wagner and Wied (2017) can be seen in Tables 1 to 3 for different settings of m. For the cases in which $d_x + d_u \leq 1$ resulting in a different convergence rate the results are displayed in italics. However, the detector behaves similarly in these cases and the size is only increased non-substantially. The degrees of memory d_u and d_x are estimated by the exact local whittle by Shimotsu and Phillips (2005) with a bandwidth of T^δ. As the estimate is reliant on the bandwidth to a certain degree there might be changes in the size and power if one chooses a different bandwidth for parameter estimation. In general, we found that bandwidths with $\delta \in [0.6 : 0.8]$ ensure good size properties, while lower bandwidths may lead to substantial size distortions. The size values seen in the tables are calculated for the value $\delta = 0.7$. The critical values are derived by bootstrap for the given parameter setting of d and m with 10000 repetitions. To cover the uncertainty in

Table 3: Size for $m = 0.25$ with adjustment, $\alpha = 0.05$

d	*Our Detector*						*Wagner and Wied (2017)*					
$d_u =$	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
$d_x = 1$	0.061	0.059	0.054	0.052	0.054	0.057	0.055	0.099	0.159	0.219	0.286	0.347
$d_x = 0.9$	*0.063*	0.063	0.059	0.056	0.056	0.062	0.030	0.064	0.107	0.166	0.227	0.299
$d_x = 0.8$	*0.073*	*0.071*	0.065	0.062	0.061	0.058	0.014	0.031	0.068	0.120	0.182	0.249
$d_x = 0.7$	*0.076*	*0.077*	*0.070*	0.070	0.059	0.057	0.005	0.014	0.036	0.080	0.143	0.208
$d_x = 0.6$	*0.081*	*0.080*	*0.076*	*0.065*	0.061	0.052	0.001	0.004	0.019	0.058	0.115	0.185
$d_x = 0.4$	0.097	0.086	0.079	0.065			0.000	0.000	0.006	0.038		

the estimates of d_u and d_x for the different calibration period lengths we use the corresponding estimate of d plus one standard deviation.

It can be seen that the size behaves decently for all kinds of values for fractional cointegration and for the different lengths of the monitoring period. For smaller m the size generally increases for our detector which is to be expected as a smaller calibration period means a longer monitoring period, a shorter calibration period with more uncertainty in the estimates and thus also a higher chance to reject the null hypothesis. The highest size can be found for $m = 0.25$, $d_x = 0.6$ and $d_u = 0$ at 0.081. For greater values of m the size only varies a bit around the targeted 5% with the lowest values being around 3.5% for fractional cointegration regressions with a rather low amount of integration reduction like the combinations of $d_x = 0.7$ or $d_x = 0.6$ and $d_u = 0.4$. On the other side the size of the detector proposed by Wagner and Wied (2017) is severely affected by fractional degrees of integration. For $m = 0.75$ their detector‘s size increases by a large margin up to over 60% for any combination with $d_u \geq 0.4$. For $m = 0.5$ the size distortions follows the general trend of increasing for "more fractional" cointegration regressions, however, the distortions here are not that severe. The maximum size is reached with around 45%. For $m = 0.25$ their detector shows decreasing size in d_x such that for $d_x = 0.6$ and $d_u = 0$ the size becomes almost zero and increasing size in d_u resulting in oversized behavior in many situations again. All in all our detector ensures a more adequate size for fractional cointegration in comparison to the one proposed by Wagner and Wied (2017), which suffers from substantial size distortions in both directions.

Tables 4 to 6 show the power properties of both detectors for different settings. As noted before the break happens during the monitoring period allowing for the possibility of a detected break before the break actually occurred. In the tables displayed here we have included the too early breaks into the power, as else the power is lower than the before calculated size in some situations. As shown in Wagner and Wied (2015) the monitoring procedure by Wagner and Wied (2017) is also consistent against breaks in integration of u from stationary into near-integrated and fractional integration of. This can also be seen in most of our results as the power stays at an adequate level for stationary settings of d_u with $m \geq 0.5$. However, these power results are inflated due to the high size distortions. To adjust for these size distortions we set the significance level of the benchmark detector to 1%, which however still leads to severe

Table 4: Power for a break in Integration to I(d_x) $m = 0.75$ all breaks, $\alpha = 0.05$, $\alpha = 0.01$ for Wagner and Wied (2017)

d	*Our Detector*						*Wagner and Wied (2017)*					
$d_u =$	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
$d_x = 1$	0.951	0.908	0.821	0.704	0.550	0.379	0.963	0.949	0.931	0.892	0.819	0.716
$d_x = 0.9$	*0.914*	0.846	0.736	0.584	0.415	0.248	0.913	0.895	0.862	0.810	0.733	0.614
$d_x = 0.8$	*0.865*	*0.767*	0.616	0.447	0.275	0.139	0.823	0.808	0.765	0.710	0.636	0.521
$d_x = 0.7$	*0.784*	*0.656*	*0.491*	0.313	0.158	0.073	0.695	0.662	0.632	0.595	0.537	0.476
$d_x = 0.6$	*0.688*	*0.529*	*0.346*	*0.191*	0.085	0.041	0.524	0.500	0.499	0.483	0.485	0.464
$d_x = 0.4$	0.414	0.253	0.126	0.059			0.136	0.178	0.299	0.419		

Table 5: Power for a break in Integration to I(d_x) $m = 0.5$ all breaks, $\alpha = 0.05$, $\alpha = 0.1$ for Wagner and Wied (2017)

d	*Our Detector*						*Wagner and Wied (2017)*					
$d_u =$	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
$d_x = 1$	0.764	0.644	0.519	0.378	0.245	0.138	0.769	0.750	0.697	0.611	0.511	0.387
$d_x = 0.9$	*0.688*	0.567	0.428	0.281	0.168	0.096	0.626	0.597	0.534	0.438	0.341	0.273
$d_x = 0.8$	*0.601*	*0.477*	0.331	0.205	0.109	0.070	0.434	0.414	0.340	0.263	0.206	0.186
$d_x = 0.7$	*0.514*	*0.367*	*0.249*	0.136	0.070	0.048	0.242	0.207	0.167	0.137	0.126	0.140
$d_x = 0.6$	*0.416*	*0.286*	*0.173*	*0.093*	0.056	0.003	0.077	0.061	0.056	0.066	0.095	0.113
$d_x = 0.4$	0.227	0.138	0.085	0.059			0.001	0.001	0.008	0.029		

Table 6: Power for a break in Integration to I(d_x) $m = 0.25$ all breaks, $\alpha = 0.05$, $\alpha = 0.01$ for Wagner and Wied (2017)

d	*Our Detector*						*Wagner and Wied (2017)*					
$d_u =$	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
$d_x = 1$	0.374	0.298	0.222	0.152	0.107	0.085	0.214	0.199	0.185	0.157	0.143	0.152
$d_x = 0.9$	*0.335*	0.259	0.187	0.326	0.092	0.077	0.065	0.063	0.063	0.068	0.081	0.098
$d_x = 0.8$	*0.298*	*0.213*	0.159	0.119	0.086	0.066	0.007	0.012	0.016	0.027	0.039	0.061
$d_x = 0.7$	*0.250*	*0.189*	*0.137*	0.097	0.075	0.060	0.001	0.001	0.003	0.008	0.020	0.039
$d_x = 0.6$	*0.210*	*0.155*	*0.112*	*0.078*	0.064	0.051	0.000	0.000	0.001	0.002	0.017	0.027
$d_x = 0.4$	0.152	0.112	0.089	0.066			0.000	0.000	0.000	0.000		

size distortions. In general it can be seen that for both detectors the power decreases for "more fractional" integration orders, as this also decreases the size of the structural break. Also, the choice of m drastically affects the power properties of both detectors. The general monitoring

principle that m should be as large as possible while the monitoring period should be as short as possible seems to play a crucial role for the monitoring of fractional cointegration. This might be due to the need to estimate the degrees of integration of both x and u, where the exact local whittle estimator is used. This estimator needs adequately large time series in order to get proper estimates. For smaller calibration periods the series might not be long enough resulting in more variation and thus reducing the power.

Our detector has a comparable or higher power for most combinations with $d_u \leq 0.1$. Increases in d_u and thus smaller breaks decrease the power as expected. Noticeably the decreases are rather small for $m = 0.75$ with the detector by Wagner and Wied (2017). This might be, however, a relic of the still increased size in these settings. A majority of the breaks found there, are located before the break actually happened, meaning that the detector rejected the null too early due to the size distortions. This can be seen exemplarily in Figure 2, where the estimated break dates of the power analysis are displayed for the setting $d_x = 0.9$, $m = 0.75$ and varying d_u. If no break was found the graph displays the break date as 1000 and the share of early detections is displayed on the left of the corresponding graph for each parameter setting. For both estimators it is visible that the estimated break dates increase with increasing d_u meaning that both face more problems finding the break in time. Yet, it seems clear that due to the higher size a larger fraction of the break points found by Wagner and Wied (2017)'s detector is located before the break in observation 850, indicated by the black vertical line, for higher d_u. This can also be seen in the share of early detections, as with an increase in d_u the portion rises from ca. 1% to over 30% for their detector. For an increasing d_u the share of early detections also increases, if one uses our detector, however it stays at a rather small level. These findings also generalize to other settings of d_x and m. For lower values of d_u our detector has an equivalent or faster detection time but also includes some more early detections. For decreasing m our detector has substantially higher power properties again for lower levels of d_u and d_x. Yet, the power in most of these situations is still relatively low especially for $m = 0.25$.

A conclusion that can be drawn from this simulation study is that the fractional cointegration monitoring approach should be used in settings, where the calibration period is rather long and consists of enough data to ensure proper estimation of the different integration levels. The power properties of our detector deteriorate for smaller m, while the size stays at a reasonable level. This fixed size is the main improvement of our proposed method in comparison to the cointegration monitoring by Wagner and Wied (2017) used in this context.

5 Application

To show the applicability of our method we apply it in the context of momentum investing. As shown by Jegadeesh and Titman (1993) the momentum investment strategy, that follows the simple rule to buy previous "winners" and short previous "losers" in the financial market, yields unreasonably high returns. A downside of this strategy, however, is that as explained by Daniel and Moskowitz (2013) the returns have a high negative skewness resulting in a relatively large risk exposure. This means they experience high negative returns in some periods, eating up previous profits especially in times of economic distress like the global financial crisis and shortly thereafter. They frame these periods "momentum crashes". Barroso and Santa-Clara (2015) argue that this

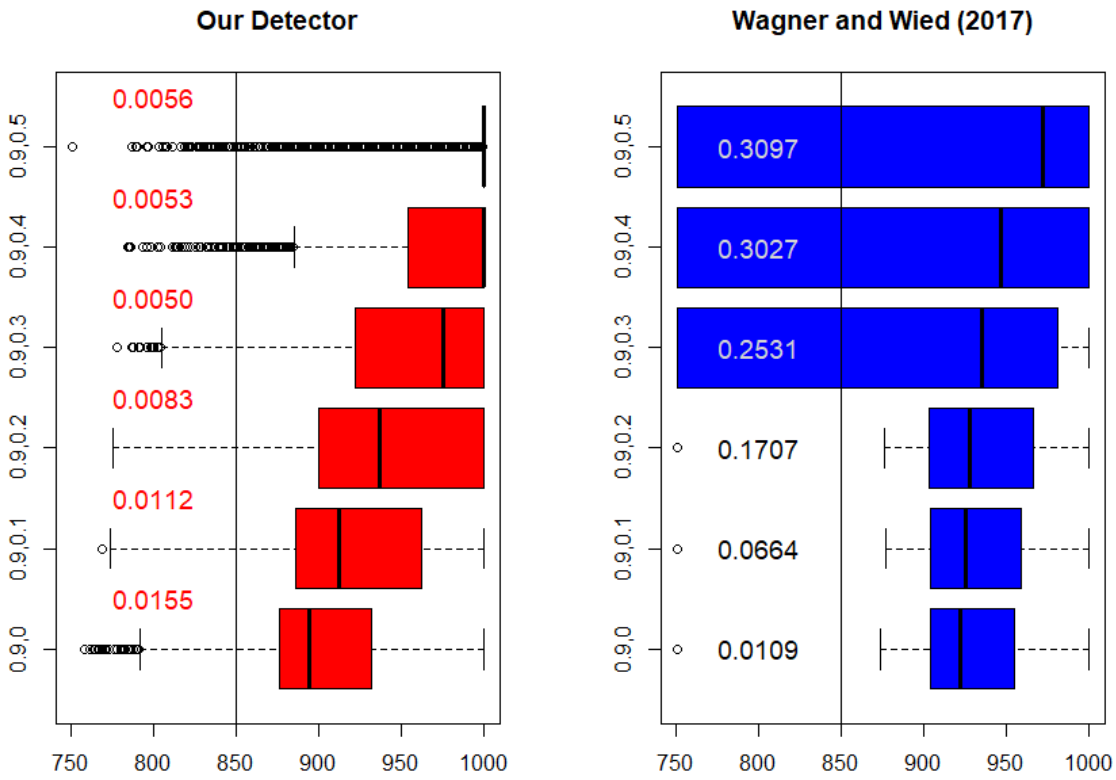

Figure 2: Detection Times for $d_x = 0.9$ and $m = 0.75$

skewness makes the strategy in itself unappealing to investors as these periods of high negative returns are too risky. Yet, it is proposed by Daniel and Moskowitz (2013) and Barroso and Santa-Clara (2015) that this risk is manageable as these crashes are somewhat predictable. Using this proposal Barroso and Santa-Clara (2015) present a risk-adjusted strategy based on volatility forecasts that are used to predict the crashes and adjust the amount of money invested in times of predicted crashes increasing the Sharpe ratio compared to pure momentum investment. The great variability of the standard strategy can be seen in figure 3. These momentum crashes are especially prominent in the time period between the years 2000 and 2010 covering the global financial crisis and the crash following the dotcom bubble resulting in losses of more than 40% in one month respectively. Predicting and or locating these crashes and thus preventing substantial losses is therefore of great interest for risk management of financial actors.

In this context we propose a fractional cointegration relationship that is used to approximate underlying relationships in the market that are driving the momentum returns in the long run, an approach that has not been discussed in the literature yet. This could fuel the literature on theoretical model building in the context of the momentum strategy. If there is a break in this fractional cointegration model, we can conclude that some of these long term market forces might be out of order inducing a momentum crash. Our approach is to create a model that relates the momentum strategy to the general state of the market. To do so we use squared monthly returns of the S&P500 to proxy the market and relate this to the squared monthly returns of the momentum strategy. The market data is obtained from Yahoo! Finance and the monthly data

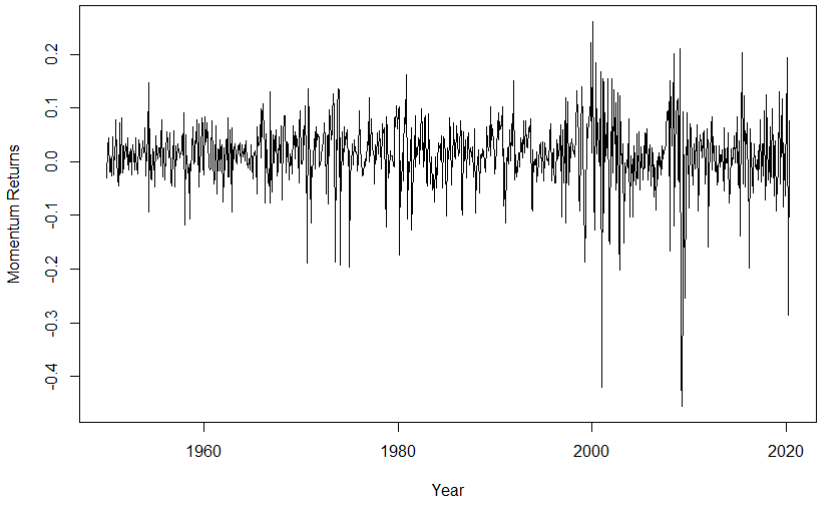

Figure 3: Momentum returns

for the momentum strategy is available in the online library provided by Fama and French. We cover the years from January 1950 to May 2020.

To test if there exists a fractional cointegration relationship between different variables several tests haven been proposed. We apply the tests by Wang et al. (2015) (WWC15), that does not assume a specfic range of d and Robinson (2008) (RO8), which is only applicable for purely stationary systems. For these tests a proper estimation of the memory parameter of the respective series is crucial. To ensure this we use the exact local whittle estimator by Shimotsu and Phillips (2005). This estimator depends on a user-chosen bandwidth parameter, which we specified to be $T^{0.75}$. In general, for a fractional cointegration model to be valid the memory parameters of the respective series must be statistically equal. Our estimates are $d_{sp} = 0.2149$ for squared market returns and $d_{mom} = 0.2169$ for the squared momentum returns. As these two estimates are already quite close to each other, this might be the first indicator of an existing cointegration relation. This presumption is also confirmed by the test for equal memory parameters proposed by Robinson and Yajima (2002), as the test is unable to reject the null of a common d in the two series. The calculated p-value using a bandwidth of $T^{0.75}$ as before is roughly 49%.

Table 7: Test results

	Calibration Period	*After estimated Break*
WWC15	Reject H_0	No Rejection
RO8	Reject H_0	Reject H_0

$H_0 =$ No Fractional Cointegration

Both tests reject the null of no fractional cointegration during the calibration period. After our estimated breakpoint only the test by Robinson (2008) is still able to reject the null, while the test by Wang et al. (2015) does not reject the null anymore. As these results implicate that

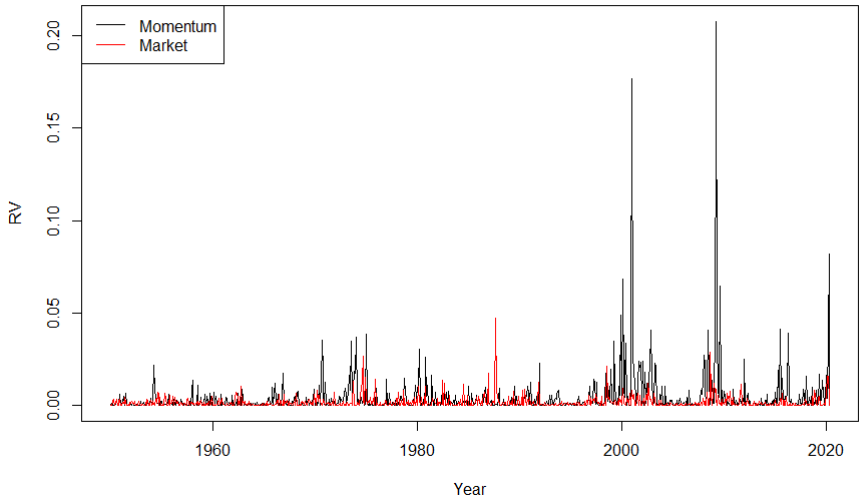

Figure 4: Squared market and momentum returns

a fractional cointegration model is suitable for the calibration period we establish the simple regression model:

$$r_{mom}^2 = c + \theta r_{sp}^2,$$

that presents the squared momentum returns as a function of the squared market returns and a constant without a trend term. For estimation we use the FM-OLS estimator by Phillips and Hansen (1990) using the Bartlett kernel and automatic bandwidth selection by Andrews (1991) to estimate the long-run variance, which is a crucial part of the detector (13). Using different kernels like the quadratic spectral kernel could influence the detection time of the break point as discussed later. The squared returns can be seen in figure 4. There does not seem to be too much movement from 1950 to ca. 2000 and both series seem to vary roughly by the same amount, with some delays and exceptions. However, it is apparent that in the years thereafter the momentum strategy experienced significantly greater variations and moves almost independent from the market returns in the years following the dotcom bubble and the global financial crisis. This fact also motivates our idea to monitor this fractional cointegration relationship. If a monitoring procedure is able to detect these times in which the momentum returns have much larger variations than the market baseline, which also are the periods of momentum crashes as in Daniel and Moskowitz (2013), it might be helpful in the context of risk-management of momentum based investments. The parameter estimates for the whole sample, calibration period and after the estimated break point can be seen in Table 8.

To apply our monitoring procedure we set $m = 0.65$, which results in a calibration period from 1950 up until September of 1995. Based on this period of time we estimate the model, which is then used to monitor the remaining data points for structural stability of the model. Here, it should be noted that a different selection for m also affects the estimated break date. We find in general, that our detector estimates a break in April 2009 for most $m \in [0.6; 0.7]$. In some cases the break is detected briefly after or before April 2009, but there unfortunately seems to be no systematic reason behind these deviations. April 2009 also corresponds to the month with the

Table 8: Regression Results

	Whole Sample	*Calibration Period*	*After estimated Break*
RVsp500	0.7315***	0.0386	1.4890***
	(0.2173)	(0.0805)	(0.3581)
Constant	0.0032***	0.0026***	0.0028*
	(0.0008)	(0.0003)	(0.0011)

Note: *p<0.1; **p<0.05; ***p<0.01

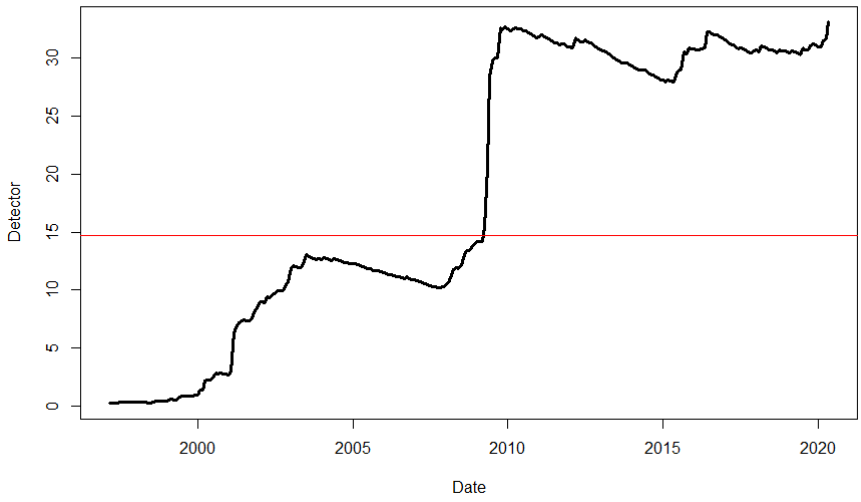

Figure 5: Detector

largest loss of the momentum strategy in the sample with a return of roughly −45% in one month. A predicted break briefly before this large loss would greatly benefit risk-managed momentum strategies similar to crash indicators proposed e.g. by Dierkes and Krupski (2022). However, as for most parameter settings the break is found in the month of this loss, this indicated market discrepancy could not be used for proper risk-management.

Figure 5 shows the detector over the monitoring period: It is clearly visible how the momentum crash in 2009 affected the detector with a sharp increase. In red one can see the critical value, which is obtained by bootstrap and passed by a large margin in April 2009 indicating the estimated break exactly in that month. It is also noticeable how the detector increases sharply during the time of the dotcom bubble and thereafter, yet this is not enough of an increase to indicate a break. Taking a look at the regression results for the period before and after April 2009 in Table 8 it is also visible that the coefficient for the squared market returns changes substantially after the breakpoint and is multiple standard errors larger than before, which might be an indicator for a break in parameter. Applying the exact local whittle estimator again with a bandwidth of $T^{0.75}$ on the respective model's residuals one also finds differences in the degree

of integration, however, all residual series stay in the stationary interval of $d_u < 0.5$. For the calibration period the residuals are integrated of order 0.2213, while the residuals after the estimated breakpoint are integrated of order 0.3620 meaning that after the break the residuals are a bit closer to nonstationary behavior. However, the smaller sample size post break also increases the uncertainty in the estimate of d. When applying the tests for fractional cointegration on the sample after the estimated break the test by Robinson (2008) still rejects the null of no cointegration, while the test by Wang et al. (2015) does not reject it. As this does not allow us to draw a concise conclusion on whether a cointegration model is still suitable in general, the estimated break is likely a break in the regression parameter. This is also supported by the expanding window estimator of the regression parameter. Figure 6 displays the series of estimates for an increasing number of observations in steps of one plus intervals of one standard deviation around them. It can be seen that for the calibration period the estimates remain stable and then start to change during the early 2000s. However, the increase in the estimate that is induced by the dotcom-bubble is not large enough to imply a break, as the detector does only find the break associated with the increase during 2009.

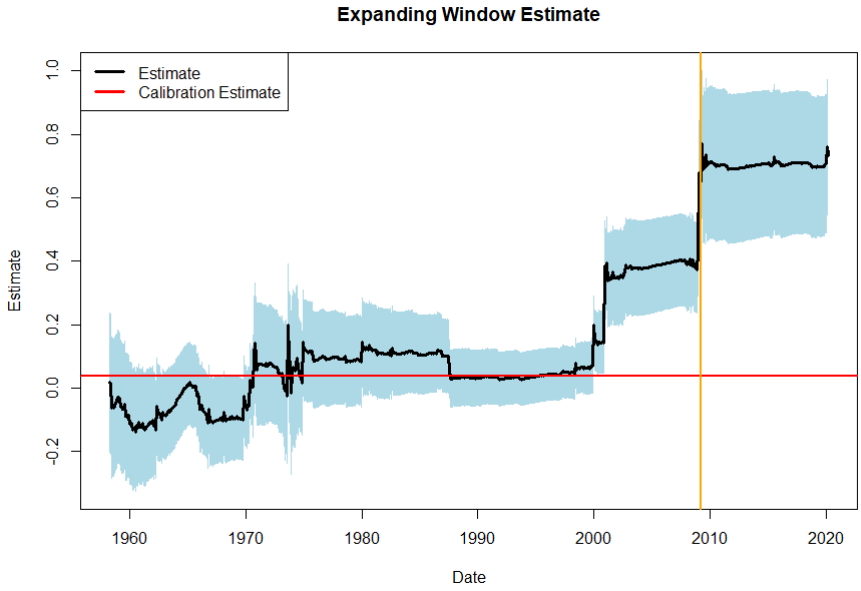

Figure 6: Expanding Window Estimates of $\theta_{RV_{sp}}$

If one uses the quadratic spectral kernel instead of the Bartlett kernel to get an estimate for the long-run covariance matrix while keeping $m = 0.65$ and automatic bandwidth selection along Andrews (1991) the detector is able to detect a break as early as November 2008 briefly after the collapse of the Lehmann Brothers bank. Using this estimated breakpoint as an indication for a disequilibrium between market and momentum returns and thus redirecting investments away from momentum could have helped preventing substantial losses during the crash of April 2009. Yet, this breakpoint is quite early before the actual crash and it might be difficult to draw a concise implication on the investment strategy and on how long the adjustments away from the momentum strategy should be implemented.

As a general note on the beforehand application it should be kept in mind that $m = 0.65$ not only means that we start monitoring in September 1995 but also that at that point it is the goal

to monitor the cointegration model until May 2020, the last observation in the data. This is due to the fact that the procedure is a closed end monitoring method, so that one always has to specify an end to the monitoring period. As discussed in footnote 7 by Wagner and Wied (2017) it is to some extend up to a semantic discussion if these procedures are to be considered online monitoring or structural break testing. They can be used for monitoring as they only depend on estimates up to a certain time point, however if they are applied in retrospect like in our application and most of the literature they are closer to classical structural break testing for a given sample as the influence of different lengths of the monitoring period on the empirical results is usually not discussed. Especially in financial markets a shorter monitoring period may be sensible in applications that focus more on the monitoring aspect.

6 Conclusion

We propose a new detector that extends the monitoring of classical cointegration regressions to also allow for fractional cointegration regressions in its null and alternative. For this we use the exact local Whittle by Shimotsu and Phillips (2005) to estimate the integration order, we estimate the regression coefficients by FM-OLS by Phillips and Hansen (1990) and use a HAC- estimator for the long-run variance. The asymptotic distribution is derived as well as the behavior under the alternative.

In our Monte- Carlo study we examine the size and power properties of our detector and use the detector by Wagner and Wied (2017) as a benchmark. While their detector suffers from substantial size distortions for most cases of fractional cointegration our detector keeps the size fixed for most cases and also for purely stationary or non-stationary cointegration. The power of our detector is comparable to their test‘s power even though it is rather low for cases of fractional cointegration in which the reduction of integration is small.

We apply the test to monitor the fractional cointegration regression between the realized volatility of the momentum strategy and the realized volatility of the market. In this context we contribute to the literature on finding and predicting momentum crashes, as our detector is able to find the momentum crash subsequent to the global financial crisis. The idea of a fractional cointegration model in this regard fuels the literature on theoretical model building for analyzing momentum crashes and sheds light on possible market forces, which have not been discussed in the literature yet. However, one should keep in mind that the monitoring methods in general are designed in a way to find breaks with some delay. Due to this it might be difficult to implement our monitoring method to forecast a momentum crash before it actually occurs.

Acknowledgement: The authors are grateful to the participants of the Statistische Woche 2025, in particular Matei Demetrescu, in Regensburg for helpful comments and discussions. Philipp Sibbertsen gratefully acknowledges financial support by Deutsche Forschungsgemeinschaft under grant 258395632.

A Proofs

Proof of Lemma 1:

The limiting distribution stems from the established asymptotic behavior of the FM-OLS estimator by Phillips and Hansen (1990) in Theorem 3.2 in combination with the convergence rate and the representation of fractionally integrated processes discussed in Davidson and Hashimzade (2009) and Leschinski (2017) for case a) and Fox and Taqqu (1987) for case b), consistent long-run variance estimation and the continuous mapping theorem.

Proof of Proposition 1:

The partial sum process of the residuals can be written analogous to Wagner and Wied (2017) for $m \leq r < s \leq 1$ as

$$\begin{aligned}
T^{d_u-1/2}\sum_{t=1}^{[sT]}\hat{u}_{t,m}^{+} =& T^{d_u-1/2}\sum_{t=1}^{[rT]}\hat{u}_{t,m}^{+} + T^{d_u-1/2}\sum_{t=[rT]+1}^{[sT]}\hat{u}_{t,m}^{+} \\
=& T^{d_u-1/2}\sum_{t=1}^{[rT]}\hat{u}_{t,m}^{+} + T^{d_u-1/2}\sum_{t=[rT]+1}^{[sT]}u_t - T^{d_u-1/2}\sum_{[rT]+1}^{[sT]}v_t'\hat{\Omega}_{vv}^{-1}\hat{\Omega}_{vu}^{-1} - \\
&- T^{d_u-1/2}\sum_{t=[\tau T]+1}^{[sT]}D_t'(\hat{\theta}_{D,m}-\theta_D) - T^{d_u-1/2}\sum_{t=[\tau T]+1}^{[sT]}X_t'(\hat{\theta}_{X,m}-\theta_X).
\end{aligned}$$

The first term converges to $\hat{\omega}B_{d_u}(s)$ as before. The second term diverges if $b = 0$ and thus $u_t \sim I(d_X)$ with $d_X > d_u$ as then the scaling factor $T^{d_u-1/2}$ is not appropriate anymore. The remaining three terms converge in distribution. If there is a break in θ_X but not in d_u the last term diverges and the others converge in distribution.

For a break in parameter consider the representation of the scaled partial sum of the adjusted residuals:

$$\begin{aligned}
T^{d_u-1/2}\sum_{t=1}^{[sT]}\hat{u}_{t,m}^{+} =& T^{d_u-1/2}\sum_{t=1}^{[sT]}\hat{u}_t - T^{d_u-1/2}\sum_{t=1}^{[sT]}v_t\text{`}\Omega_{vv,m}^{-1}\Omega_{vu} - T^{d_u-1/2}\sum_{t=1}^{[sT]}D_t\text{`}(\hat{\theta_{D,m}}-\theta_D) \\
&- T^{d_u-1/2}\sum_{t=1}^{[sT]}X_t\text{`}(\theta_{X,m}-\theta_X).
\end{aligned}$$

Under the alternative of a parameter break this can be written as:

$$\begin{aligned}
T^{d_u-1/2}\sum_{t=1}^{[sT]}\hat{u}_{t,m}^{+} =& T^{d_u-1/2}\sum_{t=1}^{[sT]}\hat{u}_t - T^{d_u-1/2}\sum_{t=1}^{[sT]}v_t\text{`}\Omega_{vv,m}^{-1}\Omega_{vu} \\
&- T^{d_u-1/2}\sum_{t=1}^{[sT]}D_t\text{`}(\hat{\theta_{D,m}}-\theta_D) - T^{d_u-1/2}\sum_{t=1}^{[sT]}X_t\text{`}(\theta_{X,m}-\theta_X) \\
&- T^{d_u-1/2}\sum_{t=[rT]+1}^{[sT]}D_t\text{`}(\hat{\theta_D}-\theta_{D,1}) - T^{d_u-1/2}\sum_{t=[rT]+1}^{[sT]}X_t\text{`}(\theta_X-\theta_{X,1})
\end{aligned}$$

The consistency follows from this representation since under a parameter break at least one of the last two terms diverges while the remaining terms converge in distribution.

Proof of Proposition 2

In this case we obtain

$$T^{d_u-1/2}\sum_{t=1}^{[sT]}\hat{u}_{t,m}^{+} \Rightarrow \omega_{u\cdot v}B_{d_u}(s) + T^{d_X-d_u-1/2}\delta\omega_{\xi}\int_r^s (B_{d_X}(z) - B_{d_X}(r))dz$$

with a non-vanishing extra term on the right hand side, that can be made sufficiently large by choosing δ large enough. Similar considerations hold true for a break in θ_X and not in d_u.

References

Abadir, K. M., Distaso, W., & Giraitis, L. (2009). Two estimators of the long-run variance: Beyond short memory. *Journal of Econometrics*, *150*(1), 56–70.

Andrews, D. W. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica: Journal of the Econometric Society*, 817–858.

Barroso, P., & Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics*, *116*(1), 111–120.

Chu, C.-S. J., Stinchcombe, M., & White, H. (1996). Monitoring structural change. *Econometrica: Journal of the Econometric Society*, 1045–1065.

Daniel, K. D., & Moskowitz, T. J. (2013). Momentum crashes. *Swiss Finance Institute Research Paper*, (13-61), 14–6.

Davidson, J., & Hashimzade, N. (2009). Representation and weak convergence of stochastic integrals with fractional integrator processes. *Econometric Theory*, *25*(6), 1589–1624.

Dierkes, M., & Krupski, J. (2022). Isolating momentum crashes. *Journal of Empirical Finance*, *66*, 1–22.

Fox, R., & Taqqu, M. S. (1987). Multiple stochastic integrals with dependent integrators. *Journal of multivariate analysis*, *21*(1), 105–127.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, *48*(1), 65–91.

Knorre, F., Wagner, M., & Grupe, M. (2021). Monitoring cointegrating polynomial regressions: Theory and application to the environmental kuznets curves for carbon and sulfur dioxide emissions. *Econometrics*, *9*(1), 12.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, *54*(1-3), 159–178.

Leschinski, C. (2017). On the memory of products of long range dependent time series. *Economics Letters*, *153*, 72–76.

Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, *61*(4), 631–653.

Phillips, P. C., & Hansen, B. E. (1990). Statistical inference in instrumental variables regression with i (1) processes. *The Review of Economic Studies*, *57*(1), 99–125.

Robinson, P. M. (2005). Robust covariance matrix estimation: Hac estimates with long memory/antipersistence correction. *Econometric Theory*, *21*(1), 171–180.

Robinson, P. M. (2008). Diagnostic testing for cointegration. *Journal of econometrics*, *143*(1), 206–225.

Robinson, P. M., & Yajima, Y. (2002). Determination of cointegrating rank in fractional systems. *Journal of Econometrics*, *106*(2), 217–241.

Shimotsu, K., & Phillips, P. C. (2005). Exact local whittle estimation of fractional integration.

Shin, Y. (1994). A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric theory*, *10*(1), 91–115.

Trapani, L., & Whitehouse, E. (2020). Sequential monitoring for cointegrating regressions. *arXiv preprint arXiv:2003.12182*.

Wagner, M., & Wied, D. (2015). Monitoring stationarity and cointegration. *Available at SSRN 2624657*.

Wagner, M., & Wied, D. (2017). Consistent monitoring of cointegrating relationships: The us housing market and the subprime crisis. *Journal of Time Series Analysis*, *38*(6), 960–980.

Wang, B., Wang, M., & Chan, N. H. (2015). Residual-based test for fractional cointegration. *Economics Letters*, *126*, 43–46.