Econometric Theory, **28**, 2012, 570–589.
doi:10.1017/S0266466611000661

TESTING FOR A CHANGE IN CORRELATION AT AN UNKNOWN POINT IN TIME USING AN EXTENDED FUNCTIONAL DELTA METHOD

DOMINIK WIED AND WALTER KRÄMER
TU Dortmund

HEROLD DEHLING
Ruhr-Universität Bochum

We propose a new test against a change in correlation at an unknown point in time based on cumulated sums of empirical correlations. The test does not require that inputs are independent and identically distributed under the null. We derive its limiting null distribution using a new functional delta method argument, provide a formula for its local power for particular types of structural changes, give some Monte Carlo evidence on its finite-sample behavior, and apply it to recent stock returns.

1. INTRODUCTION

There is quite a consensus in empirical finance and elsewhere that correlations among returns of all sorts cannot be assumed to remain constant over longer stretches of time (see, among many others, Longin and Solnik, 1995; Krishan, Petkova, and Ritchken, 2009). Although conditional correlations are easily modeled as time-varying in various ways (see, e.g., McAleer, Chan, Hoti, and Lieberman, 2008), including procedures to test for this, unconditional correlations are often taken as constant, which seems to be at odds with various stylized facts from various applications. In particular, correlations among stock returns seem to increase in times of crisis, as evidenced by the most recent joint downturn in stock markets worldwide.

Yet, there is a surprising lack of methods to formally test for a change in correlation. Existing procedures either require strong parametric assumptions (Dias and Embrechts, 2004), assume that potential break points are known (Pearson

Financial support by Deutsche Forschungsgemeinschaft (SFB 823, Statistik nichtlinearer dynamischer Prozesse) is gratefully acknowledged. We are grateful to the co-editor Benedikt M. Pötscher, three anonymous referees, Matthias Arnold, Roland Fried, Werner Ploberger, Christoph Rothe, Tatiana Vlasenco, Daniel Vogel, and Henryk Zähle for helpful criticism and comments. Address correspondence to Dominik Wied, Fakultät Statistik, TU Dortmund, D-44221 Dortmund, Germany; e-mail: wied@statistik.tu-dortmund.de.

 © Cambridge University Press 2011

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

and Wilks, 1933; Jennrich, 1970; Goetzmann, Li, and Rouwenhorst, 2005), or simply estimate correlations from moving windows without giving a formal decision rule (Longin and Solnik, 1995). Only recently, Galeano and Peña (2007) and Aue, Hörmann, Horvath, and Reimherr (2009) have proposed formal tests for a change in covariance structure that do not build upon prior knowledge as to the timing of potential shifts. Both are based on cumulated sums of second-order empirical cross moments (in the vein of Ploberger, Krämer, and Kontrus, 1989) and reject the null of constant covariance structure if these cumulated sums fluctuate too much.

Whereas Galeano and Peña (2007) operate in a parametric setting, the Aue et al. (2009) approach is quite similar to ours. However, the null hypotheses considered here and by Aue et al. (2009) are not identical. It can happen that correlations remain constant but the covariance matrix changes. This distinction is important, e.g., when testing for contagion in international finance (see, e.g., Forbes and Rigobon, 2002). What is termed *shift contagion* in this literature is equivalent in certain models to a change in covariance induced by a change in correlation. Our test will then detect such types of contagion, whereas the Aue et al. (2009) procedure might not.

Similar to Aue et al. (2009), our test statistic is a suitably standardized cumulated sum of empirical correlation coefficients. To derive its asymptotic null distribution, we use a functional delta method argument that has not been considered before, extending conventional functional central limit theorems that are the workhorse in much of the existing literature on structural changes (see, e.g., Ploberger et al., 1989; Ploberger and Krämer, 1990, 1992; Inoue, 2001, just to name a few).

2. THE TEST STATISTIC AND ITS ASYMPTOTIC NULL DISTRIBUTION

Let $(X_t, Y_t), t = 0, \pm 1, \ldots$ be a sequence of bivariate random vectors with finite $(4+\delta)$th absolute moments for some $\delta > 0$. We want to test whether the correlation between X_t and Y_t,

$$\rho_t = \frac{\mathrm{Cov}(X_t, Y_t)}{\sqrt{\mathrm{Var}(X_t)}\sqrt{\mathrm{Var}(Y_t)}},$$

is constant over time in the observation period; i.e., we test

$$H_0 : \rho_t = \rho_0 \quad \forall t \in \{1, \ldots, T\} \quad \text{vs. } H_1 : \exists t \in \{1, \ldots, T-1\} : \rho_t \neq \rho_{t+1}$$

for a constant ρ_0. Our test statistic is

$$Q_T(X, Y) = \hat{D} \max_{2 \leq j \leq T} \frac{j}{\sqrt{T}} |\hat{\rho}_j - \hat{\rho}_T|, \tag{1}$$

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

where

$$\hat{\rho}_k = \frac{\sum_{t=1}^{k}(X_t - \bar{X}_k)(Y_t - \bar{Y}_k)}{\sqrt{\sum_{t=1}^{k}(X_t - \bar{X}_k)^2}\sqrt{\sum_{t=1}^{k}(Y_t - \bar{Y}_k)^2}}$$

and $\bar{X}_k = 1/k\sum_{t=1}^{k} X_t$, $\bar{Y}_k = 1/k\sum_{t=1}^{k} Y_t$. The value $\hat{\rho}_k$ is the empirical correlation coefficient from the first k observations. The scalar $\hat{D}$ is needed for the asymptotic null distribution and will be specified subsequently. The test rejects the null hypothesis of constant correlation if the empirical correlations fluctuate too much, as measured by $\max_{2\leq j\leq T}|\hat{\rho}_j - \hat{\rho}_T|$, with the weighting factor $\frac{j}{\sqrt{T}}$ scaling down deviations at the beginning of the sample where the $\hat{\rho}_j$ are more volatile.

Of course, other functionals of the $\hat{\rho}_j$ series are likewise possible as suitable test statistics, such as some standardized version of

$$\max_{2\leq j\leq T}(\hat{\rho}_j) - \min_{2\leq j\leq T}(\hat{\rho}_j),$$

or simply some suitable average (see Krämer and Schotman, 1992; Ploberger and Krämer, 1992), but for ease of exposition we stick to expression (1) for the purpose of the present paper.

The following technical assumptions are required for the limiting null distribution.

(A1) For $U_t := \left(X_t^2 - E(X_t^2),\ Y_t^2 - E(Y_t^2),\ X_t - E(X_t),\ Y_t - E(Y_t),\ X_tY_t - E(X_tY_t)\right)'$ and $S_j := \sum_{t=1}^{j} U_t$, we have

$$\lim_{T\to\infty} E\left(\frac{1}{T}S_T S_T'\right) =: D_1 \quad \text{(finite and positive definite)}.$$

(A2) The rth absolute moments of the components of U_t are uniformly bounded for some $r > 2$.

(A3) The vector (X_t, Y_t) is L_2-NED (near-epoch dependent) with size $-(r-1)/(r-2)$, where r is from (A2), and constants $(c_t), t \in \mathbb{Z}$, on a sequence $(V_t), t \in \mathbb{Z}$, which is α-mixing of size $\phi^* := -r/(r-2)$, i.e.,

$$||(X_t, Y_t) - E((X_t, Y_t)|\sigma(V_{t-m}, \ldots, V_{t+m}))||_2 \leq c_t v_m$$

with $v_m \to 0$, such that

$$c_t \leq 2||U_t||_2$$

with U_t from Assumption (A2) and the L_2-norm $||\cdot||_2$.

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

(A4) The moments $E(X_t^2)$, $E(Y_t^2)$, $E(X_t)$, $E(Y_t)$, $E(X_tY_t)$ are uniformly bounded and "almost" constant, in the sense that the deviations d_t from the respective constants satisfy

$$\lim_{T\to\infty}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}|d_t| = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}d_t^2 = 0.$$

Of course, (A4) allows for weak stationarity; i.e., $d_t = 0$ for all t. Note that our Assumption (A3) is more general than the dependence assumption of Aue et al. (2009), because in their case, the $(V_t), t \in \mathbb{Z}$, have to be independent. As an alternative to (A4), our main results also hold when variances vary more widely (although not arbitrarily widely), but are proportional to each other, as often happens in financial markets:

(A5) For a bounded function g that is not constant and that can be approximated by step functions such that the function $\int_0^z g(u)du - z\int_0^1 g(u)du$ is different from 0 for at least one $z \in [0, 1]$.

$$E(X_t^2) = a_2 + a_2\frac{1}{\sqrt{T}}g\left(\frac{t}{T}\right),$$
$$E(Y_t^2) = a_3 + a_3\frac{1}{\sqrt{T}}g\left(\frac{t}{T}\right),$$
$$E(X_tY_t) = a_1 + a_1\frac{1}{\sqrt{T}}g\left(\frac{t}{T}\right),$$

while $E(X_t)$ and $E(Y_t)$ remain constant.

As an example, the function g might be piecewise constant with jumps in z_0 from 0 to g_0, which implies that the covariance changes in $[Tz_0]$. Assumption (A4) is violated because $E(X_tY_t) = m_{xy} + d_t$ with

$$\lim_{T\to\infty}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}|d_t| = \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{t}\left|g\left(\frac{t}{T}\right)\right| = \int_0^1|g(u)|du > 0.$$

In the case of (A5), the correlation is constant and equal to $a_1/\sqrt{a_2a_3}$. Note that we operate with a triangular array in this case (see also Section 3 and the beginning of the second section of the Appendix).

Assumption (A3) guarantees that

$$U_t^* := \left(X_t^2, Y_t^2, X_t, Y_t, X_tY_t\right)'$$

is L_2-NED with size $1/2$; see Davidson (1994). It could be modified to ϕ-mixing, requiring only finite fourth moments, but this would admit less dependence than we allow here. In particular, Assumption (A3) allows for generalized autoregressive conditional heteroskedasticity (GARCH) effects (see, e.g., Hansen, 1991; Carrasco and Chen, 2002), which are observed in financial data.

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

Given these assumptions, we next derive the limiting null distribution of our test statistic (1). To this purpose, we first rewrite it as $\sup_{0\leq z\leq 1}|K_T(z)|$ with

$$K_T(z) = \hat{D}\frac{\tau(z)}{\sqrt{T}}(\hat{\rho}_{\tau(z)} - \hat{\rho}_T), \tag{2}$$

where $\tau(z) = [2 + z(T-2)], z \in [0,1]$.

THEOREM 1. *Under H_0 and Assumptions (A1)–(A4) or (A1)–(A3) and (A5), we have*

$$\sup_{0\leq z\leq 1}|K_T(z)| \to_d \sup_{0\leq z\leq 1}|B(z)|,$$

where B is a one-dimensional Brownian bridge.

The explicit form of the distribution function of $\sup_{0\leq z\leq 1}|B(z)|$ is well known (see Billingsley, 1968); its quantiles provide an asymptotic test. The situation in (A5) explicitly allows for changing variances and covariances, while correlations remain constant. This setup is not included in the Aue et al. (2009) test, because the test is based on a different null hypothesis.

As an illustration, consider testing for a break in correlation between the S&P 500 and DAX returns, using daily data from early 2003 to the end of 2009. Figure 1a shows the evolution of the right-hand side of (1), with a maximum value of $Q_T = 2.593$. This is well beyond the 5% critical level (in fact, the corresponding p-value is less than 0.001). Interestingly enough, it is attained just after the Lehman Brothers breakdown on 15 September 2008. Figure 1b plots the evolution of the successively estimated correlations themselves. It is seen that these start at a high level and then they decrease for a long time and increase again after the breakdown, approaching a final value of 0.612.

If we apply the correlation test for the time period early 2004 to the end of 2005, then the null hypothesis of constant correlation is not rejected ($Q_T = 0.546$ with a p-value of 0.927). This fits to the fact that this time period did not contain any dramatic economy changes that would have led to changing correlations. However, the Aue et al. (2009) test (applied with the critical values from Kiefer, 1959) rejects the null hypothesis of constant covariance matrix with a value of the test statistic of 3.409 (p-value less than 0.003) and also the hypothesis of constant marginal variance of the DAX (3.194 with a p-value less than 0.004). Thus, both procedures give complementary information about the comovement of this time series; with our test we conclude that the dependence structure remains constant, which the Aue et al. (2009) test does not show. Note that there are of course other time periods where either both tests reject or neither of the tests rejects.

The proof of Theorem 1 is in the Appendix. Simply applying a standard functional central limit theorem as in Aue et al. (2009) is not possible in the present case. The proof relies on an adapted functional delta method argument that is presented and proved in the Appendix. One major difficulty in the proof is that we

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

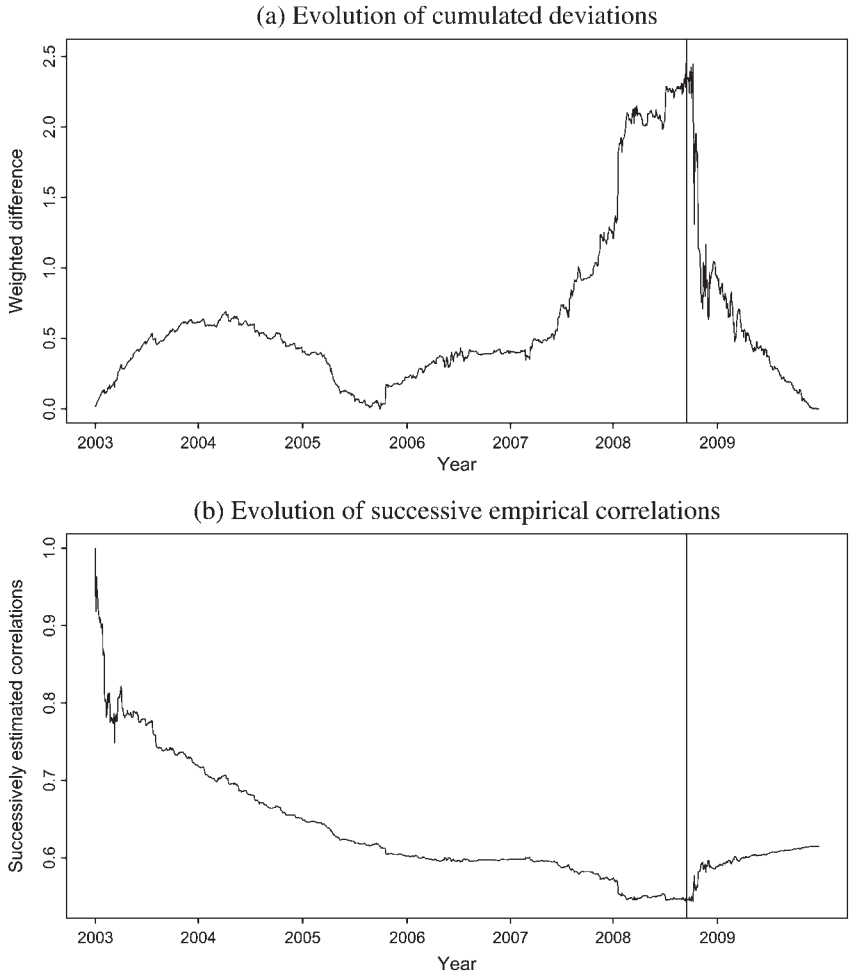


FIGURE 1. Evolution of cumulated deviations and successive empirical correlations.

first have to show convergence on the interval $[\epsilon, 1]$ for arbitrary $\epsilon > 0$ and then show that the statistic vanishes on the interval $[0, \epsilon]$ if ϵ tends to zero.

3. LOCAL POWER

Next, we consider local alternatives of the form

$$\rho_{t,T} = \rho_0 + \frac{1}{\sqrt{T}} g\left(\frac{t}{T}\right) \qquad t = 1, \ldots, T,$$

where g is as in (A5). Using a piecewise constant function g would lead to multiple change points as, e.g., Inoue (2001) deals with. This form of local

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

alternatives is similar to those of Ploberger and Krämer (1990), who analyze local power properties of the Cumulated Sum and Cumulated Sum of Squares test. Now, the random vectors $(X_t, Y_t), t \in \mathbb{Z}$, and $(V_t), t \in \mathbb{Z}$, from Assumption (A3) form a triangular array, but we stick to the former notation for simplicity, i.e., $(X_t, Y_t) := (X_{t,T}, Y_{t,T}), (V_t) := (V_{t,T}), t \in \mathbb{Z}; T = 1, 2, \ldots$. However, with a function g as in (A5), (A4) is replaced by

$$\textbf{(A6)} \quad E(X_t Y_t) = m_{xy} + (1/\sqrt{T}) g\left(\frac{t}{T}\right),$$

$$E(X_t^2) = m_x^2, \qquad E(Y_t^2) = m_y^2, \qquad E(X_t) = \mu_x, \qquad E(Y_t) = \mu_y.$$

This is a special case of the general local alternative, but we stick to it for ease of exposition. As was shown for (A5), Assumption (A4) is violated here also.

THEOREM 2. *Under Assumptions (A1)–(A3) and (A6),*

$$\sup_{z\in[0,1]} \left| \hat{D} \frac{\tau(z)}{\sqrt{T}} (\hat{\rho}_{\tau(z)} - \hat{\rho}_T) \right| \to_d \sup_{z\in[0,1]} |B(z) + C(z)|,$$

where $C(z)$ is a deterministic function that depends on the specific form of the local alternative under consideration, characterized by g.

With this theorem and Anderson's lemma, we can deduce that the asymptotic level is always larger than or equal to α; see Andrews (1997) and Rothe and Wied (2011). The proof is in the Appendix; it relies on similar arguments as the derivation of the null distribution.

The supremum is now taken over the absolute value of a Brownian bridge plus a deterministic function $C(z)$. Its distribution is rather unwieldy, but the local power of the test is easily established for large g. To this purpose, rewrite Assumption (A6) as $g(z) = Mh(z)$ for a function h and a factor M. The function h represents the structural form of the alternative, whereas M captures its amplitude.

COROLLARY 1. *Let $P_{H_1}(M)$ be the rejection probability for given M under the alternative. Let $\epsilon > 0$. Then there is an M_0 such that*

$$\lim_{T\to\infty} P_{H_1}(M) > 1 - \epsilon$$

for all $M > M_0$.

This means that local rejection probabilities become arbitrarily large as structural changes are increasing.

4. SOME FINITE-SAMPLE SIMULATIONS

First, we check the test's finite-sample null distribution. To that purpose, we consider both friendly (i.e., independent and identically distributed [i.i.d.]) and unfriendly environments defined by some serial correlation under the null. Both situations are encompassed by the AR(1) model

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \phi X_{t-1} \\ \phi Y_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t^1 \\ \epsilon_t^2 \end{pmatrix},$$

where $(\epsilon_t^1, \epsilon_t^2)'$ has correlation ρ and $(\epsilon_t^1\ \epsilon_t^2)'$ is i.i.d.-bivariate t_5 to make our data better resemble empirical data such as stock returns. The friendly environment is given by $\phi = 0$ (i.e., i.i.d. observations); the unfriendly environment is given by some nonzero value of ϕ that is compatible with our Assumption (A3). Similar to Aue et al. (2009) we use $\phi = 0.1$. We also ran additional simulations for time-series correlations as large as $\phi = 0.8$, which are likewise allowed for under H_0, given our assumptions. As autocorrelations that large are usually not observed for stock or FX returns, we do not dwell on the resulting figures here; they are available from the corresponding author upon request.

Table 1 gives our results for the cases relevant in practice, based on 5,000 replications and a nominal significance level of 5%. The table shows that convergence to the nominal significance level is not monotone and that the test overrejects quite dramatically for small samples and large values of the correlation coefficient. For correlations in the range that is relevant for, e.g., financial applications, the performance is satisfactory, however. Similar results were obtained with GARCH processes of various types, and again, there is overrejection in small samples. Detailed results are available from the corresponding author upon request.

Next, we check the size-adjusted power of our test for the following alternatives, in which variances remain constantly equal to 1 and correlations change, and compare it to Aue et al. (2009) (maximum-based version, size-adjusted):

1. $\rho_i = 0.5, i \leq \frac{T}{2}$, and $\rho_i = 0.7, i > \frac{T}{2}$,
2. $\rho_i = 0.5, i \leq \frac{T}{4}$, and $\rho_i = 0.7, i > \frac{T}{4}$,

TABLE 1. Empirical rejection frequencies under the null hypothesis

	ρ				
T	−0.9	−0.5	0	0.5	0.9
(a) i.i.d. case					
200	0.144	0.054	0.039	0.053	0.142
500	0.064	0.040	0.035	0.041	0.064
1,000	0.048	0.038	0.034	0.039	0.049
2,000	0.043	0.038	0.036	0.038	0.043
(b) Serial dependence					
200	0.150	0.052	0.038	0.053	0.144
500	0.070	0.049	0.037	0.044	0.064
1,000	0.055	0.046	0.039	0.044	0.053
2,000	0.046	0.043	0.043	0.045	0.051

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

TABLE 2. Empirical size-adjusted rejection frequencies when correlations change

	Alternative				
T	1	2	3	4	5
(a) Our test					
200	0.309	0.255	0.953	0.880	0.101
500	0.587	0.488	0.996	0.989	0.207
1,000	0.830	0.733	0.998	0.998	0.422
2,000	0.967	0.928	1	0.999	0.750
(b) Aue et al. (2009)					
200	0.241	0.159	0.977	0.842	0.083
500	0.586	0.400	1	0.998	0.163
1,000	0.858	0.690	1	1	0.284
2,000	0.980	0.929	1	1	0.609

3. $\rho_i = 0.5, i \leq \frac{T}{2}$, and $\rho_i = -0.5, i > \frac{T}{2}$,
4. $\rho_i = 0.5, i \leq \frac{T}{4}$, and $\rho_i = -0.5, i > \frac{T}{4}$,
5. $\rho_i = 0.5, i \leq \frac{T}{4}$, and $\rho_i = 0.7, \frac{T}{4} < i \leq \frac{3}{4}T$, and $\rho_i = 0.5, i > \frac{3}{4}T$.

Table 2 gives the results. It is seen that power increases rapidly with sample size, and not surprisingly, it is highest when the break in ρ is largest (i.e, ρ changes from -0.5 to 0.5). In general, the powers of both tests lie closely together, although the power of our test is often slightly higher than the power of the Aue et al. (2009) test for rather small samples and in situations when the change points appear early and is always higher in the case of several change points.

5. DISCUSSION

In this paper, we have proposed a fluctuation test for constant correlation that works under rather general assumptions. A shortcoming of our test, which it shares with Aue et al. (2009), is the requirement of finite fourth moments. Although there is evidence that many financial returns have finite second and even third moments, the existence of fourth moments remains doubtful; see Krämer (2002). If the fourth moment of one or both of the components of (X_t, Y_t) does not exist, our functional central limit theorem, from which we derive the null distribution of our test, would not apply. As the asymptotic distribution of empirical cross covariances and cross correlations would then be different (see, e.g., Davis and Mikosch, 1998; Scheffler and Meerschaert, 2001), this condition of finite fourth moments is also necessary for our limit results.

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

A multivariate extension of our approach, i.e., testing for the constancy of a whole correlation matrix, is possible by doing pairwise comparisons and rejecting the null hypothesis, e.g., if the maximum of the test statistics is too large. Circumventing the resulting multiple testing problem, however, requires some additional theory that goes beyond the scope of the present paper. The problem might be overcome by relying on other multivariate dependence measures as in Schmid, Schmidt, Blumentritt, Gaissler, and Ruppert (2010). A general approach based on multivariate empirical distributions can be found in Inoue (2001).

REFERENCES

Andrews, D. (1997) A conditional Kolmogorov test. *Econometrica* 65, 1097–1128.

Aue, A., S. Hörmann, L. Horvath, & M. Reimherr (2009) Break detection in the covariance structure of multivariate time series models. *Annals of Statistics* 37, 4046–4087.

Billingsley, P. (1968) *Convergence of Probability Measures*. Wiley.

Carrasco, M. & X. Chen (2002) Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* 18, 17–39.

Davidson, J. (1994) *Stochastic Limit Theory*. Oxford University Press.

Davis, R. & T. Mikosch (1998) The sample autocorrelations of heavy-tailed processes with applications to ARCH. *Annals of Statistics* 26, 2049–2080.

De Jong, R. & J. Davidson (2000) Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. *Econometrica* 68, 407–424.

Dias, A. & P. Embrechts (2004) Change point analysis for dependence structures in finance and insurance. In G. Szegö (ed.), *Risk Measures of the 21th Century*, pp. 321–335. Wiley.

Forbes, K. & R. Rigobon (2002) No contagion, only interdependence: Measuring stock market comovements. *Journal of Finance* 57, 2223–2261.

Galeano, P. & D. Peña (2007) Covariance changes detection in multivariate time series. *Journal of Statistical Planning and Inference* 137, 194–211.

Gill, R. (1989) Non- and semi-parametric maximum likelihood estimators and the von Mises method—Part 1. *Scandinavian Journal of Statistics* 16, 97–128.

Goetzmann, W., L. Li, & K. Rouwenhorst (2005) Long-term global market correlations. *Journal of Business* 78, 1–38.

Hansen, B. (1991) GARCH(1,1) processes are near-epoch dependent. *Economics Letters* 36, 181–186.

Inoue, A. (2001) Testing for distributional change in time series. *Econometric Theory* 17, 156–187.

Jennrich, R. (1970) An asymptotic chi-square test for the equality of two correlation matrices. *Journal of the American Statistical Association* 65, 904–912.

Kiefer, J. (1959) K-sample analogues of the Kolmogorov-Smirnov and Cramér-V. Mises Tests. *Annals of Mathematical Statistics* 30, 420–447.

Krämer, W. (2002) Statistische Besonderheiten von Finanzzeitreihen. *Jahrbücher für Nationalökonomie und Statistik* 222, 210–229.

Krämer, W. & P. Schotman (1992) Range vs. maximum in the OLS-based version of the CUSUM test. *Economics Letters* 40, 379–381.

Krishan, C., R. Petkova, & P. Ritchken (2009) Correlation risk. *Journal of Empirical Finance* 16, 353–367.

Longin, F. & B. Solnik (1995) Is the correlation in international equity returns constant: 1960–1990? *International Money and Finance* 14, 3–26.

McAleer, M., F. Chan, S. Hoti, & O. Lieberman (2008) Generalized autoregressive conditional correlation. *Econometric Theory* 24, 1554–1583.

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

Pearson, E. & S. Wilks (1933) Methods of statistical analysis appropriate for k samples of two variables. *Biometrika* 25, 353–378.

Ploberger, W. & W. Krämer (1990) The local power of the CUSUM and CUSUM of squares tests. *Econometric Theory* 6, 335–347.

Ploberger, W. & W. Krämer (1992) The CUSUM-test with OLS residuals. *Econometrica* 60, 271–285.

Ploberger, W., W. Krämer, & K. Kontrus (1989) A new test for structural stability in the linear regression model. *Journal of Econometrics* 40, 307–318.

Rothe, C. & D. Wied (2011) Misspecification Testing in a Class of Conditional Distributional Models. SFB 823 Discussion paper, TU Dortmund.

Scheffler, H.-P. & M. Meerschaert (2001) Sample cross-correlations for moving averages with regularly varying tails. *Journal of Time Series Analysis* 22, 481–492.

Schmid, F., R. Schmidt, T. Blumentritt, S. Gaissler, & M. Ruppert (2010) Copula-based measures of multivariate association. In F. Durante, W. Härdle, P. Jaworski, & T. Rychlik (eds.), *Workshop on Copula Theory and Its Applications*, pp. 209–236. Springer-Verlag.

van der Vaart, A. (1998) *Asymptotic Statistics*. Cambridge University Press.

Wooldridge, J. & H. White (1988) Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory* 4, 210–230.

APPENDIX

A.1. The Scalar $\hat{\mathbf{D}}$ from the Test Statistic. The scalar $\hat{D}$ from our test statistic (1) can be written as

$$\hat{D} = (\hat{F}_1 \hat{D}_{3,1} + \hat{F}_2 \hat{D}_{3,2} + \hat{F}_3 \hat{D}_{3,3})^{-1/2},$$

where

$$(\hat{F}_1\ \hat{F}_2\ \hat{F}_3) = \begin{pmatrix} \hat{D}_{3,1}\hat{E}_{11} + \hat{D}_{3,2}\hat{E}_{21} + \hat{D}_{3,3}\hat{E}_{31} \\ \hat{D}_{3,1}\hat{E}_{12} + \hat{D}_{3,2}\hat{E}_{22} + \hat{D}_{3,3}\hat{E}_{32} \\ \hat{D}_{3,1}\hat{E}_{13} + \hat{D}_{3,2}\hat{E}_{23} + \hat{D}_{3,3}\hat{E}_{33} \end{pmatrix}',$$

$$\hat{E}_{11} = \hat{D}_{1,11} - 4\hat{\mu}_x \hat{D}_{1,13} + 4\hat{\mu}_x^2 \hat{D}_{1,33},$$

$$\hat{E}_{12} = \hat{E}_{21} = \hat{D}_{1,12} - 2\hat{\mu}_x \hat{D}_{1,23} - 2\hat{\mu}_y \hat{D}_{1,14} + 4\hat{\mu}_x \hat{\mu}_y \hat{D}_{1,34},$$

$$\hat{E}_{22} = \hat{D}_{1,22} - 4\hat{\mu}_y \hat{D}_{1,24} + 4\hat{\mu}_y^2 \hat{D}_{1,44},$$

$$\hat{E}_{13} = \hat{E}_{31} = -\hat{\mu}_y \hat{D}_{1,13} + 2\hat{\mu}_x \hat{\mu}_y \hat{D}_{1,33} - \hat{\mu}_x \hat{D}_{1,14} + 2\hat{\mu}_x^2 \hat{D}_{1,34} + \hat{D}_{1,15} - 2\hat{\mu}_x \hat{D}_{1,35},$$

$$\hat{E}_{23} = \hat{E}_{32} = -\hat{\mu}_y \hat{D}_{1,23} + 2\hat{\mu}_x \hat{\mu}_y \hat{D}_{1,44} - \hat{\mu}_x \hat{D}_{1,24} + 2\hat{\mu}_y^2 \hat{D}_{1,34} + \hat{D}_{1,25} - 2\hat{\mu}_y \hat{D}_{1,45},$$

$$\hat{E}_{33} = \hat{\mu}_y^2 \hat{D}_{1,33} + 2\hat{\mu}_x \hat{\mu}_y \hat{D}_{1,34} - 2\hat{\mu}_y \hat{D}_{1,35} + \hat{\mu}_x^2 \hat{D}_{1,44} + \hat{D}_{1,55} - 2\hat{\mu}_x \hat{D}_{1,45},$$

$$\hat{D}_1 = \begin{pmatrix} \hat{D}_{1,11} & \hat{D}_{1,12} & \hat{D}_{1,13} & \hat{D}_{1,14} & \hat{D}_{1,15} \\ \hat{D}_{1,21} & \hat{D}_{1,22} & \hat{D}_{1,23} & \hat{D}_{1,24} & \hat{D}_{1,25} \\ \hat{D}_{1,31} & \hat{D}_{1,32} & \hat{D}_{1,33} & \hat{D}_{1,34} & \hat{D}_{1,35} \\ \hat{D}_{1,41} & \hat{D}_{1,42} & \hat{D}_{1,43} & \hat{D}_{1,44} & \hat{D}_{1,45} \\ \hat{D}_{1,51} & \hat{D}_{1,52} & \hat{D}_{1,53} & \hat{D}_{1,54} & \hat{D}_{1,55} \end{pmatrix} = \sum_{t=1}^{T} \sum_{u=1}^{T} k\left(\frac{t-u}{\gamma_T}\right) V_t V_u',$$

$$V_t = \frac{1}{\sqrt{T}} U_t^{***}, \qquad \gamma_T = [\log T],$$

$$U_t^{***} = \left(X_t^2 - \overline{(X^2)}_T\ Y_t^2 - \overline{(Y^2)}_T\ X_t - \bar{X}_T\ Y_t - \bar{Y}_T\ X_t Y_t - \overline{(XY)}_T\right)',$$

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

$$k(x) = \begin{cases} 1-|x|, & |x| \le 1 \\ 0, & \text{otherwise} \end{cases},$$

$$\hat{\mu}_x = \bar{X}_T, \qquad \hat{\mu}_y = \bar{Y}_T, \qquad \hat{D}_{3,1} = -\frac{1}{2}\frac{\hat{\sigma}_{xy}}{\hat{\sigma}_y}\hat{\sigma}_x^{-3},$$

$$\hat{D}_{3,2} = -\frac{1}{2}\frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x}\hat{\sigma}_y^{-3}, \qquad \hat{D}_{3,3} = \frac{1}{\hat{\sigma}_x\hat{\sigma}_y},$$

and

$$\hat{\sigma}_x^2 = \overline{(X^2)}_T - (\bar{X}_T)^2, \qquad \hat{\sigma}_y^2 = \overline{(Y^2)}_T - (\bar{Y}_T)^2, \qquad \hat{\sigma}_{xy} = \overline{(XY)}_T - \bar{X}_T\bar{Y}_T.$$

Proof. See the discussion preceding Lemma A.3. ■

A.2. Proof of Theorem 1 under (A1)–(A4). The proof of Theorem 1 requires several lemmas. In all proofs, we assume $(X_t, Y_t), t \in \mathbb{Z}$, to be a triangular array, which generalizes the assumptions in Section 2 and corresponds to the assumptions in Section 3. The array is defined on the probability space $(\Omega, \mathfrak{A}, \mathbb{P})$.

Let I be some interval, e.g., $I = [\epsilon, 1]$ for some $\epsilon \in [0, 1)$. For an integer $k \ge 1$, let $D(I, \mathbb{R}^k)$ be the set of all functions $\theta : I \to \mathbb{R}^k$ that are càdlàg in each of the k components, equipped with the multidimensional supremum norm

$$||\theta||_\infty := \sup_{z \in T} ||\theta(z)||,$$

where $||\cdot||$ denotes the maximum norm in $\mathbb{R}^k$. Let in addition $m_x^2, m_y^2, \mu_x, \mu_y, m_{xy}$ be the respective constants from (A4) and

$$\sigma_x := \sqrt{m_x^2 - \mu_x^2}, \qquad \sigma_y := \sqrt{m_y^2 - \mu_y^2}, \qquad \sigma_{xy} := m_{xy} - \mu_x\mu_y.$$

The first lemma is a straightforward application of the functional central limit theorem in Davidson (1994, p. 492) that relies on a univariate invariance principle from Wooldridge and White (1988).

LEMMA A.1. *On $D([\epsilon, 1], \mathbb{R}^5)$, for arbitrary $\epsilon > 0$, we have*

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\tau(\cdot)} U_t^{**} := \frac{1}{\sqrt{T}}\sum_{t=1}^{\tau(\cdot)} \begin{pmatrix} X_t^2 - m_x^2 \\ Y_t^2 - m_y^2 \\ X_t - \mu_x \\ Y_t - \mu_y \\ X_tY_t - m_{xy} \end{pmatrix} = \frac{\tau(\cdot)}{\sqrt{T}} \begin{pmatrix} \overline{(X^2)}_{\tau(\cdot)} - m_x^2 \\ \overline{(Y^2)}_{\tau(\cdot)} - m_y^2 \\ \bar{X}_{\tau(\cdot)} - \mu_x \\ \bar{Y}_{\tau(\cdot)} - \mu_y \\ \overline{(XY)}_{\tau(\cdot)} - m_{xy} \end{pmatrix}$$

$$=: U(\cdot) \to_d D_1^{\frac{1}{2}} W_5(\cdot),$$

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

where $W_k(\cdot)$ is a k-dimensional Brownian motion and

$$D_1 = D_1' = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \sum_{u=1}^{T}$$

$$\begin{pmatrix} \mathrm{Cov}(X_t^2, X_u^2) & \mathrm{Cov}(X_t^2, Y_u^2) & \mathrm{Cov}(X_t^2, X_u) & \mathrm{Cov}(X_t^2, Y_u) & \mathrm{Cov}(X_t^2, X_u Y_u) \\ \cdot & \mathrm{Cov}(Y_t^2, Y_u^2) & \mathrm{Cov}(Y_t^2, X_u) & \mathrm{Cov}(Y_t^2, Y_u) & \mathrm{Cov}(Y_t^2, X_u Y_u) \\ \cdot & \cdot & \mathrm{Cov}(X_t, X_u) & \mathrm{Cov}(X_t, Y_u) & \mathrm{Cov}(X_t, X_u Y_u) \\ \cdot & \cdot & \cdot & \mathrm{Cov}(Y_t, Y_u) & \mathrm{Cov}(Y_t, X_u Y_u) \\ \cdot & \cdot & \cdot & \cdot & \mathrm{Cov}(X_t Y_t, X_u Y_u) \end{pmatrix}.$$

Proof.

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{\tau(z)} \begin{pmatrix} X_t^2 - m_x^2 \\ Y_t^2 - m_y^2 \\ X_t - \mu_x \\ Y_t - \mu_y \\ X_t Y_t - m_{xy} \end{pmatrix} = \frac{1}{\sqrt{T}} \sum_{t=1}^{\tau(z)} \begin{pmatrix} X_t^2 - E(X_t^2) \\ Y_t^2 - E(Y_t^2) \\ X_t - E(X_t) \\ Y_t - E(Y_t) \\ X_t Y_t - E(X_t Y_t) \end{pmatrix} + \frac{1}{\sqrt{T}} \sum_{t=1}^{\tau(z)} \begin{pmatrix} E(X_t^2) - m_x^2 \\ E(Y_t^2) - m_y^2 \\ E(X_t) - \mu_x \\ E(Y_t) - \mu_y \\ E(X_t Y_t) - m_{xy} \end{pmatrix}$$

$$=: A_1 + A_2.$$

Consider the first component of A_2:

$$A_{2,1}(z) = \frac{1}{\sqrt{T}} \sum_{t=1}^{\tau(z)} \left(E(X_t^2) - m_x^2 \right).$$

We have

$$||A_{2,1}(z)||_\infty \leq \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left| E(X_t^2) - m_x^2 \right|.$$

With Assumption (A4), the right-hand side tends to zero. In the same way, all other components of A_2 converge to zero, hence also A_2.

The sum in A_1 can be separated into one term from $i = 1$ to $[Tz]$, called A_3, and one term from $[Tz+1]$ to $[Tz+(1-z)2]$, called A_4. We show that A_4 converges in probability to the zero function in the supremum norm. To this end, we first show that for fixed z, A_4 converges to zero in probability. If

$$[Tz + (1-z)2] < [Tz + 1],$$

A_4 is equal to zero. For

$$[Tz + (1-z)2] \geq [Tz + 1]$$

our argument builds on the Markov inequality. Matrix A_4 consists of two terms at most so that we have for the first component A_{41}

$$\sup_{z\in[\epsilon,1]} |A_{41}(z)| \leq \frac{2}{\sqrt{T}} \max_{1\leq t\leq T} |X_t^2 - E(X_t^2)|.$$

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

Now, for arbitrary $\delta > 0$

$$\mathbb{P}\Big(\sup_{z\in[\epsilon,1]} |A_{41}(z)| > \delta\Big) \le T \max_{1\le t\le T} \mathbb{P}\left(|X_t^2 - E(X_t^2)| > \frac{\sqrt{T}}{2\delta}\right).$$

By Assumption (A2) $E|X_t^2 - E(X_t^2)|^r$ is uniformly bounded for $r > 2$. With Markov's inequality the right-hand side converges to zero.

The same argument applies for the other components of A_4. Therefore, all finite-dimensional distributions converge to zero in probability and thus in distribution. We show the tightness of the process similarly to the method on page 138 in Billingsley (1968). At first, we show the tightness of every single component (without loss of generality for the first); the tightness of the whole vector follows.

$$B := E\left(|A_{41}(z) - A_{41}(z_1)|^{1+\alpha/2} \cdot |A_{41}(z_2) - A_{41}(z)|^{1+\alpha/2}\right) \le \frac{1}{T^{1+\alpha/2}} C$$

for $\epsilon \le z_1 \le z \le z_2 \le 1$ and a constant C because of uniform boundedness. If $[Tz_2] - [Tz_1] = 0$, then $B = 0$. If $[Tz_2] - [Tz_1] \ge 1$, we get

$$\frac{1}{T^{1+\alpha/2}} C \le ([Tz_2] - [Tz_1])^{1+\alpha/2} \frac{1}{T^{1+\alpha/2}} C = C\left(\frac{[Tz_2] - [Tz_1]}{T}\right)^{1+\alpha/2},$$

and the condition of Theorem 15.6 in Billingsley (1968) holds. Thus, A_4 as a process converges in distribution (and also in probability) to the zero function. We apply to A_3 the multivariate invariance principle from Davidson (1994, p. 492) that relies on a univariate invariance principle from Wooldridge and White (1988). The value $c_{T,i}^{\lambda}$ in this theorem is given by $(\lambda' D_1^{-1} T^{-1} \lambda)^{1/2}$ in our case. The theorem is actually given for a function space equipped with Skorokhod metric, but it also holds in our uniform topology because the limit process is continuous almost surely; see Gill (1989, p. 106) for details.

With the continuous mapping theorem (CMT) (see van der Vaart, 1998, p. 259), the lemma follows. ■

LEMMA A.2. *On* $D([\epsilon, 1], \mathbb{R})$*, for arbitrary* $\epsilon > 0$*,*

$$\frac{\tau(\cdot)}{\sqrt{T}}(\hat{\rho}_{\tau(\cdot)} - \rho_0^*) \to_d D_3 D_2 D_1^{1/2} W_5(\cdot),$$

where

$$\rho_0^* = \frac{\sigma_{xy}}{\sigma_x \sigma_y},$$

$$D_2 = \begin{pmatrix} 1 & 0 & -2\mu_x & 0 & 0 \\ 0 & 1 & 0 & -2\mu_y & 0 \\ 0 & 0 & -\mu_y & -\mu_x & 1 \end{pmatrix}, \quad \textit{and}$$

$$D_3 = \begin{pmatrix} -\frac{1}{2}\frac{\sigma_{xy}}{\sigma_y}\sigma_x^{-3} & -\frac{1}{2}\frac{\sigma_{xy}}{\sigma_x}\sigma_y^{-3} & \frac{1}{\sigma_x \sigma_y} \end{pmatrix}.$$

The proof of Lemma A.2 relies on a functional delta method argument that we present in a form slightly more general than is actually needed.

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

THEOREM A.1 (Extended Functional Delta Method). *Consider a sequence $(\theta_T)_T$ of functions in $D\big(I,\mathbb{R}^k\big)$ converging uniformly to a function $\theta \in D\big(I,\mathbb{R}^k\big)$. Furthermore, let $(s_T)_T$ be a sequence of functions $s_T : I \to \mathbb{R}\backslash\{0\}$ such that $||s_T^{-1}||_\infty \to 0$ and let $(M_T)_T$ be stochastic processes on I with values in $\mathbb{R}^k$ and bounded sample paths such that*

$$||Z_T||_\infty = O_p(1) \quad \text{with } Z_T := s_T(M_T - \theta_T).$$

Furthermore, let $f : \mathbb{R}^k \to \mathbb{R}^l$ be a mapping that is continuously differentiable on an open set $\Omega \subset \mathbb{R}^k$. Suppose that

$$\overline{\theta(I)} \quad \text{is a compact subset of } \Omega,$$

where $\overline{\theta(I)}$ stands for the closure of the set $\{\theta(t) : t \in I\}$ in $\mathbb{R}^k$. Then it holds that

(i) $s_T(\cdot)\,(f(M_T(\cdot)) - f(\theta_T(\cdot))) = Df(\theta(\cdot))Z_T(\cdot) + R_T$ *with a stochastic process such that*

$$||R_T||_\infty = o_p(1).$$

(ii) If Z_T also converges in distribution (in $D(I,\mathbb{R}^k)$) to a stochastic process Z, then

$$s_T(\cdot)\,(f(M_T(\cdot)) - f(\theta_T(\cdot))) \to_d Df(\theta(\cdot))Z(\cdot).$$

Proof. Assertion (ii) immediately follows from assertion (i) with the usual CMT.

To prove the expansion from Assertion (i), note that for any $z \in I$,

$$\begin{aligned} R_T(z) &:= s_T(z)\,(f(M_T(z)) - f(\theta_T(z))) - Df(\theta(z))Z_T(z) \\ &= s_T(z)\left(f\left(\theta_T(z) + s_T^{-1}(z)Z_T(z)\right) - f(\theta_T(z))\right) - Df(\theta(z))Z_T(z) \\ &= \int_0^1 Df\left(\theta_T(z) + us_T^{-1}(z)Z_T(z)\right)Z_T(z)du - Df(\theta(z))Z_T(z) \\ &= \int_0^1 \left(Df\left(\theta_T(z) + us_T^{-1}(z)Z_T(z)\right) - Df(\theta(z))\right)du \cdot Z_T(z), \end{aligned} \tag{A.1}$$

provided that

$$r_T := ||\theta_T - \theta||_\infty + ||s_T^{-1}||_\infty ||Z_T||_\infty = o_p(1)$$

is smaller than

$$\rho := \inf_{x \in \overline{\theta(I)},\, y \in \mathbb{R}^k \backslash \Omega} ||x - y|| > 0.$$

The latter condition is needed for (A.1) to be well defined.

Hence

$$||R_T||_\infty \le \sup\left\{||Df(y) - Df(x)|| : x \in \overline{\theta(I)},\, y \in \mathbb{R}^k, ||y - x|| \le r_T\right\} \cdot ||Z_T||_\infty. \tag{A.2}$$

Here $||Df(y) - Df(x)||$ is the usual operator norm of the matrix $Df(y) - Df(x)$ in the case of $y \in \Omega$. (In the case of $y \notin \Omega$ define $||Df(y) - Df(x)|| = \infty$.) One can easily

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

deduce from continuity of $Df(\cdot)$ on Ω, compactness of $\overline{\theta(T)} \in \Omega$, and $r_T = o_p(1)$ that the right-hand side of (A.2) converges to zero in probability. ■

Proof of Lemma A.2. We apply the extended functional delta method from Theorem A.1 twice to $U(\cdot)$ with $I = [\epsilon, 1]$ for arbitrary $\epsilon > 0$ and $s_T(\cdot) = \frac{\tau(z)}{\sqrt{T}}$. The first transformation of $U(\cdot)$ is

$$f_1 : \mathbb{R}^5 \to \mathbb{R}^3,$$

$$f_1(x_1, x_2, x_3, x_4, x_5) = \begin{pmatrix} x_1 - x_3^2 \\ x_2 - x_4^2 \\ x_5 - x_3 x_4 \end{pmatrix}$$

with, for $\theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$,

$$Df_1(\theta) = \begin{pmatrix} 1 & 0 & -2\theta_3 & 0 & 0 \\ 0 & 1 & 0 & -2\theta_4 & 0 \\ 0 & 0 & -\theta_4 & -\theta_3 & 1 \end{pmatrix}.$$

The second transformation is

$$f_2 : \mathbb{R}^3 \to \mathbb{R},$$

$$f_2(x_1, x_2, x_3) = \frac{x_3}{\sqrt{x_1 x_2}}$$

with, for $\theta = (\theta_1, \theta_2, \theta_3)$,

$$Df_2(\theta) = \left(-\tfrac{1}{2} \tfrac{\theta_3}{\sqrt{\theta_2}} \theta_1^{-3/2} \quad -\tfrac{1}{2} \tfrac{\theta_3}{\sqrt{\theta_1}} \theta_2^{-3/2} \quad \tfrac{1}{\sqrt{\theta_1 \theta_2}} \right).$$

The lemma follows then from the fact that $U(\cdot)$ converges in distribution to the stochastic process $D_1^{1/2} W_5(\cdot)$. We need the restriction on $[\epsilon, 1]$ because the function $r_T(z) = \frac{\sqrt{T}}{\tau(z)}$ would not tend to 0 in the supremum norm on $[0, 1]$. ■

Now, one can show that

$$(D_3 D_2 D_1 D_2' D_3')^{-1/2} \frac{\tau(\cdot)}{\sqrt{T}} (\hat{\rho}_{\tau(\cdot)} - \rho_0^*) \to_d W_1(\cdot)$$

on $D[\epsilon, 1]$, where W_1 is a one-dimensional Brownian motion. The parameter $(D_3 D_2 D_1 D_2' D_3')^{-1/2}$ has to be estimated consistently. It is a continuous composition of moments of X_i and Y_i that appear in the matrices D_3 and

$$E = \begin{pmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{pmatrix} := D_2 D_1 D_2'$$

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

with

$$
\begin{aligned}
E_{11} &= D_{1,11} - 4\mu_x D_{1,13} + 4\mu_x^2 D_{1,34},\\
E_{12} &= E_{21} = D_{1,12} - 2\mu_x D_{1,23} - 2\mu_y D_{1,14} + 4\mu_x\mu_y D_{1,34},\\
E_{22} &= D_{1,22} - 4\mu_y D_{1,24} + 4\mu_y^2 D_{1,44},\\
E_{13} &= E_{31} = -\mu_y D_{1,13} + 2\mu_x\mu_y D_{1,33} - \mu_x D_{1,14} + 2\mu_x^2 D_{1,34} + D_{1,15} - 2\mu_x D_{1,35},\\
E_{23} &= E_{32} = -\mu_y D_{1,23} + 2\mu_x\mu_y D_{1,44} - \mu_x D_{1,24} + 2\mu_y^2 D_{1,34} + D_{1,25} - 2\mu_y D_{1,45},\\
E_{33} &= -\mu_y^2 D_{1,33} + 2\mu_x\mu_y D_{1,34} - 2\mu_y D_{1,35} + \mu_x^2 D_{1,44} + D_{1,55} - 2\mu_x D_{1,45}.
\end{aligned}
$$

In addition,

$$
\begin{aligned}
D_3 D_2 D_1 D_2' &= \begin{pmatrix} D_{3,1}E_{11} + D_{3,2}E_{21} + D_{3,3}E_{31} \\ D_{3,1}E_{12} + D_{3,2}E_{22} + D_{3,3}E_{32} \\ D_{3,1}E_{13} + D_{3,2}E_{23} + D_{3,3}E_{33} \end{pmatrix}'\\
&=: \begin{pmatrix} F_1 & F_2 & F_3 \end{pmatrix}
\end{aligned}
$$

and

$$
(D_3 D_2 D_1 D_2' D_3')^{-1/2} = (F_1 D_{3,1} + F_2 D_{3,2} + F_3 D_{3,3})^{-1/2}.
$$

Thus, $(D_3 D_2 D_1 D_2' D_3')^{-1/2}$ is a continuous composition of moments of X_i and Y_i from the matrices D_3 and E. The only nontrivial task is to estimate the matrix D_1 consistently because the other elements can easily be estimated by maximum likelihood. We solve the problem with a kernel estimator proposed by de Jong and Davidson (2000) using the bandwidth $\gamma_T = [\log T]$ and the Bartlett kernel $k(\cdot)$ with

$$
k(x) = \begin{cases} 1 - |x|, & |x| \le 1 \\ 0, & \text{otherwise} \end{cases}.
$$

Other choices would also be possible, but we restrict to our choice for ease of exposition. Because of Assumption (A4), D_1 is asymptotically equivalent to

$$
\hat{\Sigma}_T = \sum_{t=1}^{T}\sum_{u=1}^{T} k\left(\frac{t-u}{\gamma_T}\right) V_t V_u'
$$

with $V_t = \left(1\Big/\sqrt{T}\right) U_t^{**}$ and U_t^{**} from Lemma A.1. The vector U_t^{**} depends on $\theta_0 = \left(m_x^2\ m_y^2\ \mu_x\ \mu_y\ m_{xy}\right) \in (\mathbb{R} \cap (0,\infty))^2 \times \mathbb{R}^3$; a consistent estimator for this is the sequence

$$
\theta_T = \left(\overline{(X^2)}_T\ \overline{(Y^2)}_T\ \bar{X}_T\ \bar{Y}_T\ \overline{(XY)}_T\right).
$$

Thus, from de Jong and Davidson (2000), we get a consistent estimator given in the first section of this Appendix. Now, we extend the convergence result to the interval [0, 1].

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

LEMMA A.3. *On* $D([0,1],\mathbb{R})$,

$$W_T(\cdot) := \hat{D}\frac{\tau(\cdot)}{\sqrt{T}}(\hat{\rho}_{\tau(\cdot)} - \rho_0^*) \to_d W_1(\cdot).$$

Proof. Consider the following functions:

$$W_T^\epsilon(z) = \begin{cases} W_T(z), & z \geq \epsilon \\ 0 & z < \epsilon \end{cases},$$

$$W^\epsilon(z) = \begin{cases} W_1(z), & z \geq \epsilon \\ 0 & z < \epsilon \end{cases}.$$

The previous lemmas then imply that

$$W_T^\epsilon(\cdot) \to_d W^\epsilon(\cdot)$$

for $T \to \infty$ on $D[0,1]$ and also

$$W^\epsilon(\cdot) \to_d W_1(\cdot)$$

for rational $\epsilon \to 0$ on $D[0,1]$. The convergence of $W_T(\cdot)$ on $D([0,1],\mathbb{R})$ then follows from Theorem 4.2 in Billingsley (1968) if we can show that

$$\lim_{\epsilon\to 0}\limsup_{T\to\infty}\mathbb{P}\Big(\sup_{z\in[0,1]}|W_T^\epsilon(z) - W_T(z)| \geq \eta\Big) = \lim_{\epsilon\to 0}\limsup_{T\to\infty}\mathbb{P}\Big(\sup_{z\in[0,\epsilon]}|W_T(z)| \geq \eta\Big) = 0$$

for all $\eta > 0$. Note that the separability condition of this theorem is not necessary in our case, because $\sup_{z\in I}|A(z)|$ is always a random variable when $A(\cdot)$ is a right-continuous random function. Now,

$$\begin{aligned}
&\sup_{z\in[0,\epsilon]}|W_T(z)| \\
&= \sup_{z\in[0,\epsilon]}\left|\frac{\frac{1}{\sqrt{T}}\sum_{t=1}^{\tau(z)}(X_t - \bar{X}_{\tau(z)})(Y_t - \bar{Y}_{\tau(z)}) - \frac{\rho_0}{\sqrt{T}}\sqrt{\frac{1}{\tau(z)}\sum_{t=1}^{\tau(z)}(X_t - \bar{X}_{\tau(z)})^2\frac{1}{\tau(z)}\sum_{t=1}^{\tau(z)}(Y_t - \bar{Y}_{\tau(z)})^2}}{\sqrt{\frac{1}{\tau(z)}\sum_{t=1}^{\tau(z)}(X_t - \bar{X}_{\tau(z)})^2\frac{1}{\tau(z)}\sum_{t=1}^{\tau(z)}(Y_t - \bar{Y}_{\tau(z)})^2}}\right| \\
&=: \sup_{z\in[0,\epsilon]}\left|\frac{D_1(z)}{D_2(z)}\right|.
\end{aligned}$$

By the strong law of large numbers and the CMT, D_2 goes to $\sigma_x\sigma_y$ almost surely for fixed $z > 0$. The same holds for $\bar{X}_T$ and $\bar{Y}_T$ with the limit μ_x and μ_y. Now let $\delta > 0$ be arbitrary. By Egoroff's theorem (see Davidson, 1994, Thm. 18.4), there is a set $\Omega_\delta \subset \Omega$ with $\mathbb{P}(\Omega_\delta) \geq 1-\delta$ and a number $M(\delta) > 0$ such that $|D_1(z) - \sigma_x\sigma_y| < \delta$, $|\bar{X}_{\tau(z)} - \mu_x| < \delta$, and $|\bar{Y}_{\tau(z)} - \mu_y| < \delta$ on Ω_δ for $\tau(z) \geq M(\delta)$. Hence, for $z \geq \frac{M(\delta)}{T}$, for large enough T,

$$\sup_{z\in[\frac{M(\delta)}{T},\epsilon]}\left|\frac{1}{D_2(z)}\right| \leq \frac{1}{\sigma_x\sigma_y - \delta} < \infty.$$

Straightforward calculation yields

$$\sup_{z\in[\frac{M(\delta)}{T},\epsilon]}|D_1(z)| \leq C_1(\delta)\sup_{z\in[\frac{M(\delta)}{T},\epsilon]}D_3(z)$$

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

for some constant $C_1(\delta)$, where $D_3(z)$ is the sum of finitely many functions $D_3^i(z)$ with $\sup_{z\in[\frac{M(\delta)}{T},\epsilon]}|D_3^i(z)| \to_d \sup_{z\in[0,\epsilon]}|W_1(z)|$.
We have

$$\sup_{z\in[0,\frac{M(\delta)}{T}]}|W_T(z)| \le \frac{C_2(\delta)}{\sqrt{T}} \to 0$$

for a constant $C_2(\delta)$.
Because $W(0) = 0$ $\mathbb{P}$ almost everywhere, we have

$$\lim_{\epsilon\to 0}\limsup_{T\to\infty}\mathbb{P}\Big(\sup_{z\in[0,\epsilon]}|W_T(z)| \ge \eta\Big) = 0$$

on Ω_δ. Because $\delta > 0$ was arbitrary, the lemma follows. ■

LEMMA A.4. *On* $D([0,1],\mathbb{R})$,

$$B_T(\cdot) := \hat{D}\frac{\tau(\cdot)}{\sqrt{T}}(\hat{\rho}_{\tau(\cdot)} - \rho_T) \to_d B(\cdot),$$

where $B(\cdot)$ *is a one-dimensional Brownian bridge.*

Proof. Define

$$W_T(\cdot) := \hat{D}\frac{\tau(\cdot)}{\sqrt{T}}(\hat{\rho}_{\tau(\cdot)} - \rho_0^*)$$

and

$$B_T(z) = W_T(z) - \frac{\tau(z)}{T}W_T(1) =: h\left(W_T(z), \frac{\tau(z)}{T}\right).$$

Since $\tau(z)/T$ converges to z, the lemma follows with the CMT and the definition of the Brownian bridge. ■

Applying the CMT another time proves Theorem 1.

A.3. Proofs of the Local Power.

Proof of Theorem 2. Transferring the proof of Lemma A.1, we obtain that $U(\cdot)$ converges to $D_1^{1/2}\,W_5(\cdot) + A$ with $A = \left(0\ 0\ 0\ 0\ \int_0^z g(u)du\right)'$ because

$$A_2 = \frac{1}{\sqrt{T}}\sum_{i=1}^{\tau(z)}\left(0\ 0\ 0\ 0\ \tfrac{1}{\sqrt{T}}g(\tfrac{i}{T})\right)'.$$

The fifth component converges as a process to the deterministic function $\int_0^{\cdot} g(u)du$.
Also all other proofs can be transferred, and it holds that

$$\begin{aligned}\frac{\tau(\cdot)}{\sqrt{T}}(\hat{\rho}_{\tau(\cdot)} - \rho_0^*) &\to_d D_3D_2D_1^{1/2}\,W_5(\cdot) + D_3D_2A\\ &\stackrel{L}{=} (D_3D_2D_1D_2'D_3')^{1/2}W_1(\cdot) + D_3D_2A\end{aligned}$$

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press

and

$$(D_3D_2D_1D_2'D_3')^{-1/2}\frac{\tau(\cdot)}{\sqrt{T}}(\hat{\rho}_{\tau(\cdot)}-\rho_0^*)\to_d W_1(\cdot)+(D_3D_2D_1D_2'D_3')^{-1/2}D_3D_2A$$
$$\stackrel{L}{=} W_1(\cdot)+(D_3D_2D_1D_2'D_3')^{-1/2}\cdot\frac{\int_0^{\cdot}g(u)du}{\sigma_x\sigma_y}.$$

The estimator $\hat{D}$ converges in probability to $(D_3D_2D_1D_2'D_3')^{-1/2}$. Thus,

$$\hat{D}\frac{\tau(\cdot)}{\sqrt{T}}(\hat{\rho}_{\tau(\cdot)}-\rho_T)\to_d B(\cdot)+C(\cdot),$$

where

$$C(z)=\frac{(D_3D_2D_1D_2'D_3')^{-1/2}}{\sigma_x\sigma_y}\left(\int_0^z g(u)du-z\int_0^1 g(u)du\right),$$

a deterministic function depending on z. ■

Proof of Corollary 1. Similar to the proof of Theorem 2, we have

$$\sup_{z\in[0,1]}\left|\hat{D}\frac{\tau(z)}{\sqrt{T}}(\hat{\rho}_{\tau(z)}-\hat{\rho}_T)\right|\to_d \sup_{z\in[0,1]}|B(z)+MC(z)|$$
$$=M\sup_{z\in[0,1]}\left|\frac{B(z)}{M}+C(z)\right|,$$

where $C(z)\neq 0$ for at least one z. Hence,

$$M\sup_{z\in[0,1]}\left|\frac{B(z)}{M}+C(z)\right|\geq MC_2$$

for a constant C_2. Thus, the test statistic becomes arbitrarily large, in particular, larger than every quantile of the distribution under H_0. ■

It is necessary that h is not constant because the test statistic would be equal to $\sup_{z\in[0,1]}|B(z)|$ otherwise. Because we integrate Mh from 0 to 1, even late structural changes are detected asymptotically if M is sufficiently large.

A.4. Proof of Theorem 1 under (A1)–(A3) and (A5). Analogously to Section A.3, we transfer the proof of Lemma A.1 with

$$A_2=\frac{1}{\sqrt{T}}\sum_{i=1}^{\tau(z)}\left(a_2\tfrac{1}{\sqrt{T}}g(\tfrac{i}{T})\ a_3\tfrac{1}{\sqrt{T}}g(\tfrac{i}{T})\ 0\ 0\ a_1\tfrac{1}{\sqrt{T}}g(\tfrac{i}{T})\right)'.$$

Straightforward calculation yields that $C(z)$ then equals 0. ■

https://doi.org/10.1017/S0266466611000661 Published online by Cambridge University Press