University of Cologne

Institute of Econometrics and Statistics

Specification Testing in Econometric Models

Inaugural dissertation

to obtain the Doctoral Degree from the

Faculty of Management, Economics and Social Sciences

Tim Kutzker

from

Beeskow

First Referee:	Prof. Dr. Dominik Wied	August 18, 2020
Second Referee:	Prof. Dr. Jörg Breitung	

Acknowledgements

At this point I would like to express my gratitude to the following people - I could not have written this thesis without them:

First of all, I would like to thank Dominik Wied, for the supervision of this work, his constant help, his invaluable comments and advice and, above all, his confidence, which he continually placed in me throughout this time.

I would also like to thank Jörg Breitung for his constructive suggestions and support, which contributed to the success of this work, not only during the research seminar.

Furthermore, I would like to thank my office colleague Florian Stark, with whom I not only initiated a joint project, but also always had an open ear for questions of any kind.

Moreover, I would like to thank Nadja Klein for her engagement and constant support, which particularly contributed to the success of the third project.

In general, I would also like to thank the entire institute for the always friendly, cooperative and helpful atmosphere, all academic and especially non-academic discourses that have had such a lasting impact on the work over the past years.

Finally, I would like to thank my family and friends for their unrestricted support throughout my studies and the understanding conversations of all kinds.

Contents

List of Tables iii

List of Figures iv

1 Introduction and Motivation 1

2 Testing the Correct Specification of a Spatial Dependence Panel Model for Stock Returns 11

2.1 Abstract . . . 11
2.2 Introduction and Summary . . . 11
2.3 A Cross Sectional Correlation Based Specification Test for SAR(m) Panel Models . . . 13
2.3.1 The Model . . . 13
2.3.2 The Specification Test . . . 15
2.3.3 A More Powerful Test . . . 19
2.4 Monte Carlo Simulation . . . 20
2.4.1 GARCH(1,1) . . . 26
2.5 Empirical Analysis . . . 29
2.6 Conclusion . . . 33

3 Testing for Relevant Dependence Change in Financial Data: A CUSUM Copula Approach 35

3.1 Abstract . . . 35
3.2 Introduction . . . 35
3.3 Relevant change and test statistic . . . 38
3.4 Bootstrap and Testing procedure . . . 42
3.5 Quantile- and L^2-Simulations . . . 44
3.5.1 Serially Independently Distributed Data . . . 44
3.5.2 Residual Data . . . 52
3.6 Application . . . 53
3.7 Conclusion . . . 56

4 Specification Testing in Functional Quantile Regression Models 59

4.1 Abstract . . . 59
4.2 Introduction . . . 60

4.3 Quantile regression testing 64
4.3.1 Specification test for quantile dependent regressors . . . 64
4.3.2 More powerful testing procedure using splines 68
4.3.3 Semiparamteric bootstrap procedure 70
4.4 Asymptotics 71
4.4.1 Theoretical properties for quantile dependent regressors 71
4.4.2 Theoretical properties for quantile dependent regressors using constrained polynomial spline regression 73
4.4.3 Validity of the bootstrap procedure 75
4.5 Monte Carlo simulation study 76
4.6 Conditional income disparities between East and West Germany 82
4.7 Conclusion 88

A Testing the Correct Specification of a Spatial Dependence Panel Model for Stock Returns 89
A.1 A Two Step GMM Estimation Procedure for SAR(m) Models . 89
A.2 Proofs 92

B Testing for Relevant Dependence Change in Financial Data: A CUSUM Copula Approach 96
B.1 Assumptions 96
B.2 Derivation and Asymptotic Distribution of the Test Statistic . 97
B.3 Proof of Theorem 1 104
B.4 Covariance Bootstrap 110
B.5 Simulations for the Covariance Bootstrap 112

C Specification Testing in Functional Quantile Regression Models 114
C.1 Proofs 114

Bibliography 118

List of Tables

2.1 Size and Power of S for SAR(3) 23
2.2 Size and Power of S_χ for SAR(3) 24
2.3 Size and Power of S for SAR(4) 27
2.4 Size and Power of S under GARCH model for SAR(3) 28
2.5 Partitioning of Euro Stoxx 50 members into branches and countries. 30

3.1 Size Analysis using Quantile Version 1/4 45
3.2 Size analysis using the L^2-norm 1/2 45
3.3 Power analysis 46
3.4 Size analysis using the L^2-norm 2/2 47
3.5 Size analysis using quantile version 2/4 48
3.6 Size analysis using quantile version 3/4 49
3.7 Size analysis using quantile version 4/4 49
3.8 Pretesting 50
3.9 Power analysis for residual data 51
3.10 Power with break at $\frac{T}{4}$ 51
3.11 Power analysis for varying Δ 52
3.12 Size using quantile version for GARCH-data 53
3.13 Size using quantile version for GARCH-data with breaks 54
3.14 Power analysis varying θ using GARCH-residuals 55
3.15 Empirical values 55
3.16 Empirical break points 56

4.1 Size analysis to the significance level 0.10 and 0.05 79
4.2 Power analysis location shift (LS) 80
4.3 Power analysis location shift (LS) for DGP9 81
4.4 Size and power analysis under a linear null hypothesis 81
4.5 Description of the German labor income data 84
4.6 Decomposition of the West/East income differential 85
4.7 Empirical rejection frequencies of the test statistic $S_n^{CM^*}$ 87

B.1 Size analysis with known marginals 113
B.2 Size analysis with known marginals using the L^2-norm 113

List of Figures

2.1 Rolling Window for $T = 100$ 31
2.2 Rolling Window for $T = 100$ under GARCH(1,1) adjustment . 32
2.3 VaR-Forecasts for $T = 50$ under GARCH(1,1) adjustment . . . 33
2.4 VaR-Forecasts for $T = 50$ 34

3.1 Value of the empirical copula at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)$ 57

4.1 Conditional income quantiles for East/West and entire Germany as functions of *age* 87

CHAPTER 1

Introduction and Motivation

> *"The full usefulness of a large and important group of our economic analyses will come, therefore, only as we succeed in formulating the discussions in quantitative terms. [...] But no amount of statistical information, however complete and exact, can by itself explain economic phenomena. If we are not to get lost in the overwhelming, bewildering mass of statistical data that are now becoming available, we need the guidance and help of a powerful theoretical framework" (Frisch, 1933, p. 2).*

Even if the editor's note from the very first issue of *Econometrica* does not contradict in any way the validity and explanatory power of qualitative research, the words of Ragnar Frisch, co-founder of the Econometric Society in 1930, although they are almost 90 years old, seem to be more relevant than ever, especially in today's times. This is due not only to the fact that we are experiencing a veritable explosion of data[1], but also that benefits can only be generated from data if they are appropriately reduced[2] by means of sound and scientific methods, so that connections, patterns and statements that are based on the data can be reliably concluded. However, as to the question of what constitutes a reliable conclusion in a scientific context, one finds a number of different answers, which moreover are constantly changing over time.

At the time of the founding of the Econometric Society, probably the most

[1] The International Data Corporation predicts that the global datasphere will grow from 33 zettabytes in 2018 to 175 zettabytes by 2025.

[2] Ronald A. Fisher sees reduction as the major task of statisticians.

influential group of philosophers and scientists was the Vienna Circle and its associated school of logical empiricism, which predominated the scientific discourse of the time. Important core points of the Vienna Circle and its logical empiricism are i.a. to reject metaphysical interpretations, to accept only testable statements and to demand a unified science, in which all empirical sciences should be formulated in a physical (logical) language (Keuzenkamp, 2000). So it was only coherent and entirely in the zeitgeist that the Econometric Society's main objective was to "promote studies that aim at a unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking similar to that which has come to dominate in the natural sciences" (Frisch, 1933, p. 1).

Although this paradigm is certainly still compatible with today's objectives of quantitative science, terms, definitions and their interpretations lived through the change of time. In particular the concept of probability, which is crucial for econometrics and statistics and on which the quantifiability of this scientific discipline is founded, underwent colossal value changes in the 20th century. While at the beginning of the last century Bayesian interpretations constituted the academic school of thought, the falsificationism originally developed by Karl R. Popper (*1902 - †1994) and his critical rationalism at least partially replaced the positivist interpretation of probability of logical empiricism (Efron, 2005). In contrast to the Bayesian interpretation of probability as a reasonable expectation that represents a state of knowledge, or as a quantification of a personal conviction (De Finetti, 2017), Popper's main goal was to objectify the initially subjective attributions of probability. While the approach of critical rationalism, which is based on the repeatability of (objective) frequencies, is not uncontroversial in formal logic and philosophy of science[3], to this day it has continually broadened the econometric perspective and created new types of questions. An essential distinction is that falsificationism, in contrast to logical empiricism, proceeds from the unprovable nature of a theory. Thus, in critical rationalism it is impossible to show the validity of an

[3]A widely respected critical examination of the frequentist interpretation of the concept of probability goes back to John M. Keynes (1921) and the well-known example of whether or not to go out with one's umbrella in the situation in which the pressure is high and the clouds are black.

empirical model. Instead, the validity of a model is assumed to be true until its opposite can be shown. For falsificationism, it is therefore all the more crucial that models and their implications be described in such a way that it is always determinable when a model needs to be rejected.

According to Popper, the scientific approach should consist of first generating the model's inherent properties by means of hypothetical-deductive (HD) reasoning[4] and then, where appropriate, falsifying the statements obtained with the help of statistical tests[5]. In the scientific toolbox of critical rationalism, testing and decision-making thus form the basis of scientific work (Granger, King & White, 1995). The fact that Popper's method of first generating HD statements and then testing their validity with statistical tests was not only able to answer new relevant questions, but was also in line with the beginnings of modern statistics by Jerzy Neyman (*1895 - †1980) and Egon Pearson (*1902 - †1994), was probably one reason why the methods of the critical rationalism became so established in statistics and with them the entire test theory (Lakatos, 1978)[6]. In particular, the role of specification testing was considered to be essential in the field of hypothesis testing.

This is not surprising insofar as it was already widely understood that, e.g., in the standard linear regression model, a violation of the orthogonality assumption leads to an estimation bias, whereas heteroscedasticity causes inefficient estimates[7]. If these model violations were not taken into account, this would undermine justifications for associated test decisions. Consequently, the identification of the correct model and the associated statistical verification of the model selection made is formally and logically indispensable, since it represents the reasoning basis for decisions resting upon it. Hence, in statistics and

[4] There are other methods of scientific inference, e.g. Bayesian inductive inference, although historically they have been less important (Keuzenkamp & Magnus, 1995).

[5] The positivistic interpretation of the HD method would attempt to attribute degrees of confirmation to the logical implications of certain models, which is why specification tests received far less attention in the era before Popper.

[6] Although the theoretical foundations of modern test theory can be traced back to Karl Pearson, the final breakthrough seems to have been made possible only by the dominance of critical rationalists.

[7] Trygve Haavelmo, one chief originator of the Cowles Commission received (among other members) the Nobel Prizes in economics. "[Havelmoo] is a short clear demonstration, by means of simple examples, of why least squares yields biased and inconsistent estimators in simultaneous equations models, and how to get consistent estimators in special cases that we now recognize as just identified" (Christ, 1994, p. 32).

econometrics, specification tests are the working tool for the verification of the model selection made.
One of the first specification tests can be traced back to Gregory Chow (*1930) from 1960. Although, strictly speaking, it is a test for a structural break, the Chow-test explicitly decides whether a given dataset fits the null hypothesis model (Chow, 1960)[8]. However, the breakthrough in specification testing came from Jerry A. Hausman (*1946) in 1978 with his article "Specification Tests in Econometrics" which was decisive for further developments on specification tests, since he presented a quite general method[9]. In his approach, it is first assumed that the model is correctly specified under the null hypothesis while the alternative hypothesis is misspecified. Then, two estimators are considered: one estimator θ_1 that is consistent and efficient under the null hypothesis but not consistent under the alternative hypothesis and a second estimator θ_2 that is under both hypotheses consistent but inefficient under the null hypothesis. If the model is correctly specified, then θ_1 and θ_2 should have similar values, otherwise they differ[10].
Many subsequent tests are based on the principle developed by Hausman. Well-known representatives in the at that time comparatively young subdiscipline of specification tests are the information matrix test by Halbert White (*1950 - †2012, 1982), which compares the covariance matrices of the estimators and White's test for functional misspecification (White, 1980). These tests represent only the beginning of a self-advancing, constantly growing and independent subfield of econometrics, which has seen a steadily increasing influence until today.
However, it remains to be seen, whether this subfield will continue to gain additional weight in statistics and econometrics. On the one hand, there are still justified objections to a frequentistic interpretation of the concept of probability[11]. On the other hand, the correct model specification remains of

[8] Although the work of Chow appears somewhat later than that of Ewan S. Page (1955), who is considered the father of CUSUM type tests, it seems more appropriate to include Chow in the range of specification tests, since Page only tests whether a sample of independent observations is drawn from one or two distributions.

[9] Even though Hausman's work considers the articles of De-Min Wu (1973) and James B. Ramsey (1969) specification testing received general attention through Hausman's article.

[10] In the literature, this version of the Hausman test principle is also known as the Wu-Hausman test.

[11] In recent decades, the frequentist interpretation of the concept of probability has developed

central importance in classical hypothesis testing since the model choice with all its assumption constitutes, among other things, the limiting distribution of model-based test statistics and thus, every post-model selection inference. In contrast, other, more recent approaches used to determine the critical values of model-based test decisions, such as bootstrap procedures, are better able to deal with minor misspecifications, thereby at least partially reducing the need for specification tests. Furthermore, bootstrap methods also allow for Bayesian interpretations (Efron, 2005), so the question of correct model selection loses at least some of its importance[12].

At present, it seems to be the case that "basically, there is only one way of doing physics, but there seems to be at least two ways of doing statistics, and they do not give the same answers" (Efron, 2005, p. 1). Even though this quote by Bradely Efron (*1938) is putatively an accurate description of the current state of affairs, he sees the future of statistics in merging Bayesian and frequentistic approaches in order to meet the great challenges of modernity by estimating thousands of parameters[13]. And even if procedures that are robust against minor misspecified models may gain in importance, it seems inconceivable to develop a general procedure that generates consistent estimates independent of the properties of the given data set; particularly if key variables are not taken into account in the model. After all, the consistent estimation of model-based parameters is still one of the greatest challenges in frequentistic models. Specification tests make an explicit contribution to the identification of the omitted variable bias. Thus, it can be assumed that specification tests will continue to play an important role in the validation of scientific knowledge in the future, since they can be seen as a guideline for model selection in a world in which knowledge and therewith the number of different models is growing exponentially.

into propensities, on which Popper was a well-known representative in his later years. Even if this approach has been able to eliminate many weaknesses of the frequentist concept of probability, it has not always been possible to provide a uniform interpretation of the concept of probability. One reason for this is the fact that propensities still adhere to the objectification of the concept of probability and can thus be understood as a further development of frequentism. The conception of propensities was motivated in particular to solve the interpretation of quantum mechanics.

[12]There is a variety of literature including Spokoiny & Zhilova (2014); Corradi & Swanson (2003), among others.

[13]Efron sees signs of the development of such a hybrid science in the techniques of empirical Bayes.

Notwithstanding the above, the question of the correct model selection poses a completely different problem, which was already recognized by Nikolass Tinbergen (*1907 - †1988), a pioneer of econometrics, in 1936 (Keuzenkamp, 2000, p. 142):

> *"Since it is not possible to work with a sequence of two separate analytical stages, 'first, an analysis of theories, and secondly, a statistical testing of those theories.' Modeling turned out to become an Achilles' heel for econometrics."*

The fact that this supposedly highly theoretical approach to econometric theory and specification testing also has relevant practical implications can, however, be observed very specifically in the context of the financial sector. The importance of stock performances as an indicator of economic success has become very well established in recent years. One of the difficulties in modeling stocks is the complicated correlation over cross sections. Factor models are common methods to model stock returns. In the case of one factor models, however, it has been shown that spatial autoregressive models (SAR) without exogenous covariates (Arnold, Stahlberg & Wied, 2013; Wied, 2013) are more suitable for predicting portfolio variances and Value-at-Risk (VaR) forecasts since they possess lower prediction errors (Schmitt, Schäfer, Wied & Guhr, 2016). In addition, SAR models can also be used to explain the propagation of country-specific shocks to other countries by looking at various links such as economic and monetary interdependence between countries. Again, it is essential for the quality of an estimate and for any post-estimation inference that the model prerequisites are met. In the second chapter of the present thesis, which is based on the work by Kutzker & Wied (2019), two testing procedures are proposed, which statistically verify whether an m-dimensional SAR (SAR(m)) model can be applied to a given data set and their limiting distributions are derived. The basic idea of the suggested specification tests stems from the model assumption that a SAR(m) model captures all cross sectional dependence and that the remaining covariance matrix of the errors is diagonal. Thus, if the square sum of the secondary diagonal elements of the residual matrix deviates too far from zero, it can be assumed that a SAR(m) is inappropriate. The empirical application to the Euro Stoxx 50 returns shows that the proposed specification tests can also be used as a backtest to spot inaccurate VaR forecasts. Furthermore, the empirical study revealed that, particularly in bear markets, simple three-dimensional SAR models cannot adequately capture the increased cross-dependencies, and that a disproportionately high number of VaR forecast violations can be observed, which was apparent around the time of the dot-com bubble, the time around the Lehman Brothers bankruptcy and the Euro crisis. This suggests that in high volatility periods a SAR(m) model should be extended by a factor structure or by further exogenous variables, so that the increased cross-dependencies can be better captured.

Furthermore, it is specifically noticeable that high market volatility and cross-sectional diversification occur on the stock market in times of financial crises. The question of the extent of a crisis, however, is examined in more detail in the third chapter of this thesis and corresponds to the paper "Testing for relevant dependence change in financial data: a CUSUM copula approach" (Kutzker, Stark & Wied, 2019) and was published in *Empirical Economics*. Therefore, this chapter focuses on providing a non-parametric test in order to detect relevant breaks in copula functions, thus following the tradition of structural break tests as a special case of specification tests in the sense of Chow (1960). From a portfolio manager's perspective, whose objective is to minimize the risk of the losses, these types of tests are particularly interesting since not every crisis requires a rescheduling of the portfolio, not least since transaction costs arise. Here, relevant breakpoint tests provide the possibility to compare crises and their extent. The non-parametric test approach considers the empirical copula function and assumes that the data is generated by two unequal copulas whose distance is less than or equal to a given fixed positive quantity $\Delta \in \mathbb{R}^+$. The concept of distance used here allows the functional evaluation of the copulas under the L^2-norm as well as at a fixed value. The test procedure consists of two steps: First, the breakpoint is estimated using a conventional CUSUM approach. Secondly, the test statistics, which can be understood as a transformation of the empirical copula CUSUM type processes, can be computed. Due to the non-pivotal nature of the limiting distribution that is determined by transformations of Gaussian processes and the complicated covariance structure, a bootstrap procedure is proposed, which allows to determine critical values. Interpreting Δ as the smallest admissible copula difference for which the relevant change hypothesis cannot be rejected, led to the statistical result that in the empirical application the Euro crisis was more substantial than the beginning of the financial crisis.

Gaussian processes also play a distinctive role as limiting distributions in the fourth chapter, which matches the working paper "Specification Testing in Functional Quantile Regression Models with an Application to Income Differences in Germany" by Kutzker, Klein & Wied (2020) and is devoted to a specification test for functional quantile regression models. Quantile regression models have gained increasing influence since the seminal work by Koenker & Bassett Jr (1978). Compared to simple OLS models, which are very well studied and understood, quantile regressions models have the advantage that they are less susceptible to outliers. Since they refer to every quantile and not solely to the conditional expected value, they also draw a more complete picture of the underlying problem. Due to their more complicated estimation compared to OLS, only linear quantile models have been considered for a long time. In the recent past, however, it can be observed that nonlinear and nonparametric quantile estimators have been increasingly proposed. Cardot, Crambes & Sarda (2005) even provided an estimator in which the covariates are quantile-dependent functions, which allows for very flexible quantile regressions. However, little attention has been devoted to specification testing procedures with quantile dependent covariates. In the fourth chapter, a novel consistent specification test for quantile regression models is proposed where the covariates X can have quantile-specific functional forms. The basic idea of this specification test is to compare an unrestricted estimate of the joint distribution function of the endogenous random variable Y and the exogenous random vector X with a restricted estimate that imposes the structure implied by the null hypothesis model. Based on a Cramèr-von Mises type measure of distances, the restricted estimate of the joint distribution is compared with the unrestricted one. The limiting distribution of the test statistic is non-pivotal and depends on Gaussian processes in a complex fashion. In order to obtain critical values, the validity of the suggested bootstrap method is shown. The application of the novel test procedure applied to data from the German socio-economic panel (SOEP) could statistically confirm that there are still differences in the income distributions between East and West Germany over the period considered.

Overall the thesis comprises three self-contained essays on specification testing in a frequentistic framework. The mathematical proofs are found in the corresponding appendices at the end of the thesis. All essays are joint works with Dominik Wied. In the third chapter there was an additional collaboration with Florian Stark, and in the fourth chapter with Nadja Klein. In these collaborations, my task in the first project was particularly to derive the details of the mathematical proofs and to implement the model in R both for the realization of the simulation study and the empirical application. In the second project, I contributed, besides others, to the mathematical proofs and the simulation study in Matlab and I performed the empirical analysis. In the last project I wrote down the manuscript text to a large extent, I provided the exact execution of the theoretical details and the complete implementation in R both for the simulation study and for the empirical application.

Beyond that, I also contributed to the joint working paper with Maximilian Schreiter named "The Optimal Capital Structure under Risks of Illiquidity and Over-indebtedness in a Double Barrier Option Framework" (2020) during my Ph.D. studies. Generally, dynamic capital structure models are based on single triggers, which determining bankruptcy, mainly over-indebtedness or illiquidity. The latter one tends to underestimate optimal capital structures by ignoring capital providers' flexibility to inject fresh money. The former one leans towards overestimation as it neglects agency conflicts between equity investors and debt holders while implying infinitely "deep pockets" of equity investors. The approach in the working paper incorporates both constraints, over-indebtedness and illiquidity, examining corporate debt value and optimal capital structure in a double barrier world with knock-in and knock-out options, where a closed form solution for all value components of a levered firm is provided. By testing the model for firms publicly listed in the US, evidence is gained that incorporating both triggers allows for capital structure estimates that are in accordance with empirical findings.

In addition, I had the opportunity to develop an approach for random forest algorithms, which aims to soften the uniform distribution assumption. This method proves to be particularly advantageous in the segmentation of data, as it allows individual variables to be weighted differently. An examination of the project in the master's thesis "Unsupervised Learning with Non-Uniform Random Forests" by Arkadiy Davidyan (2020) illustrates the validity and applicability of the method.

CHAPTER 2

Testing the Correct Specification of a Spatial Dependence Panel Model for Stock Returns

2.1 Abstract

This paper provides specification tests for the m-dimensional spatial autoregressive (SAR) panel model by deriving the limiting distribution of the specification test statistics and examines size and power in finite sample simulations. In the empirical application we analyzed the Euro Stoxx 50 returns. Regarding this, a 3-dimensional SAR panel model incorporating global dependencies, dependencies inside industrial branches and local dependencies is assumed. The investigation shows the tests' ability to detect inaccurate Value-at-Risk forecasts.

2.2 Introduction and Summary

In recent years the literature in economics and finance has found some interest in the connection between spatial dependence and stock returns. For example, Asgharian et al. (2013) use techniques from spatial econometrics in order to investigate in which way stock market co-movements are determined by countries' economic and geographical relations. One result shows that trade is the most important factor. Tam (2014) analyzes equity market linkages in East Asia with the result, among others, that Japan is a dominant driver. Selan & Kalatzis (2017) analyze peer effects in Brazil and find a positive spatial dependence between stock returns from peer companies, but a negative feedback effect from fundamental characteristics. A seminal methodological contribution is given by Blasques et al. (2016) who extend the spatial Durbin model by a time-varying spatial dependence parameter.

Furthermore, Arnold et al. (2013) propose a spatial autoregressive (SAR) panel model for stock returns in order to capture local dependencies and dependencies within industrial branches. Wied (2013) considers structural breaks in these models and Schmitt et al. (2016) combine the approach with local normalization techniques. Gong & Weng (2016) use the model for value at risk forecasts in the Chinese stock market. Catania & Billé (2017) generalize the SAR model with autoregressive and heteroscedastic disturbances by including methods from score-driven models. Moreover, Zhang et al. (2018) propose a dynamic spatial panel with generalized autoregressive conditional heteroscedastic model (DSP-GJR-GARCH). Lu (2017) considers a spatial panel data model, that models three effects jointly. Various empirical analyses in the aforementioned papers show that the SAR panel model is generally suitable for Value-at-Risk (VaR) forecasts and outperforms, e.g. the one-factor model.

One aspect which is often missing in recent literature is the question how good the model fits the data. In general, people tend to look at Moran's I (Moran, 1950; Li et al., 2007) to analyze if there is spatial dependence in a given data set. However, this measure is not connected to a specific model. One could apply it to somehow obtained model residuals, but even then, the question would remain in which way we can use this for a test. Born & Breitung (2011) and Su & Qu (2017) propose specification tests for SAR models, but they do not consider a panel context. Kelejian & Piras (2016) propose a J-test procedure for testing a null model against non-nested alternatives for a fixed effects spatial panel data framework. A crucial prerequisite of this test is to formulate what they call G alternative models under H_1.

In this paper, we revisit the SAR panel model from Arnold et al. (2013) and propose two methods on how to check the model fit. The basic idea stems from the model assumption that spatial weighting matrices capture all spatial dependence and that the remaining error terms are spatially uncorrelated. Therefore, we consider the model residuals such that the tests keep the null hypothesis of model fit if the covariance matrix of the residuals is basically diagonal, i.e. its off-diagonal elements are close to zero. We derive the asymp-

totic distribution of our test statistics and show in simulations, that the tests have reasonable power properties against sparse error term covariance matrices. An empirical application on stock data shows that the tests can potentially also be used as backtests for Value-at-Risk forecasts.
This paper is organized as follows: Section 2.3 describes the classical spatial autoregressive model, discusses the assumptions for a GMM estimation procedure and derives the specification tests. Section 2.4 provides an extensive Monte Carlo Simulation and Section 2.5 an empirical application. Finally, Section 2.6 concludes.

2.3 A Cross Sectional Correlation Based Specification Test for SAR(m) Panel Models

In this section, we introduce the general SAR(m), $m \in \mathbb{N}$ panel model and discuss briefly the slightly modified assumptions for the two step GMM estimator given in Arnold et al. (2013) which turn out to also hold for the m-dimensional case.

2.3.1 The Model

The SAR(m) panel model assumes that the dependent variable is correlated in the cross-sectional dimension n and that the spatial dependence can be separated into m different parts. The number m[1] and the specific form of the spatial matrices depend on the practitioner[2]. Thus, the spatial matrices $W_i,\ i = 1, ..., m$ are pre-specified and fixed. In what follows, let $\boldsymbol{y}_t$ and $\boldsymbol{\varepsilon}_t$ be n-dimensional random vectors for $t = 1, \dots, T$. The m-dimensional SAR panel model without any explanatory variables is given by

$$\boldsymbol{y}_t = \sum_{i=1}^{m} \rho_i W_i \, \boldsymbol{y}_t + \boldsymbol{\varepsilon}_t,\ t = 1, \dots, T \tag{2.1}$$

[1] In the application later on, we will introduce three different spatial matrices which are assumed to capture the structure of daily stock returns. The first part covers a general dependence which affects all subjects equally. The second part captures dependencies among industrial branches and national effects are included with the help of the third dependency structure.

[2] An overview of commonly used matrices is given in Elhorst et al. (2012).

where $\rho_i \in \mathbb{R}$ for $i = 1, ..., m$. For asymptotic results, n n is fixed and T is sent to infinity. To derive limit theorems we impose the following assumptions:

Assumption 1.

1. *The sequence of random vectors $\{\boldsymbol{y}_t\}_{t\in\mathbb{N}}$ has zero mean, is stationary and ergodic.*

2. *For $i \in \{1, ..., m\}$, $r = 1, \ldots, n$, $s = 1, \ldots, n$, $W_{i,rs} \geq 0$, $W_{i,rr} = 0$.*

3. *For $i \in \{1, ..., m\}$ and $r = 1, \ldots, n$, $\sum_{s=1}^{n} W_{i,rs} = 1$.*

4. *The parameter space S is defined as $S := \{\boldsymbol{\rho} \in \mathbb{R}^m : ||\boldsymbol{\rho}||_1 < 1\}$ where $||\cdot||_1$ defines the L_1-norm.*

5. *For $t \in \mathbb{Z}$, $\mathsf{Cov}\,(\boldsymbol{\varepsilon}_t) = diag\{\sigma_1^2, \ldots, \sigma_n^2\} =: \Sigma \in \mathbb{R}^n$.*

6. *Each element of the vector $\left(\frac{1}{\sqrt{T}} \sum\limits_{t=1}^{T} \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'\right)_{i<j}$ meets the assumption of a central limit theorem and the corresponding long-term covariances*

$$\sum_{s,t\in\mathbb{N}} Cov[\boldsymbol{\varepsilon}_{i1}\boldsymbol{\varepsilon}_{jt}, \boldsymbol{\varepsilon}_{ks}\boldsymbol{\varepsilon}_{ls}]$$

are finite for every $i < j$ and $k < l$.

In the context of daily stock returns, the zero mean and stationarity Assumption 1.1 is plausible (see Aue et al., 2009). Assumption 1.2 excludes "self influence" since the elements on the leading diagonal are zero and postulates that all elements are non-negative, which is usually the case in empirical applications. Assumption 1.3 ensures that the matrices are bounded and standardized. For the GMM estimator based on Arnold et al. (2013) we assume row-standardized weighting matrices. Depending on the underlying GM-estimation technique this assumption could be relaxed (Kelejian & Prucha, 2010; Breitung & Wigger, 2017). Assumption 1.4 restricts the parameter space such that the sum of the absolute values of the elements of $\boldsymbol{\rho} \in \mathbb{R}^m$ is smaller than 1. Even though the assumption could be slightly generalized (Elhorst et al., 2012) we follow the notation of Arnold et al. (2013) as it guarantees that the matrix $(I_n - \sum_{i=1}^{m} \rho_i W_i)$ is non-singular[3]. Hence, Assumption 1.1-1.4 ensure the model 2.1 to be well defined.

[3]The matrix $(I_n - \sum_{i=1}^{m} \rho_i W_i)$ is strictly diagonally dominant.

The crucial assumption, on which we will base our specification test, is that the covariance matrix of the error terms $\boldsymbol{\varepsilon}_t$ is diagonal. Consequently, all cross-sectional dependence is captured by the spatial terms, which corresponds to Assumption 1.5, although, heteroscedasticity is not excluded. Assumption 1.6. guarantees that the limiting distribution of our suggested test statistic is not degenerated, i.e. the dependence structure of the error vector $\boldsymbol{\varepsilon}_t$ meets certain regularity conditions, such that the serial dependence structure is bounded.
For the estimation, a two step GMM procedure is considered. First, we estimate the correlation parameters by the method of moments along the lines of Kelejian & Prucha (1999) or Kapoor et al. (2007). This step does not depend on the parameters of variance. Secondly, we estimate the variance parameters. Under some regularity assumptions the GMM estimator $\hat{\boldsymbol{\rho}}$ is consistent and as asymptotically normal. While this is worked out in Arnold et al. (2013) for the special case of $m = 3$, a detailed derivation for the GMM estimator in the general case is presented in the Appendix A.1.

2.3.2 The Specification Test

We outline the test for the case of Assumption 1, noting that simulation results in section 2.4.1 indicate that the test also works if we replace the error terms by GARCH residuals. So subsequently, the word data set can be regarded either as the original or the GARCH adjusted data.
Following the discussion given in the previous subsection, what remains is to check whether Assumptions 1.5 holds. Even if the course of action seems technical, the idea behind the test statistic is straightforward: we do not reject the null hypothesis if the covariance matrix of the errors is basically a diagonal matrix, i.e. its off-diagonal elements deviate not too far from zero. Let $\hat{H} \in \mathbb{R}^{n \times n}$ denote the empirical covariance matrix of the residuals times the square root of the time horizon, i.e. $\hat{H} := \sqrt{T}\hat{\text{Cov}}[\hat{\boldsymbol{\varepsilon}_t}]$ and $\hat{H}_{ij}$ with $i, j \in \{1, 2, ..., n\}$ its elements. Let σ_{ij}^2 denote the (i, j)-th element of the theoretical counterpart Σ, i.e. the error covariance matrix. Since $\hat{H}$ and Σ are symmetric, it is sufficient to consider only the elements of the upper triangle of the matrix

Σ. Hence, the null hypothesis is given by

$$H_0 : \sigma_{ij}^2 = 0 \text{ for all } i < j \quad \text{vs.} \quad H_1 : \exists s, t \text{ with } s < t : \sigma_{st}^2 \neq 0. \tag{2.2}$$

We opt to use χ^2-type tests for this testing problem. Instead of considering each element or the maximum of the absolute value of all off-diagonals, we take the sum of each element squared into account. Thus, the naive test statistic is given by

$$S := \sum_{i<j,i,j=1,\ldots,n} (\hat{H}_{ij})^2. \tag{2.3}$$

The following theorem identifies the limiting distribution of the empirical covariance matrix times $\sqrt{T}$.

Theorem 2.3.1. *Under the null hypothesis $H_0 : \sigma_{ij}^2 = 0$ for all $i < j$, the assumptions of Theorem A.1.3, the following holds for $1 \leq i, j \leq n$*

$$\underset{T\to\infty}{dlim} \sqrt{T} \ \hat{Cov}[\hat{\boldsymbol{\varepsilon}}_t] = A + B + B^{'} \in \mathbb{R}^{n\times n} \tag{2.4}$$

with $(A)_{ii} = \lim\limits_{T\to\infty} \sqrt{T} \sum\limits_{t=1}^{T} \sigma_{it}^2 = \infty$ and the components of A are jointly normally distributed for $i \neq j$ with $(A)_{ij} \sim N(0, \lim\limits_{T\to\infty} \mathsf{Var}\ [\frac{1}{\sqrt{T}} \sum_{t=1}^{\infty} \boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{jt}])$ and $Cov((A)_{ij}, (A)_{kl}) = 0$ for $i \neq j$ and $k \neq l$ with $(i, j) \neq (k, l)$. Moreover, $B \overset{d}{=} (\sum\limits_{g=1}^{m} X_g W_g)(I_n - \sum\limits_{g=1}^{m} \rho_g W_g)^{-1}\Sigma$, where

$$\boldsymbol{X} := (X_1, \ldots, X_m) \sim N(0, \boldsymbol{d}^{-1} S_W (\boldsymbol{d}^{-1})^{'}) \in \mathbb{R}^{1\times m}$$

with $S_W = \sum_{t=-\infty}^{\infty} \mathsf{E}[f(\boldsymbol{y}_1, \boldsymbol{\rho}) f(\boldsymbol{y}_t, \boldsymbol{\rho})']$ for $f(\boldsymbol{y}_t, \boldsymbol{\rho}) = (\boldsymbol{\varepsilon}_t' W_1 \boldsymbol{\varepsilon}_t, \cdots, \boldsymbol{\varepsilon}_t' W_m \boldsymbol{\varepsilon}_t)'$ and $\boldsymbol{d}$ defined in Assumption (4).

Here and in the following $\underset{T\to\infty}{\text{dlim}}$ denotes limit in distribution and $\overset{d}{=}$ equality in distribution. Three remarks about Theorem 4.4.1 are in order. First, the leading elements of matrix A diverge to infinity. However, the tests considers only the off-diagonal elements $(i \neq j, i, j = 1, ..., n)$, which are finite by Assumption 1.6. This in turn ensures, that the test is well defined. Second, since $\left(I_n - \sum_{g=1}^{m} \rho_g W_g\right)$ is strictly diagonally dominant, the inverse exists. Third, we note that the matrices B and its transposed appear in the limit. This is due to the effect of estimating $\boldsymbol{\rho}$ instead of using the unknown population quantity. The analysis of such a residual effect (see Demetrescu & Wied, 2019) is somewhat complicated, since the additional terms need different standardizing factors in the proof[4]. However, all terms in the limiting distribution are based on the same error terms, thus, the convergence is jointly and the limiting distribution in (2.4) is multivariate normal. If we additionally assume serially independence in the error vector, the variance of the elements in the limiting matrix A simplifies to a product, shown in the following remark.

Remark 2.3.2. *Suppose the assumptions of Theorem 4.4.1 hold.*
If $\{\boldsymbol{\varepsilon}_t\}_{t\in\{1,...,T\}}$ *is serially independent, then*

$$(A)_{ij} \sim N(0, \sigma_i^2 \sigma_j^2) \quad \textit{for} \quad i \neq j. \tag{2.5}$$

In accordance with our test statistic (2.3), we can reformulate the test in vectorial notation, i.e.

$$S = \hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\alpha}}, \tag{2.6}$$

where $\hat{\boldsymbol{\alpha}}$ represents the vector of the upper triangle of the empirical covariance matrix of the residuals times $\sqrt{T}$. Since the empirical covariance matrix consists of n^2 elements, the upper triangle matrix vector (i.e. stacking every element above the leading diagonal, but excluding elements from the leading diagonal) consists of $n(n-1)/2$ elements and has the following form:

$$\begin{aligned}
\hat{\boldsymbol{\alpha}} \; &:= \; \underset{T\to\infty}{\text{dlim}} \left(\sqrt{T}\hat{\text{Cov}}[\hat{\boldsymbol{\varepsilon}}_t]\right)_{i<j,\ i,j=1,\ldots,n} \\
&= \; \underset{T\to\infty}{\text{dlim}} \left(\frac{1}{\sqrt{T}} \sum \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t'\right)_{i<j,\ i,j=1,\ldots,n} \; = \; \underset{T\to\infty}{\text{dlim}} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{\boldsymbol{d}}_t \in \mathbb{R}^{\frac{n(n-1)}{2}}
\end{aligned}$$

$$\text{with } \hat{\boldsymbol{d}}_t := \left(\hat{\varepsilon}_{1t}\hat{\varepsilon}_{2t}, ..., \hat{\varepsilon}_{1t}\hat{\varepsilon}_{nt}, \hat{\varepsilon}_{2t}\hat{\varepsilon}_{3t}, ..., \hat{\varepsilon}_{2t}\hat{\varepsilon}_{nt}, ..., \hat{\varepsilon}_{(n-1)t}\hat{\varepsilon}_{nt}\right)'.$$

[4] For a detailed analysis of the convergence rate we refer to Lemma A.2.1 in the corresponding Appendix.

By means of Slutsky's theorem we define the theoretical counterpart

$$\begin{aligned}
\boldsymbol{\alpha} \ :&= \ (A)_{i<j,\ i,j=1,\dots,n} \\
&= \ \underset{T\to\infty}{\mathrm{dlim}} \left(\frac{1}{\sqrt{T}} \sum \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t{}'\right)_{i<j,\ i,j=1,\dots,n} \ = \ \underset{T\to\infty}{\mathrm{dlim}} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{d}_t \in \mathbb{R}^{\frac{n(n-1)}{2}} \\
\text{with } \boldsymbol{d}_t \ :&= \Big(\varepsilon_{1t}\varepsilon_{2t}, \dots, \varepsilon_{1t}\varepsilon_{nt}, \varepsilon_{2t}\varepsilon_{3t}, \dots, \varepsilon_{2t}\varepsilon_{nt}, \dots, \varepsilon_{(n-1)t}\varepsilon_{nt}\Big)'
\end{aligned}$$

which stacks the upper triangular matrix of the covariance matrix of the errors times $\sqrt{T}$ in a vector. Analogously, $\boldsymbol{\beta}$ defines the vector of the stacked upper triangular matrix of B and $\boldsymbol{\beta}^*$ of B', respectively, i.e. for
$Z_W := \underset{T\to\infty}{\mathrm{dlim}} \sum\limits_{g=1}^{m} \sqrt{T}(\rho_g - \hat{\rho}_g)W_g$ we define

$$\begin{aligned}
\boldsymbol{\beta} := (B)_{i<j,i,j=1,\dots,n} \ &= \left(Z_W(I_n - \sum_{g=1}^{m} \rho_g W_g)^{-1}\Sigma\right)_{i<j,\ i,j=1,\dots,n} \in \mathbb{R}^{\frac{n(n-1)}{2}}, \\
\boldsymbol{\beta}^* := (B')_{i<j,i,j=1,\dots,n} &= \left(\Sigma'(I_n - \sum_{g=1}^{m} \rho_g W_g')^{-1} Z_W'\right)_{i<j,\ i,j=1,\dots,n} \in \mathbb{R}^{\frac{n(n-1)}{2}}.
\end{aligned}$$

The vectors $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$ are well defined, since B is not necessarily symmetric.

Lemma 2.3.3. *$\boldsymbol{\beta}$ represents the vector of the upper triangle and $\boldsymbol{\beta}^*$ the vector of the lower triangle of the matrix $Z_W(I_n - \sum\limits_{g=1}^{m} \rho_g W_g)^{-1}\Sigma$, i.e. for $i,j \in \{1,\dots,n\}$*

$$\boldsymbol{\beta}^* \ = \ \left(Z_W(I_n - \sum_{g=1}^{m} \rho_g W_g)^{-1}\Sigma\right)_{i>j,\ i,j=1,\dots,n} \in \mathbb{R}^{\frac{n(n-1)}{2}}. \tag{2.7}$$

The next Lemma provides the limit distribution of our test statistic S (2.6).

Lemma 2.3.4. *Suppose the assumptions of Theorem 4.4.1 hold. Then the test statistic S* (2.3) *is asymptotically distributed as*

$$S \ = \ \hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\alpha}} \xrightarrow[T\to\infty]{d} (\boldsymbol{\alpha} + \boldsymbol{\beta} + \boldsymbol{\beta}^*)'(\boldsymbol{\alpha} + \boldsymbol{\beta} + \boldsymbol{\beta}^*),$$

where the covariance matrix for $\boldsymbol{\alpha}$ is given by

$$Cov[\boldsymbol{\alpha}] = \begin{pmatrix} \lim\limits_{T\to\infty} \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{1t}\boldsymbol{\varepsilon}_{2t}] & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lim\limits_{T\to\infty} \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{(n-1)t}\boldsymbol{\varepsilon}_{nt}] \end{pmatrix}.$$

Consequently, the critical value for the test statistic S (2.3) can be derived by drawing independently from the limiting distribution given in Lemma C.1.4 and computing the corresponding quantile[5]. The test takes care of the size demands and has good power properties as shown in Section 2.4. The next subsection presents another specification test, which has greater size and consequently better power properties.

2.3.3 A More Powerful Test

In Theorem 4.4.1 we have shown, that the elements of the limiting distribution follow a multivariate normal distribution. Thus, if we standardize the test statistic S (2.6) by its covariance matrix, we receive a new test statistic S_χ^* which is χ^2-distributed, i.e.

$$S_\chi^* \;:=\; \hat{\boldsymbol{\alpha}}'(\mathrm{Cov}[\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\beta}^*])^{-1}\hat{\boldsymbol{\alpha}} \sim \chi^2_{\frac{n(n-1)}{2}}. \tag{2.8}$$

The terms $\boldsymbol{\beta}$ and $\boldsymbol{\beta}^*$ can be regarded as additional noise which comes from the estimation procedure. This additional noise can be extracted by decomposing the covariance matrix given in (2.8) into two parts. Thus, we have

$$\mathrm{Cov}[\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\beta}^*] \;=\; \mathrm{Cov}[\boldsymbol{\alpha}]+\Psi \tag{2.9}$$

with $\Psi \;:=\; \mathrm{Cov}[\boldsymbol{\beta}+\boldsymbol{\beta}^*]+\mathrm{Cov}[\boldsymbol{\alpha},\boldsymbol{\beta}+\boldsymbol{\beta}^*]+\mathrm{Cov}[\boldsymbol{\alpha},\boldsymbol{\beta}+\boldsymbol{\beta}^*]'$. The first part $\mathrm{Cov}[\boldsymbol{\alpha}]$ covers the underlying variance structure while the second part Ψ can be considered as a noise term[6].

[5] The complex dependence structure of α and β can be simulated with the help of the Taylor series approximation.

[6] If we additionally assume serial independence, the covariance matrix of $\boldsymbol{\alpha}$ can easily be implemented, since only the variances need to be estimated, cf. Lemma A.2.3. Otherwise, the covariance matrix of $\boldsymbol{\alpha}$ is given in Lemma C.1.4.

If either $||(\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}\Psi|| < 1$ or $||\Psi(\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}|| < 1$ hold[7], we can estimate the inverse of covariance matrix (2.9) with the help of the Taylor series approximation and the telescoping sum[8]. It yields

$$\begin{aligned}(\mathrm{Cov}[\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\beta}^*])^{-1} &= (\mathrm{Cov}[\boldsymbol{\alpha}])^{-1} - (\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}\Psi(\mathrm{Cov}[\boldsymbol{\alpha}])^{-1} \\ &\quad + (\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}\Psi(\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}\Psi(\mathrm{Cov}[\boldsymbol{\alpha}])^{-1} - ... \\ &\leq (\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}.\end{aligned}$$

Thus, $(\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}$ is an upper bound for the inverse of the covariance matrix (2.9). Hence,

$$S_\chi := \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \hat{\boldsymbol{d}}_t{}'(\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}\sum_{t=1}^{T}\hat{\boldsymbol{d}}_t \tag{2.10}$$

provides a more powerful test, since $S_\chi \geq S_\chi^* \sim \chi^2_{\frac{n(n-1)}{2}}$. In order to study the behavior of S and S_χ in finite samples we perform an extensive Monte Carlo Simulation which can be found in the next section.

2.4 Monte Carlo Simulation

The Monte Carlo (MC) Simulation consists of three major simulations. While the first two simulations assume serial independence, the third simulation examines the behavior of the test in the case of GARCH(1,1) driven errors. The first less comprehensive simulation depicts a 3-dimensional SAR model

$$\boldsymbol{y}_t = \rho_1 W_1 \boldsymbol{y}_t + \rho_2 W_2 \boldsymbol{y}_t + \rho_3 W_3 \boldsymbol{y}_t + \boldsymbol{\varepsilon}_t,\ t = 1,\dots,T$$

with $(W_1)_{ij} = \frac{1}{n-1}$ for all $i \neq j$ and $(W_1)_{ii} = 0$. The spatial matrices W_2 and W_3 are defined as

[7] In our Monte Carlo simulation we observed that this is usually the case whenever the variance of $\boldsymbol{\varepsilon}_{it}$ is greater than 1 for all $i = 1, ..., n$.

[8] The sum and product of two symmetric positive semidefinite (psd) matrices is still psd.

$$(W_2)_{ij} = \begin{cases} 1, \text{ if } j \text{ even and } i \neq j \\ 1, \text{ if } j-1 = i \\ 0, \text{ otherwise} \end{cases} \qquad (W_3)_{ij} = \begin{cases} 1/(n/2-1), \text{ if } i,j \lesseqgtr \frac{n}{2} \\ 0, \text{ otherwise,} \end{cases}$$

where additionally the matrix W_2 is row standardized by its row sum $\sum_j (W_2)_{ij}$. The expression $i,j \lesseqgtr \frac{n}{2}$ in the definition of W_3 indicates that both i and j are either smaller or equal or greater than $\frac{n}{2}$. In terms of interpretation the matrix W_1 can be regarded as a weighting matrix, where each firm has the same weight with respect to a portfolio. Thus, the matrix W_1 captures a general effect, e.g. global crisis, market performance in the past etc.[9]. The spatial matrix W_2 can be considered as industry affiliation. W_3 may be regarded as the dichotomous component of the market which divides the market into two different fields (e.g. the beneficiaries of a given change, e.g. fiscal reform, aid payments, etc) and those who are not affected.

In the first part, the vector of observation $\boldsymbol{y}_t$ is generated by a multivariate normal error vector $\boldsymbol{\varepsilon}_t$ with zero mean and covariance matrix $\Sigma := \sigma^2 I_n$, where I_n represents the n-dimensional identity matrix. The parameter of spatial dependence is given by $\boldsymbol{\rho} = (0.45, 0.3, 0.15)$ and the homoscedastic variance equals $\sigma^2 = 2$. For calculating the power of our tests we use the following misspecification: If we consider a market with n participants, then there are $n(n-1)/2$ possible pairs (e.g. participants that are correlated with each other). The parameter ζ describes the portion of how many pairs we wish to consider[10], the parameter κ^2 describes their correlation. E.g. if we consider a market that consists of $n = 20$ actors, then there are $n(n-1)/2 = 190$ different pairs. If $\zeta = 0.1$ and $\kappa = 0.2$, we presume that there are 19 pairs that have a correlation coefficient that is equal to 0.04. No further assumptions are made about the structure of the correlation. It is possible, for instance, that the actor number of 20 has the correlation 0.04 with every other participant, i.e. Σ is a diagonal matrix with 0 in the off diagonals. Only

[9]Even if W_1 is equally weighted, ρ_1 cannot be considered as a fixed affect which affects market participant equally, since fixed affects are time independent. SAR models try to capture this time dependence structure with fixed weighting matrices.

[10]In case that $\zeta \cdot n(n-1)/2$ is odd we round down.

the last column and row of Σ is non-zero. However, in general, the correlation structure is completely random[11]. To determine the size and the power of the first test (4.4.1) we draw $B = 400$ times from the asymptotic limit distribution given in Lemma (C.1.4). The dependence between α and β is modeled by the Taylor series approximation. The overall number of MC repetitions is equal to 701. We begin by studying the size of the first test for $n = 20, 50$ and $T = 50, 100, 200, 500$. Results are presented in Table (2.1). Collectively, the test has good size. Similar properties are derived for the power analysis of the test. Whenever the ratio of T over n is small and the dependence structure in the error term is more or less negligible (cf. $\kappa = \zeta = 0.05$) the power of the test is low. However, if there are sufficient observations (i.e. $\frac{T}{n} > 10$) and if the dependence structure in the data set is not negligible ($\kappa, \zeta \geq 0.1$), then the test provides good power properties. All in all we observe an increasing power whenever the dependence structure (κ or ζ) or the number of observations (n or T) increases.

Similar results are obtained for the second test S_χ (2.3) which can be found in Table (2.2)[12]. In small sample studies S_χ performs worse than test S in terms of size power. This is due to the fact that the we used the empirical approximation for the inverse covariance matrix that is employed in S_χ, which is biased in small samples. Consequently, as T tends to infinity the size of the test S_χ converges clearly to the desired nominal level of 5% and the power increases as the level of misspecification rises.

However, additional simulations show that the tests' power decreases in the case of too large ζ, i.e. in the case of a highly non-sparse covariance matrix. Here, the population moment conditions from (A.1) are severely violated so that the model is misspecified and the behavior of the model estimators $\hat{\rho}$ is unclear (Fleming, 2004).

[11] This procedure of misspecfication ensures that the moment conditions (A.1) are violated, thus, the GMM estimator is biased (Hansen, 1982).

[12] The second test is applicable since we observed in every study and simulation we conducted that either $||(\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}\Psi|| < 1$ or $||\Psi(\mathrm{Cov}[\boldsymbol{\alpha}])^{-1}|| < 1$ hold

Table 2.1: Size and Power of S for SAR(3)

n = 20	$\zeta = 0.05$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.031	0.039	0.060	0.108	0.332
$T = 100$	0.039	0.056	0.956	0.213	0.742
$T = 200$	0.042	0.069	0.220	0.532	0.973
$T = 500$	0.034	0.114	0.576	0.943	1.00
$T = 1000$	0.045	0.257	0.929	0.999	1.00
n = 20	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.034	0.044	0.083	0.210	0.739
$T = 100$	0.038	0.067	0.173	0.526	0.984
$T = 200$	0.041	0.097	0.444	0.917	1.00
$T = 500$	0.034	0.219	0.944	1.00	1.00
$T = 1000$	0.045	1.00	1.00	1.00	1.00
n = 20	$\zeta = 0.2$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.033	0.063	0.172	0.459	0.984
$T = 100$	0.036	0.089	0.415	0.888	1.00
$T = 200$	0.029	0.166	0.830	1.00	1.00
$T = 500$	0.037	0.508	100	1.00	1.00
$T = 1000$	0.045	0.926	1.00	1.00	1.00
n = 50	$\zeta = 0.05$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.048	0.056	0.148	0.366	0.949
$T = 100$	0.041	0.079	0.264	0.763	1.00
$T = 200$	0.051	0.141	0.716	1.00	1.00
$T = 500$	0.059	0.383	1.00	1.00	1.00
$T = 1000$	0.046	0.862	1.00	1.00	1.00
n = 50	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.048	0.075	0.281	0.758	1.00
$T = 100$	0.041	0.125	0.677	0.993	1.00
$T = 200$	0.051	0.304	0.990	1.00	1.00
$T = 500$	0.059	0.810	1.00	1.00	1.00
$T = 1000$	0.045	1.00	1.00	1.00	1.00
n = 50	$\zeta = 0.2$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.049	0.104	0.608	0.988	1.00
$T = 100$	0.041	0.244	0.974	1.00	1.00
$T = 200$	0.051	0.602	1.00	1.00	1.00
$T = 500$	0.059	0.997	1.00	1.00	1.00
$T = 1000$	0.046	1.00	1.00	1.00	1.00

Power and Size Analysis of the Test (2.3) with $\rho = (0.45, 0.3, 0.15) \in \mathbb{R}^3$. The DGP follows a multivariate normal distribution where ζ describes the expected portion of pairs that are correlated with each other with correlation κ^2 and variance $\sigma_i^2 = 2$ for all $i \in \{1, ..., n\}$. The amount of draws from the limit distribution is $B = 400$ by 701 Monte Carlo repetitions.

Table 2.2: Size and Power of S_χ for SAR(3)

n = 20	$\zeta = 0.05$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.028	0.025	0.040	0.057	0.220
$T = 100$	0.042	0.030	0.072	0.175	0.679
$T = 200$	0.038	0.057	0.158	0.503	0.965
$T = 500$	0.047	0.121	0.567	0.988	1.00
$T = 1000$	0.055	0.233	0.922	0.998	1.00
n = 20	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.028	0.037	0.045	0.092	0.561
$T = 100$	0.042	0.057	0.133	0.414	0.987
$T = 200$	0.038	0.060	0.384	0.912	1.00
$T = 500$	0.047	0.238	0.937	1.00	1.00
$T = 1000$	0.055	0.591	0.998	1.00	1.00
n = 20	$\zeta = 0.2$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.028	0.033	0.060	0.253	0.954
$T = 100$	0.043	0.063	0.346	0.844	1.00
$T = 200$	0.039	0.113	0.831	0.997	1.00
$T = 500$	0.047	0.483	1.00	1.00	1.00
$T = 1000$	0.055	0.894	1.00	1.00	1.00
n = 50	$\zeta = 0.05$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.005	0.014	0.016	0.065	0.709
$T = 100$	0.018	0.027	0.156	0.601	1.00
$T = 200$	0.030	0.771	0.617	0.991	1.00
$T = 500$	0.033	0.369	0.998	1.00	1.00
$T = 1000$	0.047	0.829	1.00	1.00	1.00
n = 50	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.005	0.010	0.0411	0.330	0.995
$T = 100$	0.018	0.047	0.045	0.989	1.00
$T = 200$	0.023	0.164	0.982	1.00	1.00
$T = 500$	0.033	0.773	1.00	1.00	1.00
$T = 1000$	0.047	0.999	1.00	1.00	1.00
n = 50	$\zeta = 0.2$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.005	0.013	0.203	0.900	0.997
$T = 100$	0.018	0.106	0.927	1.00	1.00
$T = 200$	0.030	0.435	1.00	1.00	1.00
$T = 500$	0.033	0.997	1.00	1.00	1.00
$T = 1000$	0.047	1.00	1.00	1.00	1.00

Power and Size Analysis of the Test (2.8) S_χ where ζ describes the expected portion of pairs that are correlated with each other with correlation κ^2 and variance $\sigma_i^2 = 2$ for all $i \in \{1, ..., n\}$.

To summarize, both tests show good size and power properties whenever the ratio T over n is greater or equal to 10. Based on the simple limiting distribution of S_{χ}^{*}, the test S_{χ} is also very easy to implement since the approximation test S_{χ} requires only the empirical covariance matrix of the residuals.
The second MC simulation extends the investigations. Here, we consider a SAR(4) model

$$\boldsymbol{y}_t = \rho_1 W_1 \boldsymbol{y}_t + \rho_2 W_2 \boldsymbol{y}_t + \rho_3 W_3 \boldsymbol{y}_t + \rho_4 W_4 \boldsymbol{y}_t + \boldsymbol{\varepsilon}_t,\ t = 1, \ldots, T,$$

where W_1 is a group interaction matrix of the first two-thirds, W_2 is a group interaction matrix of the last one-third, W_3 a binary contiguity matrix of the third-order neighbors only (the observations $1, ..., n$ are assumed to be in a circle, i.e. 2 is a neighbor of $n-1, n, 1, 3, 4, 5$ and[13]

$$(W_4)_{ij} = \begin{cases} \frac{1}{2 \cdot \lfloor n-1 \rceil},\ \text{if } i \text{ is even and } j \text{ odd or vice versa} \\ 0,\ \text{otherwise.} \end{cases}$$

The weighting vector ρ is given by $\boldsymbol{\rho} = (-0.2\ 0.05\ 0.1\ 0.5)$. Moreover, we presuppose heteroscedastic normal error terms, i.e. $\sigma_i \sim N(0,1)$ for $i = 1, ..., n$. In order to analyze the power in case of misspecification, we choose ζ and κ likewise to the first MC simulation. To determine the size and power we follow the recommendations given in MacKinnon (2002) and draw $B = 400$ times from the asymptotic limit distribution given in Lemma (C.1.4). The overall number of MC repetitions is equal to 701. The results of the tests can be found in Table 2.3.
Even if the results of the second analysis are not one-to-one comparable with those from the first simulation,[14] it is clearly observable that the tests hold the size level. The power increases if either the correlation structure (κ or ζ) or the amount of observation grows (n or T). Thus, the results presented in the second, more complex study are in line with those given in the first simulation.

[13]Matrices W_1, W_2, W_3 are the counterparts to the matrices $G1, G2, BC3$ given in Elhorst et al. (2012).

[14]The model presupposes heteroscedasticity and the spatial structure is completely different. From this it follows, that the violation of the moment condition (A.1) is not one-to-one comparable.

In summary, the MC study has shown that the test is also applicable in case of small samples as long as the vector of observations is sufficiently large compared to the cross sectional dimension n. The next section shows that the test even holds size and power demands if the error terms follow a GARCH process.

2.4.1 GARCH(1,1)

One of the many problems researchers and practitioners face when analyzing data series in financial markets is their structure. Thus, volatility of financial assets has been extensively studied in the last twenty years. An important aspect is volatility clustering, i.e. conditional heteroskedasticity, which leads to an increase in the probability of rare events, that is often modelled with GARCH errors. Since the SAR(m) model is a powerful instrument in modelling financial data[15], the third and final Monte Carlo simulation for the suggested test (2.6) assumes that the errors of the data generating process (DGP) are driven by a GARCH(1,1) model, i.e. for $t = 1, ..., T$ and $i = 1, ...n$

$$
\begin{aligned}
&y_{it} = \sigma_{it}(I_n - \rho_1 W_1 - \rho_2 W_2 - \rho_3 W_3)^{-1}\epsilon_{it}, \\
&\sigma_{it}^2 = 0.33 + 0.33\sigma_{i(t-1)}^2 + 0.075y_{i(t-1)}^2, \\
&\epsilon_{it} \overset{i.i.d}{\sim} N(0,1).
\end{aligned}
$$

To receive comparable results, the weighting matrices W_1, W_2, W_3 are similar to those of the first MC simulation of Section 2.4. The size and power results are presented in Table 2.4. At first, it should be noted that the amount of observation of a GARCH adjusted data set needs to be significantly higher compared to a data set with no GARCH adjustment, since for the case of a GARCH adjustment an initial estimate needs to be conducted. Thus, a primarily high error of estimation distorts the stationarity assumption. However, with a sufficiently large set of observations, the test S (2.6) performs also well with reference to size and power.

[15]The empirical analysis in Section 2.5 shows, that a SAR(3) seems reasonable in times of no economic crisis.

Table 2.3: Size and Power of S for SAR(4)

$\mathbf{n = 60}$	$\zeta = 0.05$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.036	0.049	0.073	0.134	0.438
$T = 100$	0.043	0.050	0.114	0.263	0.805
$T = 200$	0.036	0.069	0.227	0.640	0.989
$T = 500$	0.035	0.117	0.681	0.979	1.00
$T = 1000$	0.050	0.329	0.960	0.997	1.00
$\mathbf{n = 60}$	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.036	0.061	0.096	0.251	0.782
$T = 100$	0.043	0.084	0.208	0.629	0.991
$T = 200$	0.036	0.097	0.509	0.939	1.00
$T = 500$	0.036	0.270	0.976	1.00	1.00
$T = 1000$	0.050	0.684	1.00	1.00	1.00
$\mathbf{n = 60}$	$\zeta = 0.2$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.036	0.074	0.191	0.535	0.988
$T = 100$	0.043	0.103	0.479	0.949	1.00
$T = 200$	0.036	0.193	0.919	1.00	1.00
$T = 500$	0.036	0.604	1.00	1.00	1.00
$T = 1000$	0.050	0.962	1.00	1.00	1.00
$\mathbf{n = 90}$	$\zeta = 0.05$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.029	0.059	0.080	0.191	0.685
$T = 100$	0.054	0.064	0.176	0.461	0.972
$T = 200$	0.044	0.101	0.398	0.893	1.00
$T = 500$	0.046	0.214	0.930	1.00	1.00
$T = 1000$	0.047	0.551	1.00	1.00	1.00
$\mathbf{n = 90}$	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.029	0.074	0.161	0.431	0.942
$T = 100$	0.054	0.080	0.382	0.853	1.00
$T = 200$	0.044	0.164	0.810	1.00	1.00
$T = 500$	0.046	0.503	0.997	1.00	1.00
$T = 1000$	0.047	0.930	1.00	1.00	1.00
$\mathbf{n = 90}$	$\zeta = 0.2$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 50$	0.029	0.094	0.365	0.806	0.999
$T = 100$	0.054	0.127	0.752	0.999	1.00
$T = 200$	0.044	0.338	0.997	1.00	1.00
$T = 500$	0.046	0.888	1.00	1.00	1.00
$T = 1000$	0.047	0.999	1.00	1.00	1.00

Power and Size Analysis of the test S with $\boldsymbol{\rho} = (-0.2\ 0.05\ 0.1\ 0.5)$. The errors are heteroscedastic, i.e. $\sigma_i \sim N(0,1)$, $i = 1, ...n$. The parameter ζ describes the portion of expected pairs of firms that are correlated to each other with correlation intensity κ^2.

Table 2.4: Size and Power of S under GARCH model for SAR(3)

$\mathbf{n=50}$	$\zeta = 0.02$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.086	0.118	0.211	0.490	0.960
$T = 1500$	0.078	0.128	0.331	0.719	0.996
$T = 2000$	0.062	0.140	0.459	0.906	1.00
$T = 2500$	0.068	0.156	0.565	0.960	1.00
$T = 3000$	0.044	0.114	0.673	0.988	1.00
$\mathbf{n=50}$	$\zeta = 0.04$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.086	0.126	0.440	0.881	1.00
$T = 1500$	0.078	0.178	0.699	0.986	1.00
$T = 2000$	0.062	0.156	0.872	1.00	1.00
$T = 2500$	0.068	0.250	0.936	1.00	1.00
$T = 3000$	0.044	0.315	0.972	1.00	1.00
$\mathbf{n=50}$	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.086	0.253	0.932	1.00	1.00
$T = 1500$	0.078	0.425	0.994	1.00	1.00
$T = 2000$	0.062	0.520	1.00	1.00	1.00
$T = 2500$	0.068	0.711	1.00	1.00	1.00
$T = 3000$	0.044	0.792	1.00	1.00	1.00
$\mathbf{n=80}$	$\zeta = 0.02$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.073	0.129	0.334	0.810	1.00
$T = 1500$	0.070	0.126	0.518	0.964	1.00
$T = 2000$	0.043	0.143	0.771	0.997	1.00
$T = 2500$	0.060	0.206	0.877	1.00	1.00
$T = 3000$	0.050	0.229	0.954	1.00	1.00
$\mathbf{n=80}$	$\zeta = 0.04$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.073	0.128	0.709	0.998	1.00
$T = 1500$	0.070	0.202	0.954	1.00	1.00
$T = 2000$	0.043	0.291	0.990	1.00	1.00
$T = 2500$	0.060	0.409	1.00	1.00	1.00
$T = 3000$	0.050	0.517	1.00	1.00	1.00
$\mathbf{n=80}$	$\zeta = 0.1$				
	$\kappa = 0.00$	$\kappa = 0.05$	$\kappa = 0.1$	$\kappa = 0.15$	$\kappa = 0.25$
$T = 1000$	0.073	0.416	0.998	1.00	1.00
$T = 1500$	0.070	0.607	1.00	1.00	1.00
$T = 2000$	0.043	0.826	1.00	1.00	1.00
$T = 2500$	0.060	0.972	1.00	1.00	1.00
$T = 3000$	0.050	0.988	1.00	1.00	1.00

Power and Size Analysis of the Test (2.3) S^* with $\rho = (0.45, 0.3, 0.15)$ under a GARCH model. The data generating process is GARCH(1,1) with constant and GARCH parameter equal to 0.33 and ARCH parameter equal to 0.075 with standard normal errors. ζ describes the expected portion of pairs that are correlated with each other with correlation κ^2 and variance $\sigma_i^2 = 2$ for all $i \in \{1, ..., n\}$.

2.5 Empirical Analysis

We analyze the spatial dependencies in the daily stock returns of the Euro Stoxx 50 members in the composition of January 2010 for the period from 2003 until 2009, using adjusted stock prices from Datastream which we transfer to log returns. Our basic model for the stock returns on day $t = 1, ..., T$, is

$$\boldsymbol{y}_t = \rho_g W_g \boldsymbol{y}_t + \rho_b W_b \boldsymbol{y}_t + \rho_l W_l \boldsymbol{y}_t + \boldsymbol{\varepsilon}_t \tag{2.11}$$

where $\boldsymbol{y}_t$ is the vector of stock returns on day t while the weighting matrices W_g, W_b, W_l capture general dependencies[16], dependencies insides branches and local dependencies[17] . In all weighting matrices, market capitalization is taken into account, i.e. the share of the respective firm is written into the non-zero entries of the rows, before the matrices are row-standardized to 1.

Thus, none of the spatial matrices W_g, W_b, W_l are symmetric. The unknown parameters ρ_g, ρ_b and ρ_l represent the corresponding factors. Our main interest is to provide statistical evidence whenever the spatial model (2.11) is applicable. Therefore, we conducted an extensive empirical analysis for the transferred initial data set to log returns with and without a GARCH(1,1) adjustment.

Figure 2.1 shows a rolling window parameter estimation for $\boldsymbol{\rho}$ for a window of size $T = 100$ in a data set of size 1861 of the Euro Stoxx 50 from 2003 until 2009. The blue line equals the ratio of the 95%-quantile of the limit distribution over the value of the test statistic of S (2.6).[18] Thus, the null hypothesis is rejected whenever the value of the blue line is smaller than 1. Figure 2.1 illustrates that in periods of economic crisis the spatial model (2.11) is not applicable. This is consistent with the observation that in times of bear markets the correlation among market participants rises dramatically. The resulting extensive dependency structure cannot be captured by the simple spatial model (2.11). Accordingly, the findings of our test give evidence that

[16]The elements of this matrix are non-zero outside the diagonal and all these entries in a single row have the same value, so that it captures impacts which affect all stocks in a similar way like prior performance of stock markets.

[17]For the partitioning of the Euro Stoxx 50 members into branches and countries we refer to Table 2.5.

[18]The results for the second test statistic S_{χ^2} are similar, so we omit them.

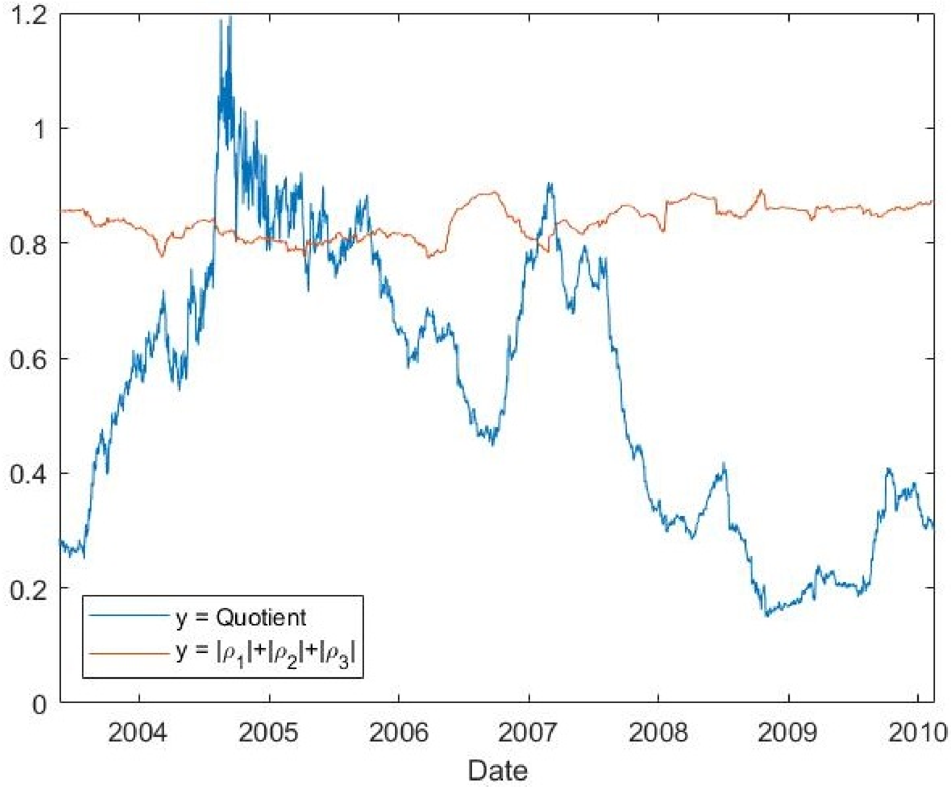
Table 2.5: Partitioning of Euro Stoxx 50 members into branches and countries.

Finance	Aegon, Allianz, AXA, Banco Bilbao, Banco Santander, BNP, Crédit Agricole, Deutsche Bank, Deutsche Börse, Generali, ING, Intesa, Münchener Rück, Société Générale, Unicredit
Automobil	Daimler, Renault, VW
Energy	Alstom, E.ON, ENEL, ENI, Iberdrola, Repsol, RWE, SUEZ, Total
Telecom and Media	Dt. Telekom, France Telecom, Telecom Italia, Telefonica, Vivendi
Pharma and Chemicals	Air Liquide, BASF, Bayer, Sanofi
Consumer Electronics	Nokia, Philips, SAP, Siemens, Schneider
Consumer retail	Anheuser Busch, Carrefour, Danone, L'Oreal, LVMH, Unilever
Basic Industry	Arcelor Mittal, CRH, Saint Gobain, Vinci
Benelux	Aegon, Anheuser Busch, Arcelor, ING, Philips, Unilever
France	Air Liquide, Alstom, AXA, BNP, Carrefour, Crédit, Agricole, France Telecom, Danone, L'Oreal, LVMH, Saint Gobain, Sanofi, Schneider, Société Générale, SUEZ, Total, Vinci, Vivendi
Germany	Allianz, BASF, Bayer, Daimler, Deutsche Bank, Deutsche Börse, Dt. Telekom, E.ON, Münchner Rück, RWE, SAP, Siemens, VW
Italy	Generali, ENEL, ENI, Intesa, Telecom Italia, Unicredito
Spain	Banco Bilbao, Banco Santander, Iberdrola, Repsol, Telefonica
Others	CRH, Nokia

Both matrices are constructed in the following way: The off-diagonal elements are nonzero if the corresponding stocks belong to the same branch (W_b) or country (W_l). In each row, the nonzero entries are identical and sum up to 1 (row-wise).

the effects of the dot-com bubble crisis around 2000 last until summer 2004, since the test declines to apply model (2.11). In the two following years, Figure 2.1 depicts evidence to apply the model, since the blue line is often greater than 1. However, roughly speaking, from the beginning of 2006 until the end of the observation period the test indicates that a spatial model is inappropriate. This in accordance with the financial crisis, that started in summer of 2006. We continue our empirical analysis by looking at Value-at-Risk (VaR) forecasts to see if our new specification test could also be used as a backtest in the spirit of Ziggel et al. (2014) among others. Figure 2.4 depicts the VaR forecast with standard normally distributed errors for the minimum variance optimal port-

Figure 2.1: Rolling Window for $T = 100$

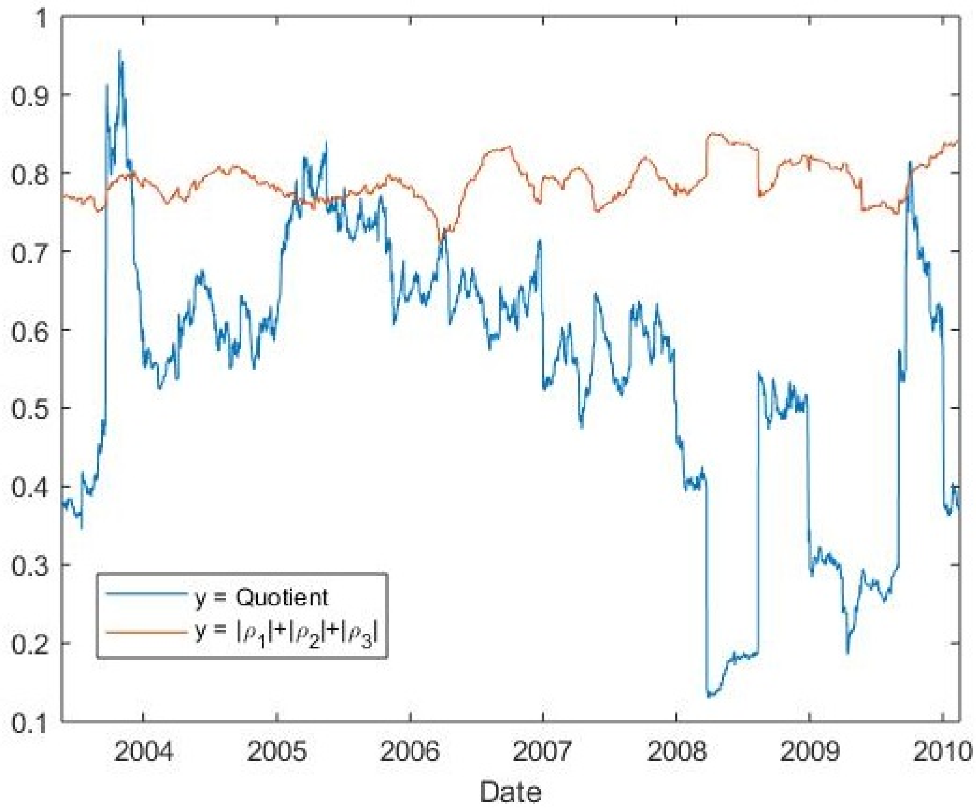


Rolling window parameter estimation for $\boldsymbol{\rho}$ for a window of size $T = 100$ in a data set of size $T = 1861$ and dimension $n = 50$. The number of Bootstrap repetitions is equal to 400. The blue line depicts the ratio of the 95%-quantile of the limit distribution given in Lemma (C.1.4) over the test statistic S from (2.3). The orange line is the accumulated spatial dependence parameter $\boldsymbol{\rho}$ within the L_1-norm.

folio based on a rolling historical window of $T = 50$. We observe that in times of moderate economic peaks (2004-2006), where the test provides statistical evidence for a spatial model, the VaR forecasts also seem to be accurate. In times of crises the test (2.6) rejects the null, such that both the spatial model (2.11) and VaR forecasts seem to be inappropriate instruments to describe the prevailing market situation. The facts that in time of a crisis both the SAR(3) is rejected and a superproportional number of VaR-exceedances are observed, are two consequences of high market volatility and cross-sectional dispersions in the stock market, such that the proposed testing procedures can also be applied as a VaR backtest.

Figure 2.2 is the analogon to Figure 2.1 under a GARCH(1,1) adjustment and it depicts that the overall structure is in accordance with those from Figure 2.1.

Figure 2.2: Rolling Window for $T = 100$ under GARCH(1,1) adjustment

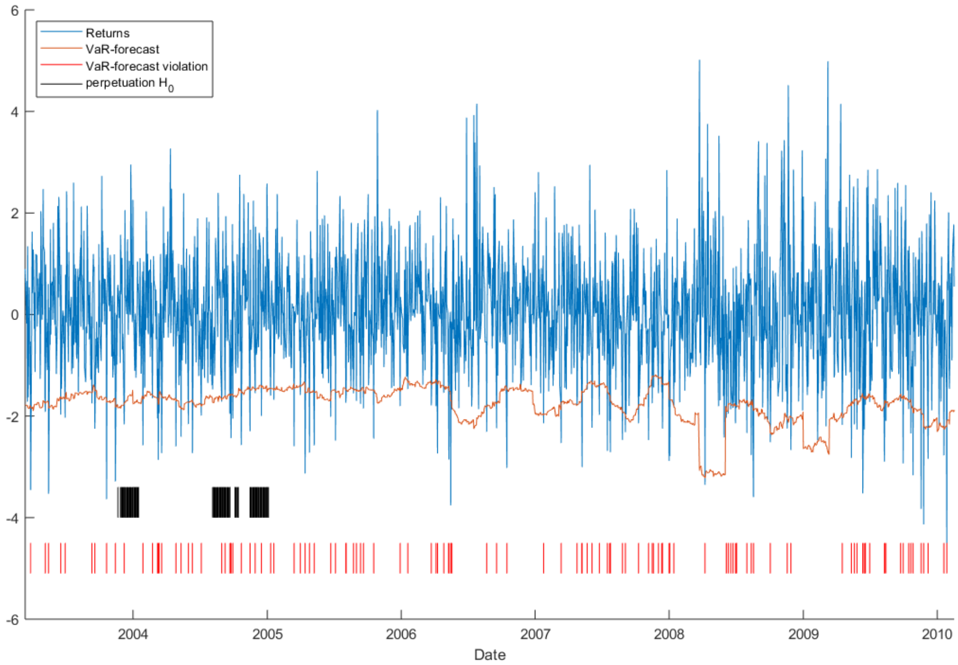


Rolling window parameter estimation for $\boldsymbol{\rho}$ for a window of size $T = 100$ in a data set of size $T = 1861$ and dimension $n = 50$ under GARCH(1,1) adjustment. The number of Bootstrap repetitions is equal to 400. The blue line depicts the ration of the 95%-quantile of the limit distribution given in Lemma (C.1.4) over the test statistic S from (2.3). The orange line is the accumulated spatial dependence parameter $\boldsymbol{\rho}$ within the L_1-norm.

Beyond that, the GARCH(1,1) filter seems to point out the typical scope of application of spatial models, that in times of economic crisis classical SAR(m) models seem to be too restrictive and not complex enough. This is consistent with the results given in Figure 2.3, which shows, that VaR-forecasts are less violated in moderate economic times compared to an economic depression[19]. Furthermore, the amount of clusters regarding VaR violations decreases from Figure 2.4 compared to Figure 2.3, where there is less clustering of VaR violations. In bear markets, however, clustering is still clearly observable which is in accordance with the findings that the extensive structure could not be fully captured by the spatial model (2.11). Overall, our empirical investigation

[19]cf. roughly summer 2006 until end of the data set

Figure 2.3: VaR-Forecasts for $T = 50$ under GARCH(1,1) adjustment



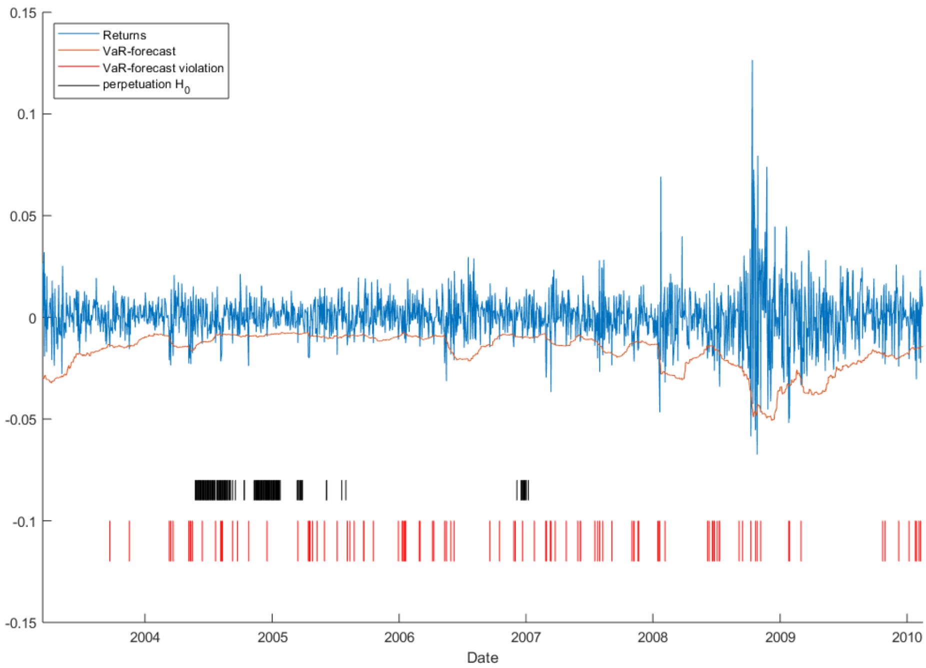
The figure depicts a Value-at-Risk forecast with standard normal distributed errors based on the Euro Stoxx 50 from 2003 to 2009 in a rolling window of size 50 under a GARCH(1,1) adjustment. The orange line represents the VaR-forecast based on the data. The blue line represents the returns. A VaR violation is reported with a red dashed line at the bottom of that figure. The black line indicates statistical significance to apply the spatial model (2.1). The VaR-level is chosen at 0.05 and the number of Bootstrap repetitions is equal to 400.

shows the tests' ability to detect misspecifications for classical SAR models for both the initial and for a GARCH adjusted data set.

2.6 Conclusion

We propose two specification tests for spatial models and analyze the size and power of these tests. The proposed tests show good size and power properties in finite samples for both initial data and GARCH adjusted data. An empirical analysis of the Euro Stoxx 50 between 2003 and 2009 substantiates that bull markets provide statistical evidence to apply a SAR(3) model. However, in bear markets a simple spatial model does not capture the extensive structure of relations and dependencies in the market. Accordingly, the test provided statistical evidence for the empirical observation that both, the time after the

Figure 2.4: VaR-Forecasts for $T = 50$



The figure depicts a Value-at-Risk forecast with standard normal distributed errors based on the Euro Stoxx 50 from 2003 to 2009 in a rolling window of size 50. The orange line represents the VaR-forecast based on the data. The blue line represents the returns. A VaR violation is reported with a red dashed line at the bottom of that figure. The black line indicates statistical significance to apply the spatial model (2.1). The VaR-level is chosen at 0.05 and the number of Bootstrap repetitions is equal to 400.

dot-com bubble and the time around the Lehman Brothers bankruptcy could not be captured correctly by a spatial model which models only a general, branches and national dependence. For that reason it seems to be useful to introduce a test which provides statistical evidence if a given data set fulfills the assumptions of a classical SAR(m) model. To the best of the authors knowledge this is the first specification test for a classical SAR(m).

An interesting task for further research would be to see if the new specification test can be reasonably combined with the test for structural changes proposed in Wied (2013). Maybe, structural changes are a key reason for misspecification. Also, one could think about extending the ideas in this paper to extensions of the SAR model including additional exogenous regressors.

CHAPTER 3

Testing for Relevant Dependence Change in Financial Data: A CUSUM Copula Approach

3.1 Abstract

We propose a new non-parametric test for detecting relevant breaks in copula functions. We assume that the data is driven by two non-equal copulas C_1 and C_2. Under the null hypothesis, the copula difference within an appropriate norm is smaller than a certain positive adjustable threshold Δ. Within the alternative hypothesis, the copula difference exceeds the fixed value Δ. The test is based on a cumulative sum approach of the empirical copula with sequentially estimated marginals. We propose a bootstrap procedure to compute critical values. The Monte Carlo simulation indicates that the test results in a reasonable sized and powered testing procedure. A real data application of the DAX30 up to cross sectional dimension $N = 30$ shows the test's ability to detect relevant break points.

3.2 Introduction

It is well known that dependencies within a portfolio increase in times of financial crisis (Aloui et al., 2011). From a portfolio manager point of view the increase of the dependencies is disadvantageous, which is known as the diversification effect. In fact, investors are interested in decreasing the dependencies by rescheduling the portfolio to lower the risk of losses. One of many approaches to detect those changes in the dependence structure is to test for changes in the copula function. For instance, Brodsky et al. (2009); Busetti F. (2011); Krämer & van Kampen (2011) have designed nonparametric tests for breaks in the copula in a fixed point considering N-dimensional random vectors. Bücher (2013) extended their approaches by testing for overall constancy of the copula in the case of known marginal distributions, while the test of constancy suggested in Bücher et al. (2014) considers sequentially estimated marginals. Wied et al. (2013) propose a test for changes in Spearman's rho, Dehling et al. (2017) consider a test for changes in Kendall's tau. Manner et al. (2019) construct a parametric test for detecting breaks in the parameters of factor copula models. The above mentioned tests can be applied to detect and quantify contagions between different financial markets or to construct optimal portfolios.

All the proposed methods test for the "classical" hypothesis, meaning that they test for stationarity in a sequence of random vectors $\{X_j\}_{j=1}^T$ with $X_j \in \mathbb{R}^N$, i.e.

$$H_0 : X_1, X_2, \ldots, X_T \sim F.$$

with the alternative in the simplest case of one structural breakpoint in time (Dette & Wied, 2016)

$$H_1 : X_1, X_2, \ldots, X_j \sim F_1 \text{ and } X_{j+1}, \ldots, X_T \sim F_2,$$

where the distribution function changes from F_1 to F_2 with $F_1 \overset{d}{\neq} F_2$ at time $j \in \{1, \ldots T\}$, i.e. F_1 and F_2 are not equal in distribution.

A general issue of such hypothesis testing is the consistency problem, i.e. any consistent test will detect any arbitrary small change in the parameters if the sample size is sufficiently large. This discrepancy was mentioned for the first time in 1938 by Berkson (1938).

Beyond that, in the case of small changes the rejection of the null might result in an unnecessary break point estimation and an expensive adjustment of the considered model. In practice, small changes in the data might not be crucial, since they do not necessarily add up to significant changes. Thus, the gain derived by the detected break point could be negatively overcompensated by the costs of adjusting the model (e.g. in case of portfolio theory these can be interpreted as transaction costs) or to be short and to the point: Significance does not necessarily imply relevance.

Therefore, we impose the more realistic assumption that our sequence of random vectors $\{X_j\}_{j=1}^T$ with $X_j \in \mathbb{R}^N$ with $j \in \{1, \dots T\}$ is driven by the distribution function F_1 and F_2, i.e. $X_1, X_2, \dots, X_j \sim F_1$ and $X_{j+1}, \dots, X_T \sim F_2$ for some $j \in \{1, \dots T\}$ such that

$$H_0 : \|F_1 - F_2\| \leq \Delta \quad \text{versus} \quad H_1 : \|F_1 - F_2\| > \Delta, \tag{3.1}$$

where $\|.\|$ is an appropriate norm and $\Delta > 0$ a fixed adjustable size. The framework in (3.1) allows for a break in the data (classical break point tests do not) and the adjustable size Δ could serve as a measure to control for the extent of the change.

Dette & Wied (2016) proposed a general approach to this problem. Later on, Dette et al. (2018) and Dette et al. (2018) extended this to the detection of changes in second-order characteristics and to high-dimensional models, respectively. Motivated by their analysis, we are interested in augmenting the literature of testing for relevant breaks in the copula of random vectors by a nonparametric testing procedure that detects relevant changes in the copula function with sequentially estimated marginal distributions. Thus, the testing problem is given by

$$H_0 : \|C_1 - C_2\| \leq \Delta \quad \text{versus} \quad H_1 : \|C_1 - C_2\| > \Delta,$$

where $C_1 \neq C_2$ are copulas and $\Delta > 0$ fixed. As the copula measures the dependence between random variables, we therefore test whether the dependence structure changes more than some given threshold Δ.

Coming back to portfolio management, a small increase in the dependence structure of a portfolio does not necessarily indicate the need to reschedule the portfolio, since transaction costs could overcompensate the benefits of the new, more risk diversified portfolio. Only a relevant change in the dependence structure, i.e. the copula difference within a certain norm is larger than Δ, should result in rescheduling the portfolio.

In our empirical application we analyzed the German DAX30 data of cross sectional dimension $N = 30$ between January 2003 and July 2015. Here, Δ could be interpreted as the largest admissible copula difference such that the relevant change hypothesis is not rejected. Every other choice of Δ that is smaller leads to a rejection of the null hypothesis. As a result, Δ can also be considered as a measure that quantifies the extent of a crisis (given that dependencies of financial returns are usually larger in times of crises).
The rest of the paper is structured as follows. Section 3.3 introduces the considered null hypothesis and test statistic, where Section 3.4 presents the bootstrap procedure to determine critical values to perform the test. Results from the Monte Carlo simulations can be found in Section 3.5. Section 3.6 presents our empirical application and Section 3.7 concludes.

3.3 Relevant change and test statistic

In this section we introduce the null hypothesis, the assumptions and the the relevant change characteristic of our testing procedure in a fully non-parametric setting.
Let $X_1, \ldots, X_T$ denote N-dimensional random vectors and $\mathbf{U}_1, \ldots, \mathbf{U}_T$ the vector of the marginal distributions, i.e. $\mathbf{U}_t := (F_1(X_{t1}), \ldots, F_N(X_{tN}))$ for $t = 1, ..., T$ where $F_i(\cdot)$ is the i-th marginal such that

$$\begin{aligned} \mathbf{U}_1, \ldots, \mathbf{U}_{\lfloor sT \rfloor} &\sim C_1(\mathbf{u}) \\ \mathbf{U}_{\lfloor sT \rfloor + 1}, \ldots, \mathbf{U}_T &\sim C_2(\mathbf{u}), \end{aligned} \tag{3.2}$$

where $\mathbf{u} \in [0,1]^N$ and $C_1, C_2 : [0,1]^N \to [0,1]$ are copulas which capture the dependencies between the components of $X_1, ..., X_{\lfloor sT \rfloor}$ and $X_{\lfloor sT \rfloor + 1}, ..., X_T$, respectively. Here, $\lfloor sT \rfloor$ denotes the change point in time, where T is the size of the sample and $s \in (0,1)$. Note, that the model set-up (3.2) is valid under both the null and the alternative hypothesis. In order to achieve reliable results, classical concepts of dependencies (e.g. $\mathbf{U}_1, ..., \mathbf{U}_T$ is stationary and strong mixing with coefficients converging sufficiently fast to 0) are not applicable any more in the setting of detecting relevant changepoints, because the general model set-up (3.2) of relevant changepoint analysis allows the se-

quence $\mathbf{U}_1, ..., \mathbf{U}_T$ to be non-stationary. That is why we have to impose the assumption of a triangular array, that is α-mixing[1].
To aggregate over $\mathbf{u}$, we consider the L^2-norm $\|\cdot\|_{L^2}$. Thus, the null hypothesis of no relevant change in the copula function is given by

$$H_0 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} \leq \Delta$$

versus the alternative

$$H_1 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} > \Delta,$$

where $\|.\|_{L^2}$ is the L^2-norm and $\Delta > 0$ fixed. For every $\mathbf{u} \in [0,1]^N$ and $t \in (0,1)$ the cumulative sum (CUSUM) type process for detecting changes in the copula is then

$$\hat{\mathbb{U}}_T^*(t,\mathbf{u}) := t(1-t)\left(\frac{1}{\lfloor tT \rfloor}\sum_{i=1}^{\lfloor tT \rfloor} Z_i^{1:\lfloor tT \rfloor}(\mathbf{u}) - \frac{1}{T-\lfloor tT \rfloor}\sum_{i=\lfloor tT \rfloor+1}^{T} Z_i^{\lfloor tT \rfloor+1:T}(\mathbf{u})\right) \quad (3.3)$$

with $Z_i^{t_1:t_2}(\mathbf{u}) := \mathbb{1}\{\hat{F}_1^{t_1:t_2}(X_{i1}) \leq u_1, \ldots, \hat{F}_N^{t_1:t_2}(X_{iN}) \leq u_N\}$, $t_1 < t_2 \in \{1, \ldots, T\}$ for $i = 1, \ldots, T$ and $\mathbb{1}\{\cdot\}$ the indicator function. Here $\hat{F}_j^{t_1:t_2}(\cdot)$ is the empirical distribution function, using data information between t_1 and t_2 and is defined as

$$\hat{F}_j^{t_1:t_2}(x) := \frac{1}{t_2 - t_1 + 1}\sum_{i=t_1}^{t_2} \mathbb{1}\{X_{ij} \leq x\},\ j = 1, ..., N.$$

For the derivation of our testing procedure we now consider $\hat{\mathbb{U}}_T(t,\mathbf{u})$ defined as

$$\hat{\mathbb{U}}_T(t,\mathbf{u}) := t(1-t)\left(\frac{1}{\lfloor tT \rfloor}\sum_{i=1}^{\lfloor tT \rfloor} Z_i(\mathbf{u}) - \frac{1}{T-\lfloor tT \rfloor}\sum_{i=\lfloor tT \rfloor+1}^{T} Z_i(\mathbf{u})\right), \quad (3.4)$$

where $Z_i(\mathbf{u}) := \mathbb{1}\{F_1(X_{i1}) \leq u_1, \ldots, F_N(X_{iN}) \leq u_N\}$, $t_1 < t_2 \in \{1, \ldots, T\}$ with F_i as known marginals, $i = 1, \ldots, T$. For fixed $s \in (0,1)$, a straightfor-

[1] Due to the fact that this discussion is very technical, the details can be found in the corresponding Appendix.

ward calculation yields[2]

$$\lim_{T\to\infty} \mathbb{E}[\hat{\mathbb{U}}_T(t,\mathbf{u})] = \begin{cases} s(1-t)\left(C_1(\mathbf{u}) - C_2(\mathbf{u})\right), & s \le t \\ t(1-s)\left(C_1(\mathbf{u}) - C_2(\mathbf{u})\right), & s > t, \end{cases} \tag{3.5}$$

where we have to distinguish between data before and after the breakpoint $\lfloor sT \rfloor$. In the next step, we want to eliminate the quantile and time dimension $\mathbf{u}$ and t, respectively. For this purpose, we consider the L^2-norm and obtain

$$L(t) := \lim_{T\to\infty} \mathbb{E}[\|\hat{\mathbb{U}}_T(t,\mathbf{u})\|_{L^2}^2] = \begin{cases} s^2(1-t)^2\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2, & t > s \\ (1-s)^2 t^2 \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2, & t \le s, \end{cases}$$

for every norm of the type $\|f(\cdot,\mathbf{u})\|_{L^2}^2 := \int_{[0,1]^N} f(\cdot,\mathbf{u})^2 d\mathbf{u}$. Integrating out t yields

$$\int_0^1 L(t)dt = \frac{s^2(1-s)^2}{3}\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2. \tag{3.6}$$

Thus, integrating out t from the empirical counterpart $\hat{L}_T(t) := \|\hat{\mathbb{U}}_T(t,\mathbf{u})\|_{L^2}^2$ yields the test statistic $\hat{\kappa}_T$ for the initial problem of detecting the relevant change

$$\hat{\kappa}_T := \int_0^1 \hat{L}_T(t)dt. \tag{3.7}$$

Due to the fact that our test statistic mainly consists of an integral over t, we disregard the possibility to trim the sample in some way, as it is sometimes done in the breakpoint literature.

We use $\hat{s} := \underset{s\in(0,1)}{\operatorname{argmax}} \|\hat{\mathbb{U}}_T(s,\mathbf{u})\|_{L^2}$ as the natural argmax estimator for the changepoint location fraction s[3]. We reject the null hypothesis of no relevant change if the test statistic (3.7) less the adjusted centering $\frac{s^2(1-s)^2}{3}\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2$ deviates too far from zero. If the marginal distributions are known, the limiting distribution of the process

[2] For the very detailed derivation of the testing procedure we refer to the corresponding Appendix.

[3] Note, $\hat{s}$ is a superconsistent estimator of the changepoint fraction s with convergence rate T (Dette & Wied, 2016).

$$\sqrt{T}\left(\int_0^1 \hat{L}_T(t)dt - \frac{s^2(1-s)^2}{3}\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2\right) \tag{3.8}$$

is normal which is shown in the corresponding Appendix.
Due to the high computational effort in high dimensions using the L^2-norm it could be reasonable to only test for specific points $\mathbf{q}$ in the copula, e.g. $\mathbf{q}$ could be chosen as the value that maximizes the copula difference, i.e. $\mathbf{q} := \sup_{\mathbf{u}\in[0,1]^N} |C_1(\mathbf{u}) - C_2(\mathbf{u})|$. For this purpose we fix $\mathbf{q} = (q_1, \ldots, q_N)'$. What we call quantile counterpart of the process (3.8) is then given by

$$\sqrt{T}\left(\int_0^1 \hat{L}_T^{\mathbf{q}}(t)dt - \frac{1}{3}s^2(1-s)^2(C_1(\mathbf{q}) - C_2(\mathbf{q}))^2\right), \tag{3.9}$$

where $\hat{L}_T(t)$ from (3.8) is replaced by its quantile version $\hat{L}_T^{\mathbf{q}}(t) := (\hat{\mathbb{U}}_T(s, \mathbf{q}))^2$ for $\mathbf{q} \in [0,1]^N$ fixed. Accordingly, the test statistic $\hat{\kappa}_T^{\mathbf{q}}$ is then defined as

$$\hat{\kappa}_T^{\mathbf{q}} := \int_0^1 \hat{L}_T^{\mathbf{q}}(t)dt. \tag{3.10}$$

Since the limit distributions of the processes (3.8) and (3.9) are not known in case of unknown marginals, we suggest a bootstrap procedure. The null hypothesis will be rejected if the expression in (3.8) or (3.9) is greater than the value of the corresponding quantile, which can be obtained by applying the bootstrap procedure presented in Section 3.4. The test holds the size level if the fixed adjustable threshold Δ is chosen as $\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}$ or for the quantile case $|C_1(\mathbf{u}) - C_2(\mathbf{u})|$. For Δ smaller than this threshold the test is oversized while a larger Δ results in a lower rejection rate. In the application later on, we set $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)$, which is in line with our Monte Carlo simulations. An Δ chosen in this way can be used, for example, to assess the extent of a crisis.
Our Monte Carlo simulations below confirm that the bootstrap results in a reasonably sized and powered testing procedure. For the bootstrap we consider the L^2-norm, but this can be easily adjusted to the quantile version simply by

interchanging the L^2-norm with the absolute value $|\cdot|$ for fixed $\mathbf{q} \in [0,1]^N$.

3.4 Bootstrap and Testing procedure

The bootstrap is based on the natural estimators of the respective terms of the process (3.8) or (3.9). We assume that our sample $\{X_i\}_{i=1}^T$ is serially independently distributed or residual data from pre-estimated time series models e.g. GARCH adjusted data. Further, $\{X_i\}_{i=1}^T$ is compounded of $\{X_i\}_{i=1}^{\lfloor sT \rfloor}$ and $\{X_i\}_{i=\lfloor sT \rfloor+1}^T$, such that there is only one breakpoint location in $\lfloor sT \rfloor$, $s \in (0,1)$ and $\{X_i\}_{i=1}^{\lfloor sT \rfloor} \sim C_1(F(X))$ and $\{X_i\}_{i=\lfloor sT \rfloor+1}^T \sim C_2(F(X))$. Then, the bootstrap procedure suggests the following course of action:

i) Estimate the breakpoint location $\lfloor sT \rfloor$ by $\lfloor \hat{s}T \rfloor$, where $\hat{s}$ is determined by

$$\hat{s} := \underset{s \in (0,1)}{\operatorname{argmax}} \|\hat{\mathbb{U}}_T^*(s, \mathbf{u})\|_{L^2}. \tag{3.11}$$

Sample separately with replacement from $\{X_i\}_{i=1}^{\lfloor \hat{s}T \rfloor}$ and $\{X_i\}_{i=\lfloor \hat{s}T \rfloor+1}^T$ to obtain B bootstrap samples $\{X_i^{(b)}\}_{i=1}^T$, for $b = 1, \ldots, B$.

ii) Estimate the break point location $\lfloor \hat{s}_b T \rfloor$ for each bootstrap sample $\{X_i^{(b)}\}_{i=1}^T$, for $b = 1, \ldots, B$, using adjusted (3.11).

iii) Estimate the copula difference $\Delta_C^b = \|\hat{C}_b^{1:\lfloor \hat{s}_b T \rfloor}(\mathbf{u}) - \hat{C}_b^{\lfloor \hat{s}_b T \rfloor+1:T}(\mathbf{u})\|_{L^2}$ for each bootstrap sample $\{X_i^{(b)}\}_{i=1}^T$, for $b = 1, \ldots, B$, where $\hat{C}_b^{t_1:t_2}$ is the empirical copula estimate with sequentially estimated marginals, using the data from t_1 to t_2.

iii) Calculate the bootstrap versions of the centered expressions (3.8) or (3.9)

$$K^{(b)} := \sqrt{T} \left(\int_0^1 \hat{L}_T^{*b}(t) dt - \frac{1}{3} \hat{s}_b^2 (1 - \hat{s}_b)^2 \Delta_C^b \right),$$

with $\hat{L}_T^{*b}(t) := \|\hat{\mathbb{U}}_T^{*b}(s, \mathbf{u})\|_{L^2}^2$, where $\hat{\mathbb{U}}_T^{*b}(s, \mathbf{u})$ is the bootstrap analogue of (3.3), using $\{X_i^{(b)}\}_{i=1}^T$.

iv) Compute B versions of $K^{(b)}$ and determine the critical value c such that

$$\frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\{K^{(b)} > c\} \stackrel{!}{=} q,$$

where $q \in (0,1)$.

With the above described bootstrap procedure we can calculate critical values for (3.8) and (3.9). The testing procedure is as follows: We reject the null of no relevant change $\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} \leq \Delta$ if

$$\hat{\kappa}_T > \frac{\hat{s}^2(1-\hat{s})^2}{3}\Delta^2 + \frac{b_{1-\alpha}}{\sqrt{T}}, \tag{3.12}$$

where $b_{1-\alpha}$ is the $1-\alpha$ quantile of the bootstrap distribution. Note that the critical values obtained by the bootstrap remain stochastically bounded both under the null and the alternative hypothesis, as the test statistic is always correctly centered.

The bootstrap and testing procedure can be easily adapted for the quantile case by adapting step i) - iii). The test given in equation (3.12) is an exact level α test if Δ is chosen as the copula difference $||C_1(\mathbf{u}) - C_2(\mathbf{u})||_{L^2}$ or $|C_1(\mathbf{q}) - C_2(\mathbf{q})|$. Otherwise, the size is smaller than α. In particular, $\hat{\kappa}_T$ converges weakly to a degenerated random variable if the copula difference is equal to zero and the Davies problem is present, i.e. the break point is unidentified under the null hypothesis. Consequently, the level of the proposed tests have practically size zero, whereas classical stationarity tests hold the asymptotic α-level. Thus, the power of the classical tests is usually larger than the power of the relevant change tests considered here. For practitioners we suggest to run a classical test first, e.g. Bücher (2013) for the case of known marginals and Bücher et al. (2014) in the case of sequentially estimated marginals. If the test rejects the null of stationarity, i.e. the copula difference is significantly larger than zero, estimate the break fraction and apply the proposed relevant change test. This two-step procedure has the drawback, however, that the statistical properties are not clear.

3.5 Quantile- and L^2-Simulations

In this section we want to analyze finite sample properties of our proposed relevant testing procedure, where we simulate multivariate data up to dimension $N = 30$ using a factor copula model following Oh & Patton (2017). We consider both serially independently distributed and residual data.

3.5.1 Serially Independently Distributed Data

In this subsection we conduct two major Monte Carlo simulations. First, we consider the following simple DGP

$$X_t = [X_{1t}, X_{2t}]' = N_2\left(\mathbf{0}, \Sigma_t(\rho)\right), \tag{3.13}$$

where $N_2(\mathbf{0}, \Sigma_t(\rho))$ with $t = 1, ..., T$ describes the bivariate normal distribution with expectation vector zero and covariance matrix $\Sigma_t(\rho) = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$ and $\rho \in [-1, 1]$. We set ρ equal to -0.3 for $t = 1, ..., \frac{T}{2}$ and $\rho = 0.8$ for $t = \frac{T}{2} + 1, ..., T$. Thus, the breakpoint $\lfloor sT \rfloor$ is chosen at $\frac{T}{2}$. We restrict the size analysis in this subsection to the two dimensional case $N = 2$. The following size study presents both L^2-norm based results and an analysis where we consider the specific point $\mathbf{q} = (0.6, 0.6)$. Note, the closer the quantile is to its boundaries, i.e. 0 or 1, the more observations are needed. Critical values of our tests are computed using the bootstrap algorithms from Sections 3.4 with $B = 300$ bootstrap replications. The tests are performed at the $\alpha = 0.05, 0.1$ significance level using 301 Monte Carlo replications. The computations were implemented in Matlab, parallelized and performed using CHEOPS, a scientific High Performance Computer at the Regional Computing Center of the University of Cologne (RRZK).

Table 3.1 presents the results of the relevant change tests under the null with Δ chosen as the estimated copula difference $|C_1(\mathbf{q}) - C_2(\mathbf{q})|$, where C_1 and C_2 are estimated by the consistent copula estimator

$$\hat{C}(\mathbf{u}) = \frac{1}{t_2 - t_1} \sum_{i=t_1}^{t_2} \mathbb{1}\{\hat{F}_1^{t_1:t_2}(X_{i1}) \leq u_1, \ldots, \hat{F}_N^{t_1:t_2}(X_{iN}) \leq u_N\}, \tag{3.14}$$

using realizations $\{X_1, \ldots X_{\lfloor \hat{s}T \rfloor}\}$ and $\{X_{\lfloor \hat{s}T \rfloor+1}, \ldots X_T\}$. The breakpoint $\lfloor \hat{s}T \rfloor$ is estimated by

$$\hat{s} := \underset{s\in(0,1)}{\text{argmax}} \, |\hat{\mathbb{U}}_T(s, \mathbf{q})|. \tag{3.15}$$

Table 3.2 reports the results of the relevant change tests under the null, where the functional difference between the copulas is determined by the L^2-norm. Similar to the quantile case, we consider for the size analysis $\Delta := \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}$ and accordingly $\hat{s} := \underset{s\in(0,1)}{\text{argmax}} \, \|\hat{\mathbb{U}}_T(s, \mathbf{u})\|_{L^2}$. Overall, the tests show good size properties and converges to the predetermined rejection level α if T gets larger. For smaller T, the differences to α are slightly, but not dramatically larger and it does not appear necessary to consider size corrections.

Table 3.1: Size Analysis using Quantile Version 1/4

	$T = 300$	$T = 500$	$T = 750$	$T = 1000$
q_{95}	0.06	0.06	0.04	0.05
q_{90}	0.12	0.11	0.10	0.10

Table 3.1 reports the rejection rate of the relevant change test for data generated with the DGP described in (3.13) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = (0.6, 0.6)$. In total, we conducted 301 Monte Carlo replications.

Table 3.2: Size analysis using the L^2-norm 1/2

	$T = 300$	$T = 500$	$T = 750$	$T = 1000$
q_{95}	0.06	0.06	0.04	0.06
q_{90}	0.11	0.11	0.10	0.12

Table 3.2 reports the rejection rate of the relevant change test for data generated with the DGP described in (3.13) using $B = 300$ bootstrap replications. The copula difference is determined using the L^2-norm. In total, we conducted 301 Monte Carlo replications.

For the power analysis, we use the quantile based test and consider two different scenarios. In the first scenario we keep Δ fix and vary ρ in the DGP (3.13). In the second scenario we vary Δ and keep the DGP (3.13) fixed.
The upper panel of Table 3.3 depicts the first scenario. In this case we deter-

mine Δ_0 as the copula difference at the point $\mathbf{q} = (0.6, 0.6)$ generated by the DGP (3.13) with $\rho = -0.3$ before and $\rho = 0.8$ after the break point at $\frac{T}{2}$. We now vary $\rho \in \{-0.4, -0.5, -0.6, -0.7\}$ before the break point and the results of the rejection rate can be seen in the upper panel of table 3.3 for different sample sizes.

The lower panel of table 3.3 depicts the second scenario. After determining the quantile value under the null, we decrease the tolerance Δ in the test (3.12) by $\Delta = d \cdot \Delta_0$ with $d \in \{0.95, 0.9, 0.85, 0.8\}$.

Note, that in both cases the rejection rate of the relevant change test holds the size level α and the rejection rate tends to 1 for increasing sample size T and decreasing d or ρ. This is the expected behavior as the null and alternative hypothesis differ the more the smaller d or ρ are. In the second major MC

Table 3.3: Power analysis

	Power Analysis varying ρ				
	$\rho = -0.3$	$\rho = -0.4$	$\rho = -0.5$	$\rho = -0.6$	$\rho = -0.7$
$T = 300$	0.06	0.48	0.83	0.98	1.00
$T = 500$	0.06	0.62	0.94	1.00	1.00
$T = 750$	0.04	0.69	0.98	1.00	1.00
$T = 1000$	0.05	0.84	0.99	1.00	1.00
	Power Analysis varying Δ				
	$\Delta = \Delta_0$	$\Delta = 0.95 \cdot \Delta_0$	$\Delta = 0.9 \cdot \Delta_0$	$\Delta = 0.85 \cdot \Delta_0$	$\Delta = 0.8 \cdot \Delta_0$
$T = 300$	0.06	0.31	0.63	0.88	0.97
$T = 500$	0.06	0.33	0.75	0.94	1.00
$T = 750$	0.04	0.36	0.83	0.96	1.00
$T = 1000$	0.05	0.42	0.86	1.00	1.00

Table 3.3 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (3.13) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = (0.6, 0.6)$. Varying $\Delta = d \cdot \Delta_0$, where $d = \{0.95, 0.9, 0.85, 0.8\}$ (lower panel) and $\rho = \{-0.4, -0.5, -0.6, -0.7\}$ in $\Sigma_t(\rho)$ (upper panel) for $t = \lfloor \frac{T}{2} \rfloor + 1, ..., T$. In total, we conducted 301 Monte Carlo repetitions.

simulation, we consider our data to be jointly distributed with a one factor copula model following Oh & Patton (2017), where the marginal distributions are in general unknown and the copula is implied by the following factor structure

$$X_t = [X_{1t}, \dots, X_{Nt}]' = \boldsymbol{\beta}_t Z + \boldsymbol{q}, \tag{3.16}$$

with $\boldsymbol{\beta}_t = \beta_t \cdot (1 \dots, 1)'$ is a parameter vector of size N, $Z \overset{i.i.d.}{\sim} \mathsf{Skew\ t}\,(\nu^{-1}, \lambda)$[4] and $\boldsymbol{q} = [q_{1t}, \dots, q_{Nt}]'$ with $q_{it} \overset{i.i.d.}{\sim} \mathsf{t}\,(\nu^{-1})$ for $i = 1, \dots, N$ and $t = 1, \dots, T$. We fix $\nu^{-1} = 0.25$ and $\lambda = -0.5$, such that our model is parametrized by the single factor loading $\theta_t = \beta_t$ for $t = 1, \dots, T$. The DGP in (3.16) provides left skewed and fat tailed data, which is a common property in financial data applications and also in line with our application in Section 5. We construct a break at $\frac{T}{2}$, where θ_0 denotes the parameter value of the model before and θ_1 the parameter value after the break. For our simulation study we choose $\theta_0 = 1$ and $\theta_1 = 2$. Note again, the test is an exact level α test if and only if Δ is chosen as the copula difference. Table 3.4 reports the results of the relevant change test under the null, where the functional difference is computed with the help of the L^2-norm. Table 3.4 shows, that the test using the proposed bootstrap procedure holds the size level using the DGP (3.16). As expected, the size converges to the corresponding rejection level $\alpha \in \{0.05, 0.1\}$ as T gets larger. This characteristic also holds for Table 3.5. In this case, we set Δ equal

Table 3.4: Size analysis using the L^2-norm 2/2

		$T = 300$	$T = 500$	$T = 750$	$T = 1000$
$N = 2$	q_{95}	0.03	0.04	0.04	0.04
	q_{90}	0.07	0.08	0.11	0.11
$N = 3$	q_{95}	0.03	0.05	0.04	0.05
	q_{90}	0.06	0.10	0.12	0.10
$N = 5$	q_{95}	0.02	0.04	0.07	0.05
	q_{90}	0.04	0.09	0.13	0.09

Table 3.4 shows the rejection rate of the relevant change test for the DGP (3.16) using $B = 300$ bootstrap replication. In total, we conducted 301 Monte Carlo repetitions.

to the copula difference evaluated at the specific point $\mathbf{q} = 0.6\cdot(1, \dots, 1)$, where $(1, \dots, 1)'$ is a N-dimensional vector. The experiment is repeated in Table 3.6 for $\mathbf{q} = 0.1 \cdot (1, \dots, 1)$. Considering such particular quantiles provides

[4]As in Oh & Patton (2017) this refers to the skewed t-distribution by Hansen (1994).

the advantage to conduct high dimensional data analysis with comparatively moderate computational efforts. Thus, the relevant change test is especially suitable for high dimensional data applications. In practice, for instance, the specific quantile $\mathbf{q}$ can be chosen as the quantile that maximizes the copula difference. The simulations show that a higher T is necessary to avoid size distortions if q is close to 0. Table 3.7 presents size results for the setting of Table 3.5 with the modification that the break already occurs at $\frac{T}{4}$. Here, the empirical size is slightly further away from the nominal size, but the differences are minor.

The size analysis in the factor copula setting is completed by analyzing the two-step procedure mentioned at the end of Section 3. This means we first perform the non-parametric copula constancy test proposed in Bücher et al. (2014) and state the rejection frequency. Then, for the rejected runs, we apply the relevant change test, where Δ is chosen as the estimated copula difference, and again state the frequency of rejections, cf. Table 3.8. The frequency in the first step gives the empirical power of the Bücher et al. (2014) test, which tends to 1 for increasing T. The frequency in the second step gives the empirical size of the relevant change test, which is close to the nominal size.

Table 3.5: Size analysis using quantile version 2/4

		$T = 300$	$T = 500$	$T = 750$	$T = 1000$
$N = 2$	q_{95}	0.06	0.04	0.05	0.06
	q_{90}	0.14	0.10	0.13	0.12
$N = 3$	q_{95}	0.08	0.07	0.06	0.06
	q_{90}	0.15	0.13	0.12	0.12
$N = 5$	q_{95}	0.04	0.04	0.05	0.05
	q_{90}	0.10	0.10	0.13	0.12
$N = 30$	q_{95}	0.03	0.04	0.05	0.06
	q_{90}	0.08	0.10	0.09	0.11

Table 3.5 reports the rejection rate of the relevant change test for data generated with the DGP described in (3.16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$. The break is constructed at $\frac{T}{2}$. In total, we conducted 301 Monte Carlo repetitions.

For the power analysis of the quantile based test we consider two different

Table 3.6: Size analysis using quantile version 3/4

		$T = 1000$	$T = 2000$	$T = 4000$
$N = 2$	q_{95}	0.08	0.09	0.08
	q_{90}	0.18	0.17	0.16
$N = 3$	q_{95}	0.09	0.10	0.07
	q_{90}	0.17	0.18	0.15
$N = 5$	q_{95}	0.06	0.09	0.05
	q_{90}	0.12	0.19	0.13
$N = 30$	q_{95}	0.04	0.05	0.05
	q_{90}	0.11	0.12	0.12

Table 3.6 reports the rejection rate of the relevant change test for data generated with the DGP described in (3.16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.1 \cdot (1, \ldots, 1)'$. The break is constructed at $\frac{T}{2}$. In total, we conducted 301 Monte Carlo repetitions.

Table 3.7: Size analysis using quantile version 4/4

		$T = 300$	$T = 500$	$T = 750$	$T = 1000$
$N = 2$	q_{95}	0.06	0.08	0.04	0.06
	q_{90}	0.14	0.15	0.11	0.12
$N = 3$	q_{95}	0.07	0.06	0.05	0.06
	q_{90}	0.15	0.13	0.10	0.11
$N = 5$	q_{95}	0.06	0.06	0.05	0.05
	q_{90}	0.14	0.13	0.12	0.11
$N = 30$	q_{95}	0.06	0.08	0.07	0.07
	q_{90}	0.09	0.13	0.13	0.12

Table 3.7 reports the rejection rate of the relevant change test for data generated with the DGP described in (3.16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$. The break is constructed at $\frac{T}{4}$. In total, we conducted 301 Monte Carlo repetitions.

scenarios. First, we set a fixed Δ while we increase the copula difference by increasing the parameter θ_1 after the break. Second, we keep the parameter values $\theta_0 = 1$ and $\theta_1 = 2$ fixed and decrease Δ, while the starting point for Δ is equal to the implied copula difference at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$.

Table 3.9 reports the rejection rate of the test (3.12) using the 95%−quantile

Table 3.8: Pretesting

	$T = 300$		$T = 500$	
	Pre	Delta	Pre	Delta
$N = 2$	0.43	0.09	0.65	0.08
$N = 3$	0.63	0.05	0.89	0.04
$N = 5$	0.76	0.03	0.95	0.04

Table 3.8 shows the rejection rate of the relevant change test for the DGP (3.16) with a break at $\frac{T}{2}$ with $\theta_0 = 1$ and $\theta_1 = 1.6$ using $B = 300$ bootstrap replication on a significance level of 0.05. In total, we conducted 301 Monte Carlo repetitions. First we performed the Bücher et al. (2014) test (Pre) and for the rejected runs we applied the relevant change test (Delta). The Δ is chosen as the estimated copula difference.

of the proposed bootstrap distribution in Section 3.4. The first column depicts the rejection rate under the null hypothesis. The values of the other columns are obtained by increasing the corresponding copula parameter $\theta_1 \in \{2.2, 2.4, 2.6, 2.8\}$, while Δ remains fixed to the initial copula difference, i.e. $\theta_0 = 1$ and $\theta_1 = 2$.

Table 3.9 illustrates that the power of the test (3.12) generally increases not only if T but also if the cross sectional dimension N increases. For example, the scenario $N = 30$, $T = 750$ and $\theta_1 = 2.6$ always rejects the null hypothesis, i.e. the rejection rate is equal to 1. This is also expected, since we increase the parameter in the factor copula model (3.16) for each component. Consequently, the error is effectively added up which leads to the gain in power.

Table 3.10 provides the power analysis for the setting of Table 3.9 for a break at $\frac{T}{4}$. According to the expectations, the table shows that the power increases for an increasing θ or T. However, the empirical power is lower compared to the setting with a break at $\frac{T}{2}$.

Finally, Table 3.11 analyzes the rejection rate if Δ decreases while the copula difference remains fixed. The value Δ_0 is equal to the copula difference computed at the point $\mathbf{q} = 0.6 \cdot (1, \dots, 1)'$. Now, we decrease Δ stepwise, i.e. $\Delta = d \cdot \Delta_0$ with $d \in \{0.95, 0.9, 0.85, 0.8, 0.75\}$. Table 3.11 shows, that the rejection rate tends to 1 if T increases. Moreover, the power is generally higher for larger N.

Table 3.9: Power analysis for residual data

		$\theta_1 = 2.0$	$\theta_1 = 2.2$	$\theta_1 = 2.4$	$\theta_1 = 2.6$	$\theta_1 = 2.8$
$N = 2$	$T = 300$	0.06	0.46	0.70	0.80	0.89
	$T = 500$	0.04	0.48	0.75	0.87	0.93
	$T = 750$	0.05	0.56	0.81	0.93	0.97
	$T = 1000$	0.06	0.57	0.87	0.97	1.00
$N = 3$	$T = 300$	0.08	0.49	0.70	0.86	0.95
	$T = 500$	0.07	0.44	0.75	0.89	0.96
	$T = 750$	0.06	0.56	0.81	0.97	0.99
	$T = 1000$	0.06	0.65	0.94	0.99	1.00
$N = 5$	$T = 300$	0.04	0.42	0.70	0.86	0.95
	$T = 500$	0.04	0.50	0.82	0.94	0.99
	$T = 750$	0.05	0.60	0.95	1.00	1.00
	$T = 1000$	0.05	0.67	0.94	1.00	1.00
$N = 30$	$T = 300$	0.03	0.56	0.80	0.93	0.97
	$T = 500$	0.04	0.55	0.92	0.99	1.00
	$T = 750$	0.05	0.68	0.97	1.00	1.00
	$T = 1000$	0.06	0.78	0.99	1.00	1.00

Table 3.9 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (3.16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$ varying $\theta_1 \in \{2.0, 2.2, 2.4, 2.6, 2.8\}$ with $\theta_0 = 1$ in the DGP (3.16). In total, we conducted 301 Monte Carlo repetitions.

Table 3.10: Power with break at $\frac{T}{4}$

		$\theta_1 = 2.0$	$\theta_1 = 2.2$	$\theta_1 = 2.4$	$\theta_1 = 2.6$	$\theta_1 = 2.8$
$N = 2$	$T = 1000$	0.08	0.27	0.53	0.72	0.86
	$T = 2000$	0.08	0.41	0.76	0.94	0.98
$N = 3$	$T = 1000$	0.07	0.30	0.71	0.82	0.95
	$T = 2000$	0.08	0.45	0.84	0.96	0.99
$N = 5$	$T = 1000$	0.06	0.33	0.67	0.90	1.00
	$T = 2000$	0.07	0.51	0.89	0.99	1.00

Table 3.10 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (3.16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$ varying $\theta_1 \in \{2.0, 2.2, 2.4, 2.6, 2.8\}$ with $\theta_0 = 1$ in the DGP (3.16) where the break point ist constructed at $\frac{T}{4}$. In total, we conducted 301 Monte Carlo repetitions.

Table 3.11: Power analysis for varying Δ

$\Delta =$		Δ_0	$0.95 \cdot \Delta_0$	$0.9 \cdot \Delta_0$	$0.85 \cdot \Delta_0$	$0.80 \cdot \Delta_0$	$0.75 \cdot \Delta_0$
$N = 2$	$T = 300$	0.06	0.18	0.36	0.57	0.72	0.95
	$T = 500$	0.04	0.19	0.37	0.55	0.72	0.94
	$T = 750$	0.05	0.15	0.35	0.56	0.74	0.95
	$T = 1000$	0.06	0.22	0.43	0.64	0.80	0.98
$N = 3$	$T = 300$	0.08	0.17	0.37	0.57	0.77	0.95
	$T = 500$	0.07	0.15	0.32	0.54	0.73	0.96
	$T = 750$	0.06	0.20	0.41	0.64	0.84	0.97
	$T = 1000$	0.06	0.17	0.45	0.60	0.87	0.98
$N = 5$	$T = 300$	0.04	0.15	0.29	0.51	0.69	0.90
	$T = 500$	0.04	0.16	0.35	0.56	0.75	0.94
	$T = 750$	0.05	0.19	0.43	0.63	0.83	0.99
	$T = 1000$	0.05	0.17	0.43	0.73	0.92	1.00
$N = 30$	$T = 300$	0.03	0.18	0.32	0.47	0.65	0.88
	$T = 500$	0.04	0.15	0.34	0.55	0.76	0.95
	$T = 750$	0.05	0.15	0.42	0.66	0.84	0.98
	$T = 1000$	0.06	0.19	0.50	0.75	0.91	1.00

Table 3.11 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (3.16) using $B = 300$ bootstrap replications and 301 Monte Carlo repetitions. The third column Δ_0 depicts the size. For the Power Analysis (see column $4-8$) the copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$, while $\Delta = d \cdot \Delta_0$ with $d \in \{1, 0.95, 0.9, 0.85, 0.8, 0.75\}$.

3.5.2 Residual Data

In this subsection we consider residual data X_t from pre-estimated time series models for $t = 1, \ldots, T$. For our simulation we consider a GARCH(1,1) model, i.e.

$$\begin{aligned} r_{it} &= \sigma_{it} X_{it} \\ \sigma_{it}^2 &= \alpha_0 + \alpha_1 r_{i,t-1}^2 + \beta_1 \sigma_{i,t-1}^2 \end{aligned} \tag{3.17}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. To get serial correlated data we first simulate residual data using the factor copula model (3.16) with a break constructed at $\frac{T}{3}$ and $\theta_0 = 1$ and $\theta_1 = 2$. Then, we transform the residual data in serial correlated data r_{it} using the GARCH(1,1) model with fixed parameter values $\alpha_0 = \frac{1}{10}, \alpha_1 = \frac{1}{15}$ and $\beta_1 = \frac{1}{3}$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$.

With the simulated serial correlated data r_{it} we estimate the time series models using a GARCH(1,1) model and determine the residual data X_{it} for $i = 1, \dots, N$, which is used to perform the test. We vary the sample size $T = 1000, 2000, 4000$ and cross sectional dimension $N = 3, 5, 10$. The results can be seen in Table 3.12, which indicates that the test using residual data holds the size level.

Table 3.13 presents size results for the case of breaks in the GARCH parameters, where the GARCH residuals are calculated by means of the known GARCH parameters. Here, the coefficient β from (3.17) increases from 0.4 to 0.7 at $\frac{T}{2}$. Also in this case, the empirical size is close to the nominal size.

Finally, the power of our test in the case of GARCH residuals (with constant parameters) is examined in Table 3.14. The power is slightly lower than in the case without GARCH effects (Table 3.9), but also converges to 1 for increasing θ_1.

Table 3.12: Size using quantile version for GARCH-data

		$T = 1000$	$T = 2000$	$T = 4000$
$N = 2$	q_{95}	0.06	0.07	0.06
	q_{90}	0.15	0.14	0.12
$N = 3$	q_{95}	0.05	0.07	0.06
	q_{90}	0.12	0.13	0.13
$N = 5$	q_{95}	0.05	0.06	0.06
	q_{90}	0.11	0.13	0.11
$N = 10$	q_{95}	0.06	0.05	0.05
	q_{90}	0.11	0.10	0.10

Table 3.12 reports the rejection rate of the relevant change test where residual data from pre-estimated GARCH(1,1) models is considered. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \dots, 1)'$. In total, we conducted $B = 300$ bootstrap replications and 701 Monte Carlo repetitions.

3.6 Application

In this section, we apply the quantile based test to a multivariate data set of cross-sectional dimension $N = 30$. First, we apply a $GARCH(1,1)$ filter to the daily aggregated stock log-returns over a time span ranging from January

Table 3.13: Size using quantile version for GARCH-data with breaks

		$T = 1000$	$T = 2000$	$T = 4000$
$N = 2$	q_{95}	0.04	0.06	0.05
	q_{90}	0.09	0.09	0.11
$N = 3$	q_{95}	0.03	0.05	0.05
	q_{90}	0.08	0.11	0.09
$N = 5$	q_{95}	0.03	0.06	0.05
	q_{90}	0.08	0.12	0.11
$N = 10$	q_{95}	0.04	0.06	0.07
	q_{90}	0.13	0.14	0.13

Table 3.13 reports the rejection rate of the relevant change test where residual data from pre-estimated GARCH(1,1) models is considered. The coefficient of β_1 increases from 0.4 to 0.7 for each time series at $\frac{T}{2}$. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$. In total, we conducted $B = 300$ bootstrap replications and 701 Monte Carlo repetitions.

2003 to July 2015 from the German DAX30, implying $T = 3200$ and $N = 30$[5]. Second, we estimate a possible break point location in our $GARCH(1,1)$ adjusted data set, using (3.15), with the quantile $\mathbf{q} = 0.6\cdot(1, \ldots, 1)'$. This gives $\lfloor \hat{s}T \rfloor = 1884$ (15.02.2011), cf. the black dotted line in the middle in Figure 1. The first estimated break point in February 2011 now divides the data set into two parts (Dec. 04 - Feb. 11 and Feb. 11 - Jul. 15). As the test indeed indicates a break (see below), we repeat the breakpoint estimation in each part and obtain $\lfloor \hat{s}T \rfloor = 676$ (17.05.2006) for the first part and $\lfloor \hat{s}T \rfloor = 2653$ (26.02.2014) for the second part, respectively, which are both represented by the black dotted line in Figure 1. We do not search further for any breaks, as it seems to be unlikely that there are more than three change points (see Manner et al., 2019).

In the next step, we calculate (for each interval with one estimated breakpoint inside) $\Delta^{smallest}$ for each estimated break point as the smallest Δ for which the null hypothesis of no relevant change cannot be rejected, i.e., $|C_1(\mathbf{0.6}) - C_2(\mathbf{0.6})| \leq \Delta^{smallest}$. The number of bootstrap replications is 300. For each estimated break point, we also calculate the difference of the two resulting

[5]We adjusted the estimate for 5% of their outliers by setting these values equal to the expected value.

Table 3.14: Power analysis varying θ using GARCH-residuals

		$\theta_1 = 2.0$	$\theta_1 = 2.2$	$\theta_1 = 2.4$	$\theta_1 = 2.6$	$\theta_1 = 2.8$
$N = 2$	$T = 1000$	0.05	0.46	0.73	0.90	0.99
	$T = 2000$	0.07	0.47	0.82	0.98	1.00
$N = 3$	$T = 1000$	0.05	0.46	0.76	0.94	0.99
	$T = 2000$	0.06	0.50	0.93	1.00	1.00
$N = 5$	$T = 1000$	0.04	0.47	0.82	0.96	1.00
	$T = 2000$	0.05	0.64	0.97	1.00	1.00

Table 3.14 reports the rejection rate of the quantile relevant change test for data generated with the DGP described in (3.16) using $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$ varying $\theta_1 \in \{2.0, 2.2, 2.4, 2.6, 2.8\}$ with $\theta_0 = 1$ in the DGP (3.16) using GARCH-Residuals. In total, we conducted 301 Monte Carlo repetitions.

empirical copulas δ for $\mathbf{u} = 0.6 \cdot (1, ..., 1)'$, i.e. $\delta := |\hat{C}_1(\mathbf{0.6}) - \hat{C}_2(\mathbf{0.6})|$ and the change of the pairwise averaged Spearman's rhos before and after the estimated break point. The results of $\Delta^{smallest}$, δ and the change of the pairwise averaged Spearman's rhos can be found in Table 3.16. Table 3.15 provides the estimated Spearman's rho using the initial dataset for the breakpoint in Feb. 11 and the resulting subdatasets for the change in Spearman's rho in May 06 and Feb. 14.

Table 3.15: Empirical values

	Dec. 04 - May 06	May 06 - Feb. 11	Feb.11 - Feb. 14	Feb. 14 - Jul. 15	Dec. 04 - Feb. 11	Feb.11 - Jul. 15
$\rho_{Spearman}$	0.3469	0.4216	0.4677	0.5237	0.3998	0.4841

Table 3.15 reports the mean of Spearman's ρ in the corresponding intervals of the $GARCH(1,1)$ adjusted log returns.

Given $\mathbf{q} = 0.6 \cdot (1, \ldots, 1)'$, one possible reason for the first estimated break point in Feb. 2011 can be the beginning of the Euro crisis, i.e. the period in which considerable peaks of several Euro government bond yields were observed. Both, Delta and Spearman's rho are at their highest values here, which suggests that this crisis is having the strongest impact. It is well known that dependencies increase in times of crisis. The fact that Spearman's rho is rising

Table 3.16: Empirical break points

	May 06	Feb. 11	Feb. 14
$\Delta^{smallest}$	0.0098	0.0293	0.0273
δ	0.0292	0.0316	0.0387
$\rho^{diff}_{Spearman}$	0.0747	0.0843	0.0560

Table 3.16 reports the smallest Delta $\Delta^{smallest}$ such that the null hypothesis of no relevant break cannot be rejected. It also provides the empirical copula difference before and after the break point. The last row depicts the change in Spearman's rho before and after the estimated break points.

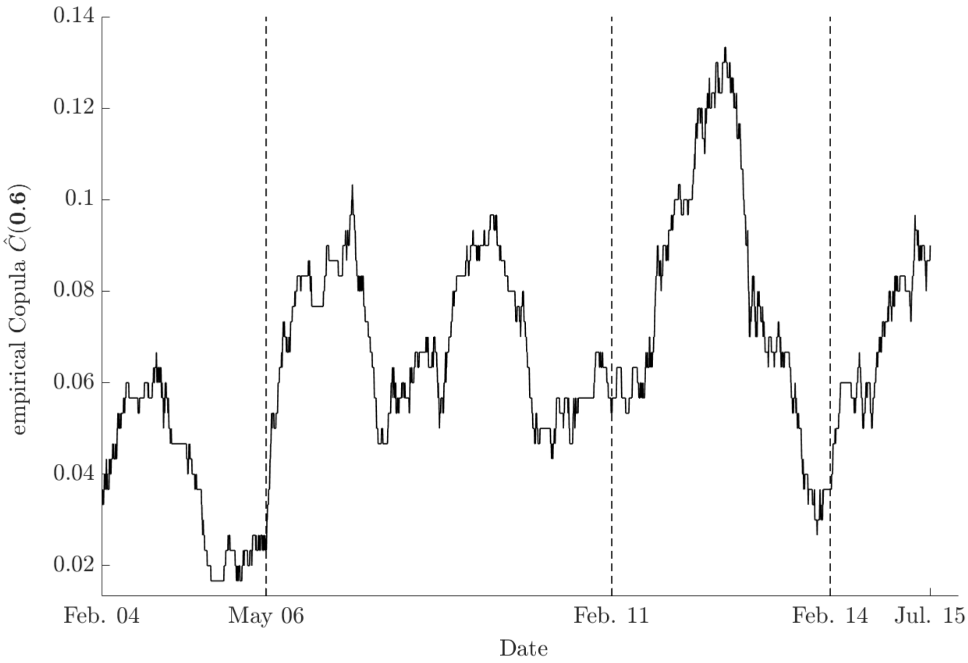
strongly and positively after the break point in Feb. 11 is therefore an indication. For the break point in May 06, $\Delta^{smallest}$ is equal to 0.0098, while the difference in Spearman's rho from the period Dec. 04 - May 06 and May 06 - Feb. 11 is equal to 0.0747. Analogously, for the estimated break point in Feb. 14, $\Delta^{smallest}$ is equal to 0.0273, while the change in Spearman's rho is equal to 0.0560. To sum up, if Δ is chosen to be the smallest value for which the null hypothesis of no relevant change cannot be rejected, the testing procedure provides a formula to determine Δ biuniquely. In addition, we observe that large values of $\Delta^{smallest}$ are related to large values of Spearman's rho. This means, the test can not only be used to test for relevant changes in the copula, but also as a selection tool to assess the effects of breaks.

3.7 Conclusion

In summary, the classical break point testing framework has two severe issues: On the one hand it considers a null which is theoretically never fulfilled and on the other hand any consistent test detects any arbitrary small change if the sample size is sufficiently large. Relevant change point analysis offers a way out.

We propose a new non-parametric test for detecting relevant breaks in copula functions, where the hypothesis is of the form $H_0 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\| \leq \Delta$ versus $H_1 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\| > \Delta$ with Δ a positive adjustable size to allow for difference in the copulas C_1 and C_2. Here, the norm in the hypothesis represents two different approaches: Either it measures the distance of the

Figure 3.1: Value of the empirical copula at $\mathbf{q} = 0.6 \cdot (1, \dots, 1)$



Value of the empirical copula defined in (3.14) evaluated at $\mathbf{q} = 0.6 \cdot (1, \dots, 1)$, computed in a rolling window of size 300. The estimated breakpoint, using (3.15), is displayed with the vertical black dotted line (15.02.2011) in the middle. The remaining two outer dotted lines represent estimated breakpoints of the resulting subdatasets on 17. May 2006 and 26. February 2014, respectively. Observed data between January 2003 and July 2015, implying $T = 3200$ and $N = 30$.

copulas given a certain value $\mathbf{q}$ or it equals the L^2-norm.

As a starting point, we consider a natural CUSUM-type test statistic fitting to the underlined testing problem. For the estimation of the limiting distribution, we construct a new non-parametric bootstrap based on natural estimates of the constructed testing process, which is applicable in the case of unknown sequentially estimated marginal distributions.

In the case where the copula distance is measured at a given value $\mathbf{q}$, we consider simulated data up to cross sectional dimension $N = 30$. For the L^2-norm, we investigate the behavior of our test up to $N = 5$. The Monte Carlo simulations show considerable size and power properties for both serially independent and residual data.

In our empirical application we analyze German DAX30 data of cross sectional dimension $N = 30$ between January 2003 and July 2015. Here, Δ is interpreted as the smallest admissible copula difference for which the relevant change hypothesis cannot be rejected. Every other choice of Δ that is smaller leads to a rejection of the null hypothesis. Cutting the empirical data into three parts leads to a detection of the very start of the financial crisis in 2006, the start of the Euro crisis in 2011 and to a break in 2014 given that the quantile $\mathbf{q}$ is chosen to be equal to $0.6 \cdot (1, \dots, 1)'$.

In conclusion, Δ can be regarded not only as the upper bound of an admissible copula distance, but also as a measure of the extent of a crisis.

CHAPTER 4

Specification Testing in Functional Quantile Regression Models

4.1 Abstract

We propose a novel consistent specification test for quantile regression models where we allow the covariate effects to be quantile dependent and nonlinear. To achieve this, we parameterize the conditional quantile functions by appropriate basis functions, rather than parametrically and hence allowing to test for functional forms beyond linearity while retaining the linear cases as special cases. Due to the dependence on the quantile itself covariate-quantile relations can differ for distinct quantiles. The induced class of conditional distribution functions can finally be tested with a Cramér-von Mises type test statistic. We derive the theoretical limit distribution and propose a practical bootstrap method. To increase the power of our test, we suggest a modified test statistic using quantile regression splines. A detailed Monte Carlo experiment shows that the test results in a reasonable sized testing procedure with large power. An application to conditional income disparities between East and West Germany over the period $2001-2010$ indicates that there are still significant differences across the quantiles of the conditional income distributions, when conditioning on age.

4.2 Introduction

Hypothesis testing plays a central role in many economic research areas. A necessary prerequisite for the statistical validity of the decisions to be made is the correct specification of the underlying model. Specification tests can be used to validate the correctness of theoretical assumptions. Within the framework of linear regression, a whole range of specification tests are available that can test both, parametric and non-parametric approaches. In general, testing misspecification in linear OLS models is well understood and developed. For the parametric setup, e.g., Bierens (1990) showed that any conditional moment test of functional form of nonlinear regression models can be converted into a consistent chi-square test that is consistent against all deviations from the null hypothesis. Härdle & Mammen (1993) suggested a wild bootstrap procedure for regression fits in order to decide whether a parametric model could be justified. Stute (1997) proposed a general method for testing the goodness of fit of a parametric regression model. For the nonparametric case, among others, Gozalo (1993) proposed a general framework for specification testing of the regression function in a nonparametric smoothing estimation context and Stute et al. (1998) suggested a goodness of fit test using a wild bootstrap procedure that checks whether a function belongs to a certain model class.

However, OLS estimates are sensitive to outliers and draw only a part of the whole picture, since they only model the conditional expected value. As it provides more robust estimates compared to OLS and allows a more comprehensive picture and flexible analysis of the economic problem, quantile regression has become increasingly popular since the seminal article by Koenker & Bassett Jr (1978). But it also applies to quantile regressions, that post-estimation inference procedures essentially depend on the validity of the underlying parametric functional form for the quantiles considered (Angrist et al., 2006). For example, assuming the same fixed linear relationship between covariates for all quantiles is the connecting element of the Machado-Mata decomposition (used in particular to describe wage inequalities) by Machado & Mata (2005) and the Khmaladazation (which is based on the Doob-Meyer decomposition of the martingale) by Koenker & Xiao (2002). Since such a linearity assumption considerably limits the number of possible models and hence the hypothesis space, there have recently been successful attempts to weaken the linearity assumption for quantile estimation and inference with independently and identically distributed (i.i.d.) data.

In this context, more general parametric quantile models have been developed that, among others, include works by Hallin et al. (2009) suggesting an estimator for local linear spatial quantile regression and Guerre & Sabbah (2012) investigating the Bahadur representation of a local polynomial estimator of the conditional quantile function and its derivatives. But also nonparametric approaches for estimating conditional quantile functions have attracted much attention. Li & Racine (2008) proposed a nonparametric conditional cumulative distribution function kernel estimator along with an associated nonparametric conditional quantile estimator. Belloni et al. (2019) developed a nonparametric quantile regression-series framework for performing inference on the entire conditional quantile function and its linear functionals and Qu & Yoon (2015) presented estimators for nonparametrically specified condtional quantile processes that are based on local linear regressions. Li et al. (2020) investigated the problem of nonparametrically estimating a conditional quantile function with mixed discrete and continuous covariates suggesting a kernel based approach. But regardless of whether parametric or non-parametric approaches are chosen, the theory concerning the validity of the correct model choice seems to keep up with the rapid development of new estimation methods only to a limited extent. To the best of our knowledge, there does not exist a testing procedure that allows for quantile-specific functional covariate effects.

In a parametric framework, one of the first specification tests for linear location shift and location-scale shift quantile models with i.i.d. data is the test by Koenker & Xiao (2002). Shortly after that, Chernozhukov (2002) proposes a resampling testing procedure avoiding to estimate further objects, such as the score function using the same principle as Koenker & Xiao (2002). However, these two tests proposed do not test the validity of the quantile regression model itself. Escanciano et al. (2010) and Escanciano & Velasco (2010) both tested the validity of the null hypothesis that a conditional quantile restriction is valid over a range of quantiles. Rothe & Wied (2013) proposed a

specification test for a larger class of models, including quantile regression models. In case of nonparametric instrumental quantile regression, Breunig (2019) develops a methodology for testing the hypothesis whether the instrumental quantile regression model is correctly specified. However, all models have in common that they require linearity in the regressors. In this paper, we consider a broad approach that tackles the following two challenges simultaneously, hence proposing a general specification test that is an important contribution in the field with potential in a wide range of applied questions. *i*) We suggest a testing procedure for quantile regression models, where the regressors can explicitly depend on the quantile considered, which allows to test for the correct specification of large number of models. *ii*) Due to our general model set up, our proposed methodology does also allow to test for semi-parametric models, e.g., B-splines for quantile regressions, where the covariates have a functional form (cf. Cardot et al. (2005) for the estimation procedure of such processes). Such a testing procedure not only increases the range of applications but also offers the advantage that effects can be tested in isolation, depending on the quantile. As such, it extends the literature on quantile regression specification tests for better answering relevant questions in economics and further sciences; wherever specific regressors have a functional, non-linear influence on distinct quantiles may be present.
To illustrate the power and potential of our test, we consider the case of income inequality, with a focus on differences in the conditional income quantiles between East and West Germany in a balanced panel data set. Such disparities have received considerable attention in the economic literature (e.g. Biewen, 2000), and also consistently played a major role in the domestic political debate. Our empirical analysis uses the German socio-economic panel (SOEP) and shows that age has a predominant linear influence on income development in Germany, but for the upper 90% quantile the influence of age is solely quadratic. Such statistically proven statements on income distributions also provide further evidence for investigations on inequalities in income distributions. Importantly, and in line with other studies on this topic, we find through an initial Machado-Mata (Machado & Mata, 2005) decomposition that there are still income differences between East and West Germany, which can be confirmed by our proposed testing procedure.

But also further use cases in, e.g., finance can be addressed with our testing approach. For instance, the correct specification of the left tail of the distribution function is essential to adequately assess risks. Our null hypothesis can be specified in a way that only some parts of the quantile function follow an explicit parametric approach. Thus, we provide a statistical verification procedure for the part of the distribution function that is of relevance for the calculation of the value at risk. Previous procedures usually require a complete characterization of the quantile function, in which the covariates must also be independent of the considered quantiles.

The basic idea of our procedure is based on the principle characterized by Rothe & Wied (2013): We compare an unrestricted estimate of the joint distribution function of the random variable Y and the random vector X with a restricted estimate that imposes the structure implied by the null hypothesis model. Based on a Cramèr-von Mises type measure of distances, the restricted estimate of the joint distribution can then be compared with the unrestricted one. We derive the non-pivotal limiting distribution of our test statistic and show the validity of our suggested parametric bootstrap procedure for the approximation of the critical values. To increase the power of our test, we replace the unrestricted model estimate with a quadratic B-spline, meeting the assumptions of a quantile function. Due to the generality of our test procedure we can subsume previous specification tests for quantile regression models with i.i.d. data as marginal cases of our procedure. The Monte Carlo simulation study shows that the proposed testing procedure has superior power properties than existing methods.

In sum, we believe that the testing procedure proposed in our paper is a useful extension of existing methods for testing the correct specification in quantile regression models, both in terms of the improvement in power, and also in terms of the extension to quantile dependent regressors it offers.

The paper is organized as follows. Section 2 formulates the test problem and derives the test statistics including the parametrized bootstrap approximation. In Section 3, we provide the theoretical properties of the testing procedure and the bootstrap. Section 4 contains an intensive Monte Carlo simulation including comparisons to existing tests and in Section 5 we present the empirical application. The last section concludes. In order to increase the readability of the paper, all proofs are to be found in the Appendix.

4.3 Quantile regression testing

4.3.1 Specification test for quantile dependent regressors

We observe an outcome variable $Y_i \in \mathbb{R}$ and a vector of explanatory variables $X_i \in \mathbb{R}^K$ for $i = 1, ..., n$, $K \in \mathbb{N}$. We assume the data points to be independent and identically distributed (i.i.d.). Our aim is to test the validity of certain classes of parametric specifications for the conditional cumulative distribution function (cdf) F of Y given X, i.e. $F_{Y|X}$ and with corresponding conditional quantile function (qf) $F^{-1}_{Y|X}$. Since for $y \in \mathbb{R}$ in holds that

$$F_{Y|X}(y \,|\, x) = \int_0^1 \mathbb{1}_{\left\{F^{-1}_{Y|X}(\tau|x) \leq y\right\}} d\tau, \tag{4.1}$$

let $\mathcal{F}$ be the set of conditional cdfs $F_{Y|X}$ induced by $F^{-1}_{Y|X}$, i.e.

$$\begin{aligned} \mathcal{F} := \{F_{Y|X}(y|x, \theta) \mid F^{-1}_{Y|X}(\tau \mid x) = P(x, \tau)'\theta \\ \text{for some } \theta \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and } (y, x) \in \mathcal{S}\}, \end{aligned} \tag{4.2}$$

where $\mathcal{S}$ denotes the support of $(y, x) \in \mathbb{R}^{K+1}$ and $\mathcal{B}(\mathcal{T}, \Theta)$ the class of functions $\tau \mapsto \theta(\tau) \in \Theta \subset \mathbb{R}^p$ for $\tau \in \mathcal{T} \subseteq [0, 1]$ with p the dimension of the parameters. These conditional qfs $F^{-1}_{Y|X}$ are assumed to be of the form $P(X, \tau)'\theta(\tau)$, i.e. $F^{-1}_{Y|X}(\tau|X) = P(X, \tau)'\theta$ for every $\tau \in \mathcal{T}$ and $\theta \in \mathcal{B}(\mathcal{T}, \Theta)$. $P(X, \tau)$ is a vector of transformations of the realization X (also known as basis function evaluations in the literature) depending on $\tau \in \mathcal{T}$ such as polynomials or B-splines (which are also called basis functions) evaluated at X that depend on τ and thus may differ for distinct quantiles. In the simplest case, e.g., $P(X, \tau)$ can be equal to X for all $\tau \in \mathcal{T}$, such that the quantiles are linear in X. In this case, the parameterization of the quantile function corresponds to the classical linear quantile regression model, where the vector of transformations $P(X, \tau)$ is constant for all quantiles τ.

The hypothesis we want to test is that the conditional cdf $F_{Y|X}$ coincides with an element of a class of distributions $\mathcal{F}$ of the form (4.2) which corresponding $F^{-1}_{Y|X}$ can be decomposed to a vector of transformations and a functional parameter $\theta \in \mathcal{B}(\mathcal{T}, \Theta)$:

$$H_0 : F_{Y|X} \in \mathcal{F} \quad \text{vs.} \quad H_1 : F_{Y|X} \notin \mathcal{F}. \tag{4.3}$$

Again, in comparison to existing parametric quantile regression models we explicitly allow the vector of transformations $P(X, \tau)$ to depend on τ. Hence, this framework allows to model a quantile function that, e.g., contains a linear regressor in the lower fifty percent quantile and a highly non-linear functional regression form in the upper fifty percent quantile. Naturally, models in which the vector of transformations does not depend on τ are captured by our approach as a special case, when $P(X, \tau) \equiv P(X) = \mathit{const}$ for all $\tau \in \mathcal{T}$. Consequently, the aim of this article is to present a testing procedure that is able to give statistical insights if the quantile model assumption (4.3) holds statistically true. We assume that there is a unique $\theta_0 \in \mathcal{B}(\mathcal{T}, \Theta)$ under the null hypothesis. Accordingly, we can reformulate our testing problem (4.3) to

$$\begin{aligned}
H_0 &: F_{Y|X}(y \mid x) \in \mathcal{F}^0 := \{F_{Y|X}(y \mid x, \theta_0) \mid F^{-1}_{Y|X}(\tau \mid x) = P(x, \tau)'\theta_0 \\
&\qquad \text{for some } \theta_0 \in \mathcal{B}(\mathcal{T}, \Theta), \text{ for all } (y, x) \in \mathcal{S}\} \\
H_1 &: P\left(F_{Y|X}(y^* \mid x^*) \neq F_{Y|X}(y^* \mid x^*, \theta)\right) > 0 \\
&\qquad \text{for all } \theta \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and for some } (y^*, x^*) \in \mathcal{S}.
\end{aligned} \tag{4.4}$$

In addition, we assume that under the null hypothesis any functional parameter $\theta \in \mathcal{B}(\mathcal{T}, \Theta)$ satisfying $F_{Y|X}(y|x) = F^*_{Y|X}(y|x, \theta)$ with $F^*_{Y|X}(y|x, \theta) \in \mathcal{F}^0$ for all $(y, x) \in \mathcal{S}$ also satisfies $\theta(\tau) = \theta_0(\tau)$ for all $\tau \in \mathcal{T}$.
To propose a testing procedure for the problem (4.4) we assume that the true value of the functional parameter, i.e. $\theta_0(\tau)$ for every $\tau \in \mathcal{T}$, is identified under the null hypothesis through a moment condition. Specifically, let

$$g : \mathcal{S} \times \Theta \times \mathcal{T} \to \mathbb{R}^p$$

be a uniformly integrable functions whose exact form depends on $\mathcal{F}^0$, and suppose that for every $\tau \in \mathcal{T}$, the equation[1]

$$G(\theta, \tau) := \mathbb{E}[g(Y, X, \theta, \tau)] = 0 \in \mathbb{R}^p \tag{4.5}$$

has a unique solution $\theta_0(\tau)$. Furthermore, under the alternative, $\theta_0(\tau)$ remains well defined for all $\tau \in \mathcal{T}$ due to Assumption (4.5) and can thus be thought of as a pseudo-true value of the functional parameter in this case.
The null hypothesis can be equivalently stated as

$$F_{Y|X}(y|x) = F_{Y|X}(y|x, \theta_0) \text{ for all } (y, x) \in \mathbb{R}^{K+1}, \tag{4.6}$$

with $\theta_0(\tau)$ as the unique solution to (4.5) for all $\tau \in \mathcal{T}$. This holds true since $\mathcal{F}^0$ is a singleton containing $F_{\cdot|\cdot}(\cdot|\cdot, \theta_0)$. Since $F_{Y|X}(y|X) = \mathbb{E}[\mathbb{1}_{\{Y \le y\}}|X]$ we can write the joint cdf of Y and X, $F_{Y,X}$, as[2]

$$F(y, x) = \int_{\mathbb{R}^K} F_{Y|X}(y \mid x^*) \mathbb{1}_{\{x^* \le x\}} dF_X(x^*) \tag{4.7}$$

$$F(y, x, \theta_0) = \int_{\mathbb{R}^K} F_{Y|X}(y \mid x^*, \theta_0) \mathbb{1}_{\{x^* \le x\}} dF_X(x^*), \tag{4.8}$$

where F_X denotes the marginal cdf of X. From Billingsley (1995) Theorem 16.10 (iii) it follows that the testing problem (4.4) can be restated as

$$\begin{aligned} &H_0 : F(y, x) = F(y, x, \theta_0) \text{ for some } \theta_0 \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and for all } (y, x) \in \mathcal{S} \\ &\qquad\qquad\qquad\qquad\text{versus} \\ &H_1 : F(y, x) \neq F(y, x, \theta) \text{ for all } \theta \in \mathcal{B}(\mathcal{T}, \Theta) \text{ and for some } (y, x) \in \mathcal{S}. \end{aligned} \tag{4.9}$$

for some $\theta_0 \in \mathcal{B}(\mathcal{T}, \Theta)$. With the help of the above representation of the null hypothesis (4.9) we introduce a function $S : \mathbb{R}^{K+1} \times \Theta \to \mathbb{R}$ that measures the difference of the non-parametric $F(y, x)$ and the parametrized cdf $F(y, x, \theta)$ defined as

$$S(y, x, \theta) := F(y, x) - F(y, x, \theta). \tag{4.10}$$

[1] The representation of the moment function g is given by
$g(Y, X, \theta, \tau) := \left(\tau - \mathbb{1}_{\{Y \le P(X,\tau)'\theta(\tau)\}}\right) P(X, \tau)'$.

[2] Due to readability we will suppress the index Y, X for the joint cdf $F_{Y,X}$ in the following, i.e. $F = F_{Y,X}$.

The null hypothesis is true by assumption if $S(y, x, \theta_0) = 0$ for all $(y, x) \in \mathcal{S}$, whereas $S(y, x, \theta) \neq 0$ for all $\theta \neq \theta_0 \in \mathcal{B}(\mathcal{T}, \Theta)$ and for some $(y, x) \in \mathcal{S}$. To obtain an applicable test statistic we will replace $F(y, x)$ and $F(y, x, \theta_0)$ by its empirical counterparts $\hat{F}_n(y, x)$ and $\hat{F}_n(y, x, \theta)$. Thus, we have

$$S_n(y, x, \hat{\theta}) := \hat{F}_n(y, x) - \hat{F}_n(y, x, \theta), \tag{4.11}$$

with $\hat{F}_n(y, x, \theta) = F(y, x, \hat{\theta}_n)$, a parametric estimate of F based on a consistent estimate $\hat{\theta}_n$ of θ_0. In order to emphasize that the parametric empirical cdf $\hat{F}_n(y, x, \theta)$ particularly estimates the parameter θ_0 by $\hat{\theta}_n$, we also use the notation $\hat{F}_n(y, x, \hat{\theta}_n)$. Under the null hypothesis, $\hat{F}_n(y, x)$ and $\hat{F}_n(y, x, \hat{\theta}_n)$ are consistent estimators for $F(y, x)$ and $F(y, x, \theta_0)$, respectively. In that case, $S_n(y, x, \theta)$ should be close to zero for all $(y, x) \in \mathcal{S}$. If, however, the alternative holds true, then there is a vector (y, x) for each $\theta \in \mathcal{B}(\mathcal{T}, \Theta)$ such that the function S_n is greater than zero.

To obtain an estimate for the parametrized empirical cdf $\hat{F}_n(y, x, \hat{\theta}_n)$ we follow Chernozhukov et al. (2013) and take $\hat{\theta}_n$ to be an approximate Z-estimator satisfying

$$||\hat{G}(\hat{\theta}_n, \tau)|| = \inf_{\theta \in \Theta} ||\hat{G}(\theta, \tau)|| + \eta_n \tag{4.12}$$

where the function $\hat{G}(\hat{\theta}_n, \tau) := n^{-1} \sum_{i=1}^{n} g(Y_i, X_i, \theta, \tau)$ is the sample analogue of the moment condition (4.5) for every $\tau \in \mathcal{T}$ and for some possibly random variable $\eta_n = o_p(n^{-1/2})$. For every $\tau \in \mathcal{T}$ and every $(y, x) \in \mathbb{R}^{K+1}$, the estimator based on the testing problem (4.4) takes the form

$$\hat{F}_n(y|x, \hat{\theta}_n) = \delta + \int_{\delta}^{1-\delta} \mathbb{1}_{\{P(x,\tau)'\hat{\theta}_n(\tau) \leq y\}} d\tau, \qquad (y, x) \in \mathcal{S}, \tag{4.13}$$

$$\hat{\theta}_n(\tau) = \underset{\theta \in \Theta}{\operatorname{argmin}} \sum_{(y,x) \in \mathcal{S}} \left(\tau - \mathbb{1}_{\{y \leq P(x,\tau)'\theta\}}\right) (y - P(x, \tau)'\theta) \tag{4.14}$$

for some arbitrary constant $\delta > 0$. The trimming by δ avoids estimation of tail quantiles (Koenker, 2005) and is valid under the conditions in Theorem 4.4.1 in Section 4.4. Thus, the test statistic (4.11), that is based on the differences of the non-parametric and parametrized empirical distribution functions, can be expressed as

$$
\begin{aligned}
S_n(y,x,\hat{\theta}_n) &= \hat{F}_n(y,x) - \hat{F}_n(y,x,\hat{\theta}_n) \qquad (4.15)\\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{1}_{\{Y_i\le y\}}\mathbb{1}_{\{X_i\le x\}}\right)\\
&\quad - \int_{\mathbb{R}^K}\mathbb{1}_{\{x^*\le x\}}\left(\delta + \int_{\delta}^{1-\delta}\mathbb{1}_{\{P(x^*,\tau)'\hat{\theta}_n(\tau)\le y\}}d\tau\right)\, d\hat{F}_X(x^*)\\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\mathbb{1}_{\{Y_i\le y\}}\mathbb{1}_{\{X_i\le x\}} - \mathbb{1}_{\{X_i\le x\}}\left[\delta + \int_{\delta}^{1-\delta}\mathbb{1}_{\{P(X_i,\tau)'\hat{\theta}_n(\tau)\le y\}}d\tau\right]\right).
\end{aligned}
$$

We propose a Cramér-von Mises type test statistic S_n^{CM} defined as

$$
S_n^{CM} := \int ||\sqrt{n}S_n(y,x,\hat{\theta}_n)||^2 d\hat{F}_n(y,x), \qquad (4.16)
$$

which is a generalization of existing quantile regression tests. However, if the vector of transformations $P(X,\tau)$ in (4.4) is independent of τ, then the test statistic coincides with test statistic proposed in Rothe & Wied (2013). Notwithstanding the above, it is also possible to consider a Kolmogorov-Smirnov-type test statistic

$$
S_n^{KS} := \sqrt{n}\sup_{(y,x)\in\mathcal{S}} ||S_n(y,x,\hat{\theta}_n)||, \qquad (4.17)
$$

but the Cramér-von-Mises-type test provides better (power) results, since it is less susceptible to outliers (Chernozhukov, 2002; Rothe & Wied, 2013).

4.3.2 More powerful testing procedure using splines

In order to obtain better power results, we consider two different test statistics of the form (4.11), using two estimators for the quantile regression model specified under the null hypothesis: one estimator corresponds to the model, the other employs a spline approach, i.e.

$$
\begin{aligned}
S_n^{*}(y,x,\hat{\theta}_n) &= \hat{F}(y,x) - \hat{F}_n^{S}(y,x,\hat{\theta}_n) - \left(\hat{F}(y,x) - \hat{F}_n(y,x,\hat{\theta}_n)\right) \qquad (4.18)\\
&= \hat{F}_n^{S}(y,x,\hat{\theta}_n) - \hat{F}_n(y,x,\hat{\theta}_n) \qquad (4.19)
\end{aligned}
$$

where $\hat{F}_n^S$ is the estimate of the cumulative distribution function by a quantile regression spline that meets some regularity assumptions (cf. Assumptions 3 in Section 4.4 and Cardot et al. (2005)) and $\hat{F}_n(y, x, \hat{\theta}_n)$ the estimate using the null hypothesis model. In case of non-varying covariates, i.e. $P(X, \tau) = P(X)$ for all $\tau \in \mathcal{T}$, $\hat{F}_n^S$ could be, e.g., estimated by a quadratic B-spline with monotone increasing parameters.

Besides standard assumptions, the monotonicity assumption is of central importance for the estimation of the quantile regression function by splines[3]. However, the Monte Carlo simulation clearly shows that these additional assumptions significantly increase the rejection rates in the case of misspecified null hypotheses. Xue & Wang (2010) have shown, e.g., that the estimate of the cumulative distribution function with a smooth monotone polynomial spline has better finite sample properties than the empirical distributional estimate. However, the goodness and convergence rate of the spline approximation depends, in general, in a complex fashion on the degree of the spline, the number of knots and the position of those knots. He & Shi (1997) have pointed out that if the number of knots $k_n \sim (n/\log n)^{2/5}$ and under some mild assumptions,[4] the order of approximation of a quadratic monotone B-spline is $(\log n/n)^{2/5}$ for a quantile regression model with non-varying covariates. Cardot et al. (2005) have generalized the limiting result for quantile regression models with varying covariates. This result is of particular interest since, together with the Donsker-class property[5], it provides the basis for the convergence of the Cramer-von Misès type test statistic that is defined as

$$S_n^{CM^*} := \int ||\sqrt{n} S_n^*(y, x, \hat{\theta}_n)||^2 d\hat{F}_n(y, x). \tag{4.20}$$

[3]There is a whole series of assumptions that guarantee the monotonicity of the quantile function, e.g., derivatives of the (quadratic) spline to be non-negative or estimators are monotonically increasing as quantiles increase assuming static covariates, all having different computational properties. To discuss all these assumptions is beyond the scope of this project. We refer here to the relevant literature, i.e. Koenker et al. (1994), Bondell et al. (2010) among others.

[4]For a detailed description of the requirements we refer to the assumptions $C1 - C3$ from He & Shi (1997).

[5]Yu et al. (2017) have shown the Donker-class property for functional linear partial quantile regressions.

4.3.3 Semiparamteric bootstrap procedure

As we show in more detail in the next section, the asymptotic null distribution of S_n^{CM} and S_n^{CM*}, respectively, depend on the data generating process in a complex fashion. To obtain critical values for our test, we therefore propose a semiparametric bootstrap procedure. This procedure is reasonable from a practical point of view, since it avoids the complicated problem of estimating the null distribution directly, including the complex covariance structure. The idea of our semiparametric bootstrap is to generate synthetic data that is line with the assumptions under the null hypothesis. Thus, the bootstrap mimics the distribution of the data under the null hypothesis, even though the data might be generated by an alternative distribution. The procedure works as follows:

i.) Draw B bootstrap samples of covariates $\{X_{b,i}, 1 \leq i \leq n\}_{b=1,...,B}$ of size n with replacement from the realized values $\{X_i, 1 \leq i \leq n\}$.

ii.) Generate independently B n-dimensional vectors U_b with $b = 1, ..., B$ of standard uniform distributed random variables, i.e. $U_b = (U_{b,i})_{i=1}^n$ with $U_{b,i} \overset{i.i.d.}{\sim} U(0,1)$ for $i = 1, ..., n$ and $b = 1, \ldots, B$, that represent the randomly chosen quantiles.

iii.) For each $b = 1, ..., B$, estimate the conditional quantile function $\hat{F}^{-1}(U_{b,i} \mid X)$ for every $i = 1, ..., n$ by the model specified under the null hypothesis using the realized values X and compute n-dimensional estimates $\hat{Y}_b := (\hat{Y}_i)_{b,i=1}^n$ for $b = 1, ...n$ by means of the bootstrap sample of covariates $\{X_{b,i}, 1 \leq i \leq n\}_{b=1,...,B}$, i.e. $\hat{Y}_{b,i} = \hat{F}^{-1}(U_{b,i} \mid X_{b,i})$ for $i = 1, ...n$ and $b = 1, ..., B$.

iv.) Calculate B bootstrap versions of the test statistic (4.20), i.e. for $b = 1, ..., B$ compute

$$S_{n,b}^{CM} := \int ||\sqrt{n} S_{n,B}(\hat{y}, x_b, \hat{\theta}_n)||^2 d\hat{F}_n(\hat{y}, x_b). \tag{4.21}$$

v.) Determine the critical value c such that

$$\frac{1}{B}\sum_{b=1}^{B} \mathbb{1}_{\{S_{n,b}^{CM}>c\}} \overset{!}{=} q, \tag{4.22}$$

where $q \in (0,1)$.

With the above described bootstrap procedure we can calculate critical values $c(q)$ for (4.16). Critical values for (4.20) can be obtained in the same manner if the test statistic $S_{n,B}^{CM}$ from (4.21) is replaced by its spline counterpart, i.e. $S_{n,B}^{CM^*}$.

4.4 Asymptotics

4.4.1 Theoretical properties for quantile dependent regressors

This section shows that the test statistic S_n^{CM} has the correct asymptotic size which is summarized in Theorem 4.4.1 at the end of that subsection. Before we derive large sample properties of our test statistic (4.16), we need to impose and to discuss some mild assumptions that are in line with Chernozhukov et al. (2013). However, since our proposed test statistic is a generalization of existing tests we need to slightly adjust the standard assumptions. Additionally, we assume that there is a finite compact decomposition of $\mathcal{T} := [\varepsilon, 1-\varepsilon]$, $\varepsilon \in (0, 0.5)$. Hence, we can formulate the assumptions for Θ being an arbitrary subset of $\mathbb{R}^p$ as

Assumption 2.

i.) *$P(X,\tau)$ is L^2-bounded in $[0,1]$.*

ii.) *Let $\bigcup_{l=1}^{L} I_l = \mathcal{T}$, $L \in \mathbb{N}$, I_l compact for $l = 1,...,L$ and $I_{l_1} \cap I_{l_2}$ a singleton or the empty set for $l_1 \neq l_2$.*

iii.) *For each $\tau \in I_l$ with $l = 1,...,L$, $G(\cdot,\tau) : \Theta \to \mathbb{R}^p$ possesses a unique zero at $\theta_0(\tau) \in interior(\Theta)$ such that $G(\theta_0(\tau),\tau) = 0$, and, for some $\delta > 0$, $\mathcal{B} := \bigcup_{\tau\in\mathcal{I}_l} B_\delta(\theta_0(\tau))$ is a compact subset of $\mathbb{R}^p$ contained in Θ for $l = 1,...,L$.*

iv.) *Further, $G(\cdot,\tau)$ has a inverse $G^{-1}(x,\tau) := \{\theta \in \Theta \mid \mathbb{G}(\theta,\tau) = x\}$ that is continuous at $x = 0$ uniformly in $\tau \in I_l$ for all $l = 1, ..., L$ with respect to the Hausdorff distance.*

v.) *There is a derivative $\dot{G}_{\theta_0(\tau),\tau}$ such that*

$$\lim_{t\to 0} \sup_{\tau\in I_l, ||h||=1} \left|\frac{G(\theta_0(\tau)+th,\tau) - G(\theta_0(\tau),\tau)}{t} - \dot{G}_{\theta_0(\tau),\tau}h\right| = 0,$$

where $\dot{G}_{\theta_0(\tau),\tau}$ is non-singular at $\theta_0(\cdot)$ uniformly over $\tau \in I_l$ with $l = 1, ..., L$, i.e. $\inf_{\tau\in I_l} \inf_{||h||=1} ||\dot{G}_{\theta_0(\tau),\tau}h|| > 0$ for all $l = 1, ..., L$.

vi.) *The maps $\tau \mapsto \theta_0(\tau)$ and $\tau \mapsto \dot{G}_{\theta_0(\tau),\tau}$ are continuous on $\mathcal{T}$.*

vii.) *The function set $\mathcal{G}_l = \{g(Y, X, \theta, \tau) | (\theta, \tau) \in \Theta \times I_l)\}$ is F_{YX}-Donsker*[6] *for all $l = 1, ..., L$ with a square integrable envelope $\tilde{G}$ for $\bigcup_{l=1}^{L} \mathcal{G}_l$. The map $(\theta,\tau) \mapsto g(\cdot,\theta,\tau)$ is continuous at each $(\theta,\tau) \in \Theta \times I_l$ for all $l = 1, ..., L$.*

viii.) *The mapping $\theta \mapsto F(\cdot|\cdot,\theta)$ is Hadamard differentiable for all $\theta \in \Theta$ with derivative $h \mapsto \dot{F}(\cdot|\cdot,\theta)[h]$*

Assumption 2 *i.*) claiming there is a finite, compact decomposition of the unit interval is required since we consider Donsker classes in the proof. We are using the fact that the union of Donsker classes is also Donsker (see Dudley, 2014, section 3.8). Assumptions 2 *i.*) $-$ *v.*) guarantee the regularity of our estimator $\hat{\theta}_n$ and ensure that a functional central limit theorem can be applied to Z-estimator processes (cf. Corollary C.1.2 in the Appendix C.1). Assumption 2 *vi.*) is a smoothness condition, that implies together with the functional delta method that the restricted cdf estimator process

$$(y,x) \mapsto \sqrt{n}\left(\hat{F}_n(y,x,\hat{\theta}) - F(y,x,\theta)\right) \tag{4.23}$$

is F_{YX}-Donsker. The convergence (4.23) can be shown to be jointly with that

[6] Consider the empirical process $\mathbb{G}_n := \sqrt{n}(\mathbb{F}_n - F)$, where $\mathbb{F}_n$ is the empirical distribution function and F the theoretical cdf. If $\mathbb{G}_n$ converges weakly to a tight Borel measurable element in $\ell^\infty(\mathcal{F})$, then the class $\mathcal{F}$ for which this is true is called F-Donsker. $\ell^\infty(\mathcal{F})$ is the set of all uniformly bounded real functions from $\Omega \to \mathbb{R}$.

of the empirical cdf process

$$(y,x) \mapsto \sqrt{n}\left(\hat{F}_n(y,x) - F(y,x)\right) \tag{4.24}$$

to a Brownian Bridge by some standard arguments given in Lemma C.1.1. The limiting distribution of our test statistic S_n^{CM} then follows from an application of the continuous mapping theorem (the proof and further details are shifted to the Appendix C.1). We are now able to derive our main result:

Theorem 4.4.1. *If Assumptions* (2) *is satisfied, then the following statements hold:*

i.) Under the null hypothesis H_0 (4.9),

$$S_n^{CM} \xrightarrow{d} \int ||\mathbb{G}_1(y,x) - \mathbb{G}_2(y,x)||^2 dF_{YX}(y,x), \tag{4.25}$$

where $(\mathbb{G}_1, \mathbb{G}_2)$ *are Gaussian processes with zero mean and covariance function*

$$Cov[\mathbb{G}_1(y,x), \mathbb{G}_1(y',x')] = \sum_{k=-\infty}^{\infty} Cov[\mathbb{1}_{\{Y_0 \le y\}}\mathbb{1}_{\{X_0 \le x\}}, \mathbb{1}_{\{Y_k \le y'\}}\mathbb{1}_{\{X_k \le x'\}}]$$

$$\mathbb{G}_2(y,x) := \int \mathbb{G}_2^{+}(y,x^*)\mathbb{1}_{\{x^* \le x\}} dF_X(x^*) + \int F(y \,|\, x^*)\mathbb{1}_{\{x^* \le x\}} d\mathbb{G}_1(\infty, x^*)$$

with $\mathbb{G}_2^{+}(y,x)$ *the limiting Gaussian process of* $\sqrt{n}\left(\hat{F}_n(y \,|\, x, \hat{\theta}_n) - F(y \,|\, x)\right) \in \ell^\infty(\mathcal{S})$.

ii.) Under any fixed alternative, i.e., when the data are distributed according to some F *that satisfies the alternative hypothesis* H_1 *in* (4.9),

$$\lim_{n\to\infty} P(S_n^{CM} > \varepsilon) = 1 \text{ for all constants } \varepsilon > 0. \tag{4.26}$$

4.4.2 Theoretical properties for quantile dependent regressors using constrained polynomial spline regression

Theorem 4.4.1 represents a generalization of previous tests for quantile regression models, since it allows the covariate to depend on the quantile. Imposing the assumptions from Cardot et al. (2005) on a quantile regression spline enables us to replace the quantile function by an appropriate spline estimator.

Let $B_{k,r} := (B_1, ..., B_{k+r})'$ denote the basis of the vectorial space of spline functions, where r is the degree of the piecewise polynomials. We will estimate $\Psi_\tau := \sum_{i=1}^{k+r} \theta_i(\tau) B_i$ for all $\tau \in \mathcal{T}$.

Assumption 3.

i.) *The function Ψ_τ is supposed to have a q'th derivative $\Psi_\tau^{(q')}$ such that*

$$\left|\Psi_\tau^{(q')}(t) - \Psi_\tau^{(q')}(s)\right| \leq C_1 |t-s|^v, \ \ s,t \in [0,1], \tag{4.27}$$

where $C_1 > 0$ and $v \in [0,1]$. In what follows, we set $q = q' + v$ and we suppose that $r \geq q \geq m$.

ii.) *The eigenvalues of $\mathbb{E}[\int_0^1 P(X,\tau) d\tau X]$ are strictly positive.*

iii.) *The errors defined by $\varepsilon = Y - \int_0^1 \Psi_\tau P(X,\tau) d\tau$ are i.i.d. and have density $f_{\varepsilon|X=x}$ given $X = x$, continuous and bounded below by a strictly positive constant at 0, uniformly for x.*

iv.) *The choice of knots corresponds to $k_n \sim n^{\frac{1}{2r+1}}$ with $r > 1/2$ and they are quasi-uniformly placed.*

Since finite sums of Donkser classes (cf. Assumption 2) are again Donsker, we can now formulate the second theorem

Theorem 4.4.2. *If Assumptions 2 and 3 are satisfied, then the following statements hold:*

i.) *Under the null hypothesis H_0 in (4.9),*

$$S_n^{CM^*} \xrightarrow{d} \int ||\mathbb{G}_2^*(y,x)||^2 dF_{YX}(y,x), \tag{4.28}$$

where $(\mathbb{G}_2^*)$ is the difference of tight zero mean Gaussian processes with a corresponding covariance structure according to Theorem 4.4.1.

ii.) *Under any fixed alternative, i.e., when the data are distributed according to some F that satisfies the alternative hypothesis H_1 in (4.9),*

$$\lim_{n\to\infty} P(S_n^{CM^*} > \varepsilon) = 1 \text{ for all constants } \varepsilon > 0. \tag{4.29}$$

In the simulation part we are using a quadratic B-spline for the test statistic $S_n^{CM^*}$.

4.4.3 Validity of the bootstrap procedure

Finally, we show that the proposed bootstrap procedure computes the correct critical value for out test statistics (4.16) and (4.20). This does not require any further assumptions. Under the null hypothesis, Assumptions 2 ensure that the bootstrap consistently estimates the limiting distribution for (4.16). For the more powerful test statistic $S_n^{CM^*}$, Assumption 3 has to be additionally fulfilled in order to ensure the Donsker property of the empirical cdf estimator. Under any fixed alternative, the bootstrap critical values can be shown to be bounded in probability. Together with Theorem 4.4.1ii.) and Theorem 4.4.2ii.), respectively, this implies that the proposed tests (4.16) and (4.20) are consistent.

Theorem 4.4.3. *Under Assumption 2, the following statements hold true for every* $\alpha \in (0,1)$

i.) Under the null hypothesis H_0 *in* (4.9)*, we have that*

$$\lim_{n\to\infty} P(S_n^{CM} > \hat{c}_n(\alpha)) = \alpha$$

ii.) Under any fixed alternative H_1 *in* (4.9)*, we have that*

$$\lim_{n\to\infty} P(S_n^{CM} > \hat{c}_n(\alpha)) = 1$$

If additionally Assumption 3 is fulfilled, then the following statements hold true for every $\alpha \in (0,1)$

iii.) Under the null hypothesis H_0 *in* (4.9)*, we have that*

$$\lim_{n\to\infty} P(S_n^{CM^*} > \hat{c}_n(\alpha)) = \alpha$$

iv.) Under any fixed alternative H_1 *in* (4.9)*, we have that*

$$\lim_{n\to\infty} P(S_n^{CM^*} > \hat{c}_n(\alpha)) = 1$$

In order to study the behavior of the CS induced Cramér-von Mises type test statistic $S_n^{CM^*}$ in finite samples we perform an extensive Monte Carlo Simulation, presented in the next section.

4.5 Monte Carlo simulation study

In this section, we show that our test $S_n^{CM^*}$ from (4.28) holds the size level and has superior power properties by means of twelve different data generating processes (DGPs). Thereby, the different DGPs cover location shift models (LS) and location-scale shift models (LSS) including heteroscedastic errors, both, in an univariate and multivariate setting. In order to assess the quality and validity of our proposed test against existing procedures, we will compare the test results (cf. Table 4.1 - Table 4.4) with the benchmark tests of Koenker & Xiao (2002), Chernozhukov (2002) and Rothe & Wied (2013) where comparisons are possible. Finally, we consider predominantly linear models and show that our test detects such only weakly misspecified models well.
For the definition of the twelve DGPs we introduce the following variables: Let $x_1 \sim Bin(1, 0.5)$, $x_2 \sim N(0, 1)$, $x_3 \in U(0, 1)$, $x_4 \in \chi^2(1)$, $u \sim N(0, 1)$, $w \sim N(0, 0.1)$, $v = (1-2x_1)\cdot v_2^* \cdot 8^{-0.5}$ with $v_2^* \sim \chi^2(2)$, where $Bin(\cdot, \cdot)$ describes the Binomial, $N(\cdot, \cdot)$ the normal, $U(\cdot, \cdot)$ the uniform and $\chi^2(\cdot)$ the chi squared distribution. Further, let $x_0 \in [0, 2\pi]$ and the variables $x_1, x_2, x_3, x_4, u, v, w$ be ally mutually independent.
DGPs $1-3$ from (4.30) represent the univariate case and serve as preliminary for our empirical application, since they model a linear and quadratic univariate processes. Hereby, DGP 1 describes a simple LS model, DGP 2 a more complex LSS model with a linear regressor and, finally, DPG 3 generates a LSS model with a quadratic influence factor. The multivariate case (cf. (4.31) and (4.32)) is specified by the DGPs $4-8$ that are from Rothe & Wied (2013) and DGP 9 from Chernozhukov (2002). Here, DGP 4 is a simple multivariate LS model with normal distributed errors. DGP 5 is again a simple LS model, but now the errors follow a mixture of a "positive" and "negative" χ^2 distribution with two degrees of freedom (normalized to have unit variance). DGPs $6-8$ are multivariate LSS models where the level of heteroscedasticity increases. DGP 9 is considered in order to compare our proposed testing procedure with

those provided in Chernozhukov (2002) and Koenker & Xiao (2002). When $\gamma_1 = 0$ DGP 9 is a LS model, otherwise it is a LSS model. DGPs $10-12$ (cf. (4.33)) are processes in which the functional form appears predominantly linear. DGP 10 is implemented by modeling the lower 50%-quantile linearly, while the upper 50%-quantile is modeled quadratically. Due to the quantile dependence of the regressors, DGP 10 cannot be tested with previous tests and therefore represents an extension of our test. DGP $11-12$ are appearing mainly linear in the interval $[0,1]$ and exhibit non-linear growth only at values close to 1. Assuming a linear model, DGPs of the form $10-12$ often impede the detection of misspecification.

$$\begin{aligned}
&\text{(DGP 1):} && f_1(x_0) := 0.25x_0 + 1 + u \\
&\text{(DGP 2):} && f_2(x_0) := 0.25x_0 + 1 + u \cdot x_0 \\
&\text{(DGP 3):} && f_3(x_0) := 0.25x_0^2 + 1 + u \cdot x_0^2
\end{aligned} \tag{4.30}$$

$$\begin{aligned}
&\text{(DGP 4):} && f_4(x_1, x_2) := x_1 + x_2 + u \\
&\text{(DGP 5):} && f_5(x_1, x_2) := x_1 + x_2 + v \\
&\text{(DGP 6):} && f_6(x_1, x_2) := x_1 + x_2 + (0.5 + x_1)u \\
&\text{(DGP 7):} && f_7(x_1, x_2) := x_1 + x_2 + (0.5 + x_1 + x_2^2)^{0.5}u \\
&\text{(DGP 8):} && f_8(x_1, x_2) := x_1 + x_2 + 0.2(0.5 + x_1 + x_2^2)^{1.5}u
\end{aligned} \tag{4.31}$$

$$\text{(DGP 9):} \qquad f_9(x_2) := x_2 + (1 + \gamma_1 \cdot x_2)u \tag{4.32}$$

$$\begin{aligned}
&\text{(DGP 10):} && f_{10}(x_4) := \begin{cases} 0.25 \cdot x_4^2 + 1 + 0.5 \cdot \epsilon \cdot x_4^2, & \text{for } \tau \geq 0.5 \\ -0.25 \cdot x_4 + 1 + u \cdot x_4, & \text{otherwise} \end{cases} \\
&\text{(DGP 11):} && f_{11}(x_3) := \sin(-\frac{\pi}{2} + x_3^3) + w \\
&\text{(DGP 12):} && f_{12}(x_3) := e^{f_5(x_3)}
\end{aligned} \tag{4.33}$$

In order to illustrate the performance of our test, we draw comparisons to common test procedures in the scope of quantile regression. The test proposed in

Koenker & Xiao (2002) (denoted as KX), which is based on Khmaladazation, which in turn refers to the Doob-Mayer decomposition of martingales, provides the starting point for quantile regression specification tests. We also consider the enhancement proposed in Chernozhukov (2002) (denoted as $Cher$). Furthermore, we compare our test with Rothe & Wied (2013) (denoted as RW) since our test is based on a similar principle but more flexible. The aforementioned tests are characterized by the following properties:

- The KX-test models the conditional qf paramterically by assuming a LS or a LSS model. In addition, the regressors are fixed for all quantiles considered and the estimation of non-parameter sparsity and score functions are required.
- In order to avoid such estimation, $Cher$ proposes a resampling testing procedure based on KX that results in better power and accurate size. However, he still assumes a fully parametrized model under the null hypothesis with non-varying regressors for distinct quantiles.
- RW propose a testing procedure for a wide range of parametric models that is based on a Cramèr-von Mises distance between an unrestricted estimate of the joint cdf and the estimate of the joint cdf under the null hypothesis. However, the regressors are assumed to be constant for all quantiles.

To analyze finite sample properties of our testing procedure, we consider different sample sizes n and set the number of Monte Carlo replications to 701, while the number of bootstrap replication is equal to $B = 500$.
Table 4.1 shows the comparison with RW for the univariate DGPs $1-3$. It can be noted that

- compared to RW our proposed testing procedure $S_n^{CM^*}$ consistently has better size properties.
- In particular, the test $S_n^{CM^*}$ also manages to maintain the size level when the structure of the error terms is highly heteroscedastic (cf. 5% column of DGP 3 in Table 4.1).

- In addition, the rejection rate for misspecified models (for DGP 3 we are assuming a linear LSS model in the last column of Table 4.1) in small samples ($n \leq 300$) is approximately three times higher than for the RW test.

Table 4.1: Size analysis to the significance level 0.10 and 0.05

	$DGP1$		$DGP2$		$DGP3$	
RW	10%	5%	10%	5%	5%	$Power$
$n = 30$	0.077	0.019	0.093	0.039	0.005	**0.032**
$n = 50$	0.061	0.016	0.095	0.038	0.016	**0.045**
$n = 100$	0.056	0.024	0.087	0.033	0.024	**0.075**
$n = 300$	0.055	0.028	0.078	0.032	0.026	**0.312**
$n = 500$	0.056	0.016	0.069	0.029	**0.010**	0.486
$n = 1000$	0.043	0.016	0.069	0.030	0.014	0.883
$n = 2000$	0.064	0.020	0.066	0.030	0.014	1.000
S_n^{CM*}	10%	5%	10%	5%	5%	$Power$
$n = 30$	0.101	0.035	0.089	0.037	0.028	**0.095**
$n = 50$	0.103	0.046	0.074	0.027	0.037	**0.147**
$n = 100$	0.094	0.043	0.112	0.061	0.064	**0.407**
$n = 300$	0.090	0.043	0.159	0.084	0.047	**0.988**
$n = 500$	0.086	0.043	0.111	0.058	**0.050**	1.000
$n = 1000$	0.095	0.048	0.095	0.038	0.056	1.000
$n = 2000$	0.098	0.049	0.092	0.042	0.044	1.000

The number of Monte Carlo repetitions is equal to 701 with 500 bootstrap replications. For the size analysis the wrap speed bootstrap procedure is applied. Here, the quantile is modeled by a B-spline of second order with penalty term $\lambda = 1$ and $\sqrt{n}$ knots, meeting monotonicity assumptions. The 7^{th} and last column named $Power$ depicts the power analysis while the quantile function is assumed to follow linear LSS model under the null hypothesis.

Table 4.2 additionally illustrates the comparison with KX for the DGPs $4-8$, whereby a location shift model is assumed under the null hypothesis. Thus, the results of DGPs 4 and 5 reflect size properties, while DGPs $6-8$ measure the power of our and the benchmark tests RW and KX.

- It can be observed that our test S_n^{CM*} holds the size for multivariate processes (cf. DGP $4, 5$).
- KX has difficulties to detect misspecification when heteroscedasticity prevails (cf. DGP $6-8$).
- RW usually detects misspecification. However, the rejections rate of the test S_n^{CM*} are clearly higher compared to those from RW even in small

samples (cf. $n = 100$ DGP 7 of Table 4.2).

Table 4.2: Power analysis location shift (LS)

	RW		*KX*		S_n^{CM*}	
$n = 100$	10%	5%	10%	5%	10%	5%
$DGP4$	0.093	0.048	0.067	0.035	0.122	**0.068**
$DGP5$	0.085	0.033	0.069	0.037	0.114	**0.065**
$DGP6$	0.829	**0.669**	0.082	**0.047**	0.870	**0.838**
$DGP7$	0.404	**0.239**	0.097	**0.049**	0.669	**0.565**
$DGP8$	0.874	**0.746**	0.055	**0.027**	0.970	**0.944**
$n = 300$	10%	5%	10%	5%	10%	5%
$DGP4$	0.109	0.056	0.107	0.039	0.125	**0.068**
$DGP5$	0.096	0.043	0.066	0.024	0.120	**0.056**
$DGP6$	1.000	0.997	0.336	0.231	1.000	1.000
$DGP7$	0.847	0.679	0.147	0.076	0.950	0.908
$DGP8$	1.000	0.997	0.099	0.050	1.000	1.000

All results are one-to-one transferred from Rothe & Wied (2013) (RW). Other details of the set up are as those reported there. The null hypotheses assumes a LS quantile regression model. The number of MC repetitions is equal to 701 with 500 bootstrap replications. Here, the quantile is modeled by a B-spline of second order with penalty term $\lambda = 1$ and $\sqrt{n}$ knots, meeting monotonicity assumptions.

Table 4.3 provides a comparison with the standard testing procedure proposed in Koenker & Xiao (2002) and the enhancement from Chernozhukov (2002), where the results of Table 4.3 of the benchmark tests KX and $Cher$ are taken from Chernozhukov (2002).

- Even if the structure of DGP 9 is less complex compared to the other DGPs from (4.30)-(4.33), the test S_n^{CM*} has consistently better finite sample properties compared to the benchmarks KX and $Cher$.
- In small samples (cf. $n = 100$) the strong results of KX and $Cher$ could be improved further.

Finally, Table 4.4 now examines size and power properties for the DGPSs $10 - 12$.

- In each of the DGPs considered, the test holds the significance level.
- Assuming a linear model, misspecification is detected even in small sample sizes.

Table 4.3: Power analysis location shift (LS) for DGP9

	KX			$Cher$			$S_n^{CM^*}$		
	Size	Power		Size	Power		Size	Power	
$\gamma_1 =$	0	0.2	0.5	0	0.2	0.5	0	0.2	0.5
$n = 100$	0.101	0.264	0.898	0.014	0.348	0.980	**0.0495**	**0.396**	**0.99**
$n = 200$	0.070	0.480	0.988	0.052	0.752	1.000	0.063	0.772	1.000
$n = 300$	0.062	0.622	0.998	0.058	0.910	1.000	0.068	0.930	1.000

All results are one-to-one transferred from Koenker & Xiao (2002) (KX) and Chernozhukov (2002) ($Cher$), respectively. Other details of the set up are as those reported there. The null hypotheses assumes a LS quantile regression model. The Monte Carlo study for the proposed test uses 701 replications with $B = 500$ bootstrap replications. The significance level is 0.05.

- DGP 10 cannot be tested with previous approaches due to the quantile dependent regressors. The slightly lower power for DGP 10 is due to the fact that half of the observations actually follow a linear relationship.

Table 4.4: Size and power analysis under a linear null hypothesis

	$DGP10$		$DGP11$		$DGP12$	
$S_n^{CM^*}$	5%	$Power$	5%	$Power$	5%	$Power$
$n = 30$	0.068	0.177	0.014	0.055	0.009	0.069
$n = 50$	0.057	0.189	0.018	0.285	0.013	0.318
$n = 100$	**0.051**	**0.192**	**0.033**	**0.979**	**0.023**	**0.989**
$n = 300$	0.039	0.469	0.040	1.000	0.031	1.000
$n = 500$	0.042	0.519	0.039	1.000	0.029	1.000
$n = 1000$	0.046	0.658	0.034	1.000	0.035	1.000
$n = 2000$	0.042	0.743	0.041	1.000	0.049	1.000

The number of Monte Carlo repetitions is equal to 701 with 500 bootstrap replications. For the size analysis the wrap speed bootstrap procedure is applied. Here, the quantile is modeled by a B-spline of second order with penalty term $\lambda = 1$ and $\sqrt{n}$ knots, meeting monotonicity assumptions. The columns named $Power$ depict the power analysis. Under the null hypothesis, the quantile function is modeled as a linear LSS function for DGP 10 and a linear LS function for DGPs 11, 12, respectively.

In summary, the Monte Carlo study has thus shown that our proposed test procedure holds the significance level and also has superior power properties compared to three benchmark tests, even in small samples. The procedure works for both, univariate and multivariate DGPs and can also test models with quantile-dependent regressors. Even weakly misspecified models are detected in sufficiently large sample sizes.

4.6 Conditional income disparities between East and West Germany

In this section, we apply the bootstrap version of the specification test for generalized quantile regression models to conditional income distributions in Germany. For this purpose, we utilize information from the German Socio-Economic Panel (SOEP, Wagner et al., 2007). More specifically, we consider real gross annual personal labor income in Germany as defined in Bach et al. (2009) for the years 2001 to 2010. Following the standard literature, we only consider the income of males in full-time employment (see, among others, Dustmann et al. (2009); Card et al. (2013)) in the age range 20-60. This yielded 7220 individuals and is the data set that was also used in Klein et al. (2015). The variables age, $origin$ (East or West Germany) and $years$ are available as covariates (cf. Table 4.5 for a full description of the data). To obtain an estimate of the quantile function and to take full advantage of the spline approximation, we first regressed the income on the dummy coded variable $years$ and then performed a quantile regression using the variables age or age^2 on the residuals[7]. This approach takes account of the fact that income grows solely with increasing age. Rather, it can be observed that income increases at the beginning of employment, peaks in middle age and finally decreases (Creedy & Hart, 1979; Luong & Hébert, 2009; Klein et al., 2015).
In order to gain further insights, we next conduct the Machado-Mata decomposition of the year dummy adjusted data set conditioned on the $origin$ according to Melly (2005) and Machado & Mata (2005)[8]. For the decomposition we assume that the quantile function of the income Y can be represented as a function of the form

$$F_{Y|X}^{-1}(\tau|X) = P(X,\tau)'\theta(\tau), \tag{4.34}$$

where X depicts the matrix of covariates, that consists of the variables age or

[7]We consider this approach justified since four out of six tests did not reject the null hypothesis that there is no correlation between *age* and *year* dummies and age^2 and *year* dummies, respectively.

[8]Another application of the Machado and Mata decomposition for differences in incomes can be found in Landmesser (2016).

age^2 and the quantile $\tau \in (0,1)$. Specifically, we consider here three different linear quantile regression models: The first model describes an entirely linear effect of the regressor age on income for all quantiles $\tau \in (0,1)$, i.e. $P(X,\tau) = age$ for all $\tau \in (0,1)$. The second models a quadratic influence of age on income for all quantiles $\tau \in (0,1)$, i.e. $P(X,\tau) = age^2$ for all $\tau \in (0,1)$. And finally, the third model considers the sum of the regressors age and age^2 that are constant for all quantiles $\tau \in (0,1)$, i.e. $P(X,\tau) = age + age^2$ for all $\tau \in (0,1)$. Due to the probability integral transform theorem the sequence $P(X,\tau_i)'\hat{\theta}(\tau_i)$ for $\tau_i \overset{i.i.d.}{\sim} Uni(0,1)$, $i = 1,...,n$ constitutes a random sample from the estimated conditional distribution of income Y given the covariates X (Machado & Mata, 2005). In order to obtain the difference between East and West, first, the coefficients for East $\hat{\theta}_E(\tau)$ and West $\hat{\theta}_W(\tau)$ for $\tau \in \{0.1, 0.2, ..., 0.9\}$ are estimated on the basis of the disjoint subsets of the covariates for East X_E and West X_W and the corresponding income in the East Y_E and West Y_W. Second, we draw with replacement B random samples X_E^i and X_W^i for $i = 1,...,B$ from the corresponding covariate subsets X_E and X_W, respectively to obtain a random sample via (4.34) for the distribution of the income Y_l^i, $i = 1,...,B$, $l = E,W$. Thus, the estimated income difference $\hat{\Delta}_y$ for incomes in East Y_E and incomes in West Y_W can now be decomposed according to Machado-Mata as

$$\begin{aligned}
\hat{\Delta}_Y &= \hat{F}^{-1}_{Y_E|X_E}(\tau|X_E) - \hat{F}^{-1}_{Y_W|X_W}(\tau|X_W) && (4.35)\\
&= \left(P(X_E^B,\tau) - P(X_W^B,\tau)\right)\hat{\theta}_E(\tau) + \left(\hat{\theta}_E(\tau) - \hat{\theta}_W(\tau)\right)P(X_W^B,\tau), && (4.36)
\end{aligned}$$

where the first summand of (4.36) is the explained while the second summand depicts the unexplained difference.

Table 4.6 depicts the counterfactual analysis of the effect of $origin$ on income. The covariates used for the quantile regressions are age (row $4-9$ of Table 4.6), age^2 (row $11-16$ of Table 4.6) and the sum of these two variables (row $18-23$ of Table 4.6). The results in Table 4.6 suggest that there is a significant income gap between East and West Germany over the period considered, which is particularly striking in the first line, where the observed income differences ranges from 26.21% to 35.49%. However, the income difference between the smallest quantile $\tau = 0.1$ and the largest $\tau = 0.9$ decreases by about eight

Table 4.5: Description of the German labor income data

	Description		
Y	gross market labor income, (continuous $1,257€ \leq Y \leq 280,092€$, average $= 46,641€$)		
origin	indicator for East or West (binary, -1=West (73.8%), 1=East (26.2%))		
age	age of the male in years (continuous, $20 \leq$ age ≤ 60, average $= 38$)		
years	time in years (categorical, $2001 \leq years \leq 2010$, 10 years)		
(Sub)sample	Description	Average income (Std.)	Observations
Ger	Entire sample	$51,026$ $(30,569)$	$n = 7220$
West	Subsample with $origin = -1$	$55,141$ $(31,494)$	$n = 5325$
East	Subsample with $origin = 1$	$39,463$ $(24,336)$	$n = 1895$

percent. If income is to be explained by the single covariate *age* or age^2, it cannot be assumed that the model is sufficiently well specified for all quantiles due to high residuals (4.37 for $\tau = 0.1$ and 7.33 for $\tau = 0.9$), indicating misspecification. However, the covariate age^2 seems to be appropriate for the smallest quantile 0.1 while a linear effect of age to income seems to prevail in higher quantiles. In contrast, the additive model $age + age^2$ seems to capture the income effect for all quantiles quite well due to moderate residuals. For all decompositions it holds, that *age* and age^2 contribute a maximum of 16% to the explanation of the income difference between East and West Germany (except highest quantile in age^2, i.e. 25.41). Due to the different residuals and the different explanatory power of the income difference between East and West for the quantile regressions based on *age* or age^2, it seems reasonable to assume that *age* and age^2 have different effects for different quantiles. For example, the residual of the 30% quantile of *age* is about 18 times smaller than the residual of the corresponding quantile regression using age^2 as explanatory variable. It is therefore suspected that the a linear effect of age dominates in this quantile. The emerging, more general question, at which quantile age has a linear or quadratic effect on income, can be answered with the help of the proposed test.

For this purpose, we have defined five different model specifications (4.37), which should take into account the observations of the Machado-Mata decom-

Table 4.6: Decomposition of the West/East income differential

$\tau =$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
raw gap	-35.49	-32.4	-33.28	-29.06	-28.38	-26.44	-26.21	-26.93	-27.38
					age				
M-M gap	-39.87	-36.64	-33.52	-31.62	-31.45	-29.37	-29.68	-28.83	-25.89
Explained	-1.84	-3.65	-2.06	-2.99	-2.35	-2.19	-0.85	-0.31	-0.1
Unexpl.	-38.02	-32.98	-31.47	-28.63	-29.1	-27.18	-28.83	-28.52	-25.8
% Explained	4.63	9.97	6.13	9.45	7.46	7.46	2.87	1.07	0.38
% Unexpl.	95.37	90.03	93.87	90.55	92.54	92.54	97.13	98.93	99.62
Residuals	4.37	4.23	0.24	2.56	3.07	2.93	3.46	1.9	-1.49
					age^2				
M-M gap	-36.41	-38.13	-37.52	-35.49	-31.21	-31.96	-32.26	-31.55	-34.72
Explained	-3.07	-6.08	-4.49	-6.43	-2.81	-2.6	-3.92	-4.35	-8.82
Unexpl.	-33.35	-32.05	-33.04	-29.06	-28.4	-29.36	-28.34	-27.2	-25.89
% Explained	8.42	15.94	11.96	18.12	8.99	8.13	12.15	13.8	25.41
% Unexpl.	91.58	84.06	88.04	81.88	91.01	91.87	87.85	86.2	74.59
Residuals	0.92	5.73	4.24	6.44	2.83	5.51	6.05	4.62	7.33
					$age+age^2$				
M-M gap	-33.39	-31.80	-33.16	-30.28	-28.49	-28.90	-27.61	-28.25	-25.69
Explained	2.03	1.55	-1.44	-1.5	0.13	0.31	0.33	-3.09	1.67
Unexpl.	-35.42	-33.35	-31.72	-28.78	-28.62	-29.21	-27.94	-25.16	-27.36
% Explained	6.09	4.89	4.34	4.94	0.45	1.09	1.19	10.95	6.49
% Unexpl.	93.91	95.11	95.66	95.06	99.55	98.91	98.81	89.05	93.51
Residuals	-2.10	-0.61	-0.12	1.22	0.11	2.45	1.40	1.32	-1.69

The covariates used for the quantile regressions are *age* (row 4-9), age^2 (row 11-16) and the sum of these two variables (row 18-23). The second row *raw gap* depicts the observed income gap between East and West. Remaining rows show three different Machado-Mata decompositions using *age*, age^2 and $age+age^2$ as covariates for the quantile regression models. The rows *M-M gap* are the estimated gap of the income difference depending on the underlying quantile regression model. The quantiles τ range from 0.1 to 0.9. The number of bootstrap replications is equal to 2500. All numbers are in percent. Totals may not sum exactly due to rounding.

position in Table 4.6. Figure 4.1 visualizes the testing problem and provides further indications of when age might have a quadratic or linear effect. The blue line in Figure 4.1 describes the empirical 90 percent quantile. The corresponding dashed blue line represents the corresponding estimate of the quantile regression. The green (50 percent quantile) and red (10 percent quantile) lines are the equivalent counterparts. In the 5 different quantile regression models considered, the effect of *age* depends on the quantile. Specifications $1-3$ from (4.37) describe quadratic dependencies in the upper or lower quantiles. Specification 4 and 5 model a completely linear and quadratic dependence structure in the covariate, respectively.

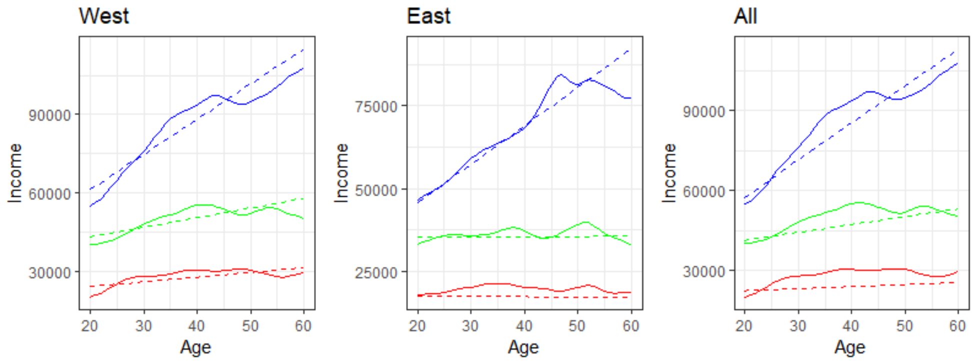
The testing procedure is applied to the subsamples East (only individuals from East Germany are considered) and West (only individuals from West Germany are considered) as well as to the complete data set (cf. last column *All* in Table 4.7).

$$
\begin{aligned}
&\text{Specification 1 :} && F^{-1}_{Y|X}(\tau|x) = \begin{cases} x^{2\prime}\theta_0, & \text{for } 0 \leq \tau \leq 0.1 \\ x'\theta_0\ , & \text{for } 0.1 \leq \tau \leq 0.9 \\ x^{2\prime}\theta_0, & \text{otherwise} \end{cases} \\
&\text{Specification 2 :} && F^{-1}_{Y|X}(\tau|x) = \begin{cases} x^{2\prime}\theta_0, & \text{for } 0 \leq \tau \leq 0.1 \\ x'\theta_0\ , & \text{for } 0.1 \leq \tau \leq 0.9 \\ x'\theta_0\ , & \text{otherwise} \end{cases} \\
&\text{Specification 3 :} && F^{-1}_{Y|X}(\tau|x) = \begin{cases} x'\theta_0\ , & \text{for } 0 \leq \tau \leq 0.1 \\ x'\theta_0\ , & \text{for } 0.1 \leq \tau \leq 0.9 \\ x^{2\prime}\theta_0, & \text{otherwise} \end{cases} \\
&\text{Specification 4 :} && F^{-1}_{Y|X}(\tau|x) = x'\theta_0 \\
&\text{Specification 5 :} && F^{-1}_{Y|X}(\tau|x) = x^{2\prime}\theta_0
\end{aligned}
\tag{4.37}
$$

Since the sample sizes for East, West and All differ and in order to make the results comparable, we computed the rejection rates of subsamples of East, West and All of size $n = 500, 1000$. We repeated this procedure for every subsample a total of 501 times. The results are listed in Table 4.7.

First, it can be observed that age does not have a completely linear influence on income, as the rejection rates for the 4 specifications are sufficiently high, 0.828 for $n = 1000$, respectively). Assuming a complete quadratic relationship between age and income, this statement cannot be upheld, since the rejection rates for the income distribution in West Germany are below the significance level of 5%. However, the results clearly show that neither in East Germany nor in all of Germany (cf. columns *East* and *All* of Table 4.7) can the income distribution be adequately described by a quadratic process due to their

Figure 4.1: Conditional income quantiles for East/West and entire Germany as functions of *age*



Figures show the 0.9 (blue) 0.5 (green) 0.1 (red) smoothed quantiles (using a cubic smoothing spline with smoothing factor 0.5 (cf. R function smooth.spline)) of the income conditioned on the *origin* and the unconditioned data set (*All*). The dashed line depicts the corresponding quantile regression estimate with *age* as covariate.

Table 4.7: Empirical rejection frequencies of the test statistic S_n^{CM*}

	West	East	All
$n = 500$			
Specification 1	0.023	0.048	0.054
Specification 2	0.122	0.142	0.118
Specification 3	0.010	0.030	0.025
Specification 4	0.080	0.410	0.345
Specification 5	0.066	0.295	0.242
$n = 1000$			
Specification 1	0.014	0.106	0.098
Specification 2	0.242	0.301	0.215
Specification 3	0.019	0.056	0.036
Specification 4	0.128	0.828	0.705
Specification 5	0.082	0.557	0.463

The table depicts the subsample rejection rate of size n of the specification being used from (4.37). The number of subsamplings is 501 and the critical values were calculated at a significance level of 5%.

rejection rates. Second, the Specifications 2 assuming a quadratic structure in the 0.1 and lower quantiles while the remaining quantiles follow a linear model seems for all subsamples considered inappropriate. Third, the Specifications 1 and 3 have the lowest rejection rates for all subsamples indicating that age has a quadratic influence for quantiles 0.9 and higher. In particular, Specification 3 seems to model the income structure sufficiently well for all 3 samples considered. However, as the rejection rates, especially for Specifications 4 and 5, differ sufficiently between East and West, conditional different income distributions between East and West are likely. The test results are in line with the findings of other studies: Based on the different structure of the conditional quantile functions and the corresponding rejections rates for different specifications in Table 4.7 significant structural differences between East and West Germany can still be assumed (Kluge & Weber, 2018).

4.7 Conclusion

We believe there are many different areas of application in which the influence of the regressors depends on the quantile linearly or nonlinearly or even in a more complex functional form. A well-known example is the effect of age on the income distribution, which we have taken as illustration. Previous testing procedures of quantile regression are not able to test such influences separately. The present paper proposes a test for generalized quantile regression that addresses these two issues jointly. To improve finite sample properties, we replace quantile regression function by a quadratic monotone B-spline. Our Monte Carlo study illustrates that the proposed method has superior test properties compared to several existing benchmarks from the literature. In addition, a detailed investigation of the conditional income distributions between East and West Germany using the Machado-Mata decomposition reveals that still income differences between the regions in Germany are present, even more than two decades after the reunification. The application of our test could statistically confirm a different functional correlation between the income distributions in East and West Germany.

APPENDIX A

Testing the Correct Specification of a Spatial Dependence Panel Model for Stock Returns

A.1 A Two Step GMM Estimation Procedure for SAR(m) Models

Given the assumptions given in 2nd section hold true. The covariance matrix of $\boldsymbol{y}_t = (I_n - \sum_{i=1}^{m} \rho_i W_i)^{-1} \boldsymbol{\varepsilon}_t$ is given by

$$\mathsf{Cov}[\boldsymbol{y}_t] = \left(I_n - \sum_{i=1}^{m} \rho_i W_i\right)^{-1} \Sigma \left(I_n - \sum_{i=1}^{m} \rho_i W_i'\right)^{-1} =: V.$$

For the estimation, a two step procedure is considered. First, we estimate the correlation parameters by the method of moments which does not depend on the parameters of variance. Second, we estimate the variance parameters.
The moment estimator for the correlation parameters uses the following m-moment conditions:

$$\mathsf{E}\left[\boldsymbol{\varepsilon}_t' W_i \boldsymbol{\varepsilon}_t\right] = \mathsf{tr}(W_i \Sigma) = 0 \qquad \text{for} \qquad i = 1, ..., m. \tag{A.1}$$

Clearly, the variance parameters σ_i^2 for $i = 1, ..., m$ do not enter the moment conditions. Replacing $\boldsymbol{\varepsilon}_t$ by

$$\boldsymbol{\varepsilon}_t = \left(I_n - \sum_{i=1}^{m} \rho_i W_i\right) \boldsymbol{y}_t$$

and averaging over t gives the theoretical system of equations

$$\Gamma \boldsymbol{\lambda} + \boldsymbol{\gamma} = 0,$$

where $\boldsymbol{\lambda} := \boldsymbol{\lambda}(\boldsymbol{\rho})$ is a functional vector of $\boldsymbol{\rho} := (\rho_1, ...\rho_m)$ of dimension $M := \binom{m}{1} + \binom{m+2-1}{2}$, $\binom{\cdot}{\cdot}$ denoting the binomial coefficient, such that

$$\begin{aligned}
\lambda_i &= \rho_i && \text{for} \quad i = 1, ..., m && \text{(A.2)}\\
\lambda_{m+i} &= \rho_i^2 && \text{for} \quad i = 1, ..., m && \text{(A.3)}\\
\lambda_{2m+\#\{ij\,|\,i<j, i<l, j\le k\}} &= \rho_l \rho_k && \text{for} \quad l, k = 1, ..., m, && \text{(A.4)}
\end{aligned}$$

where $\#\{ij\,|\,i < j, i < l, j \le k\}$ represents the number of integer pairs ij such that the conditions $i < j, i < l$ and $j \le k$ are fulfilled for $l, k = 1, ..., m$. The elements of $\Gamma \in \mathbb{R}^{m \times M}$ and $\gamma \in \mathbb{R}^m$ are defined by for $i, j = 1, ..., m$,

$$\begin{aligned}
\Gamma_{i,j} &= \mathsf{E}\left[-\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{y}_t' \left(W_i + W_i'\right) W_j \boldsymbol{y}_t\right], && \text{(A.5)}\\
\Gamma_{i,m+j} &= \mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{y}_t' W_j' W_i W_j \boldsymbol{y}_t\right], && \text{(A.6)}\\
\Gamma_{i,2m+\#\{ij\,|\,i<j,i<l,j\le k\}} &= \mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{y}_t' W_l' \left(W_i + W_i'\right) W_k \boldsymbol{y}_t\right], && \text{(A.7)}\\
\gamma_i &= \mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{y}_t' W_i \boldsymbol{y}_t\right].
\end{aligned}$$

Let G and $\boldsymbol{g}$ be the empirical counterparts of Γ and $\boldsymbol{\gamma}$, i.e. the expectation operator is left out. The moment estimator for $\boldsymbol{\rho} = (\rho_1, ..., \rho_m)'$ is defined as

$$\hat{\boldsymbol{\rho}} := (\hat{\rho}_1, ..., \hat{\rho}_m)' := \arg\min_{\boldsymbol{\rho}\in S} ||G\boldsymbol{\lambda} + \boldsymbol{g}||$$

where $||\cdot||$ represents the euclidean norm.

Remark A.1.1. *For $k, l \in \{1, ..., m\}$, the entries of $\mathsf{E}[G] = \Gamma$ given in (A.5)-(A.7) can be calculated as*

$$\begin{aligned}
&\Gamma_{k,l} = \mathsf{tr}\left(\left(W_k + W_k'\right) W_l V\right),\\
&\Gamma_{k,m+l} = \mathsf{tr}\left(W_l' W_k W_l V\right),\\
&\Gamma_{i,2m+\#\{ij\,|\,i<j,i<l,j\le k\}} = \mathsf{tr}\left(W_l' \left(W_i + W_i'\right) W_k V\right).
\end{aligned}$$

The following remark illustrates the results for the SAR(3) model.

Remark A.1.2. *For the case $m = 3$ we have to estimate the spatial vector $\boldsymbol{\rho} := (\rho_1, \rho_2, \rho_3)$. The corresponding theoretical system of equations is given by $\Gamma\boldsymbol{\lambda} + \boldsymbol{\gamma} = 0$ with $\Gamma := \left(\Gamma_{(1)}, \Gamma_{(2)}, \Gamma_{(3)}\right) \in \mathbb{R}^{m\times M}$, $\boldsymbol{\lambda} \in \mathbb{R}^{M\times 1}$ and $\boldsymbol{\gamma} \in \mathbb{R}^{m\times 1}$ with $M = 3 + \binom{4}{2} = 9$*

which are defined as

$$\Gamma := \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_t'\left(\Gamma_{(1)},\Gamma_{(2)},\Gamma_{(3)}\right)\boldsymbol{y}_t\right]$$

with

$$\Gamma_{(1)} := \begin{pmatrix} (W_1+W_1')W_1 & (W_1+W_1')W_2 & (W_1+W_1')W_3 \\ (W_2+W_2')W_1 & (W_2+W_2')W_2 & (W_2+W_2')W_3 \\ (W_3+W_3')W_1 & (W_3+W_3')W_2 & (W_3+W_3')W_3 \end{pmatrix},$$

$$\Gamma_{(2)} := \begin{pmatrix} W_1'W_1W_1 & W_2'W_1W_2 & W_3'W_1W_3 \\ W_1'W_2W_1 & W_2'W_2W_2 & W_3'W_2W_3 \\ W_1'W_3W_1 & W_2'W_3W_2 & W_3'W_3W_3 \end{pmatrix},$$

$$\Gamma_{(3)} := \begin{pmatrix} W_1'(W_1+W_1')W_2 & W_1'(W_1+W_1')W_3 & W_2'(W_1+W_1')W_3 \\ W_1'(W_2+W_2')W_2 & W_1'(W_2+W_2')W_3 & W_2'(W_2+W_2')W_3 \\ W_1'(W_3+W_3')W_2 & W_1'(W_3+W_3')W_3 & W_2'(W_3+W_3')W_3 \end{pmatrix},$$

$$\boldsymbol{\lambda} := \left(\rho_1,\rho_2,\rho_3,\rho_1^2,\rho_2^2,\rho_3^2,\rho_1\rho_2,\rho_1\rho_3,\rho_2\rho_3\right)$$

and

$$\boldsymbol{\gamma} := \mathsf{E}\left[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_t'\left(W_1',W_2',W_3'\right)'\boldsymbol{y}_t\right].$$

Since the theoretical term $\Gamma\boldsymbol{\lambda}+\boldsymbol{\gamma}$ is equal to zero for the true parameter values, the moment estimator for $\hat{\boldsymbol{\rho}}$ minimizes the corresponding empirical system $G\boldsymbol{\lambda}+\boldsymbol{g}$. Arnold et al. (2013) prove consistency and asymptotic normality of the moment estimator (cf. Theorem A.1.3) for $T\to\infty$, for which an additional assumption is needed.

Assumption 4.

1. *The true parameter $\boldsymbol{\rho}\in S$ is the unique solution of the theoretical system of equations, i.e.*

$$\Gamma\boldsymbol{\lambda}+\boldsymbol{\gamma}=0 \Leftrightarrow \hat{\boldsymbol{\rho}}=\boldsymbol{\rho}.$$

2. *The matrix $\mathsf{E}\left(\frac{\partial(G\boldsymbol{\lambda}+\boldsymbol{g})}{\partial\hat{\boldsymbol{\rho}}}(\boldsymbol{y}_t,\boldsymbol{\rho})\right) =: \boldsymbol{d} = \Gamma\boldsymbol{\lambda}^{(1)}$ exists, is finite and has full rank with $\boldsymbol{\lambda}^{(1)}$ a $(M\times m)$ dimensional matrix defined as*

$$\boldsymbol{\lambda}^{(1)}(l,l)=1, \qquad \boldsymbol{\lambda}^{(1)}(2m+\#\{ij\,|\,i<j,i<l,j\le k\},l)=\rho_k$$

$$\boldsymbol{\lambda}^{(1)}(m+l,l)=2\rho_l, \qquad \boldsymbol{\lambda}^{(1)}(2m+\#\{ij\,|\,i<j,i<l,j\le k\},k)=\rho_l$$

for all $l, k = 1, ..., m$.

3. *For*

$$f(\boldsymbol{y}_t, \boldsymbol{\rho}) = \begin{pmatrix} \boldsymbol{\varepsilon}_t' W_1 \boldsymbol{\varepsilon}_t \\ \vdots \\ \boldsymbol{\varepsilon}_t' W_m \boldsymbol{\varepsilon}_t \end{pmatrix},$$

it holds that, for $j \to \infty$, $\mathsf{E}[f(\boldsymbol{y}_t, \boldsymbol{\rho}) \,|\, f(\boldsymbol{y}_{t-j}, \boldsymbol{\rho}), f(\boldsymbol{y}_{t-j-1}, \boldsymbol{\rho}), \ldots]$ *converges in mean square to zero and that, for*

$$\begin{aligned} \boldsymbol{v}_j :&= \mathbb{E}[f(\boldsymbol{y}_t, \rho) \,|\, f(\boldsymbol{y}_{t-j}, \rho), f(\boldsymbol{y}_{t-j-1}, \rho), ...) \\ &\quad - \mathbb{E}[f(\boldsymbol{y}_t, \rho) \,|\, f(\boldsymbol{y}_{t-j-1}, \rho), f(\boldsymbol{y}_{t-j-2}, \rho), ...] \end{aligned}$$

the infinite sum $\sum_{t=-\infty}^{\infty} \mathsf{E}[(\boldsymbol{v}_j \boldsymbol{v}_j)^{\frac{1}{2}}]$ *is finite.*

Under the Assumptions 1 and 4 the GMM estimator $\hat{\boldsymbol{\rho}}$ is consistent and asymptotic normal as the following theorem shows:

Theorem A.1.3. *Let Assumption 1 and 4 hold. Then, for*

$$S_W = \sum_{t=-\infty}^{\infty} \mathsf{E}[f(\boldsymbol{y}_1, \boldsymbol{\rho}) f(\boldsymbol{y}_t, \boldsymbol{\rho})']$$

and $T \to \infty$ *it holds:*

1. $\hat{\boldsymbol{\rho}} \overset{p}{\to} \boldsymbol{\rho}$
2. $\sqrt{T}(\hat{\boldsymbol{\rho}} - \boldsymbol{\rho}) \overset{d}{\to} N(0, \boldsymbol{d}^{-1} S_W (\boldsymbol{d}^{-1})')$.

A.2 Proofs

Theorem 4.4.1 is proved by means of the following Lemmas.

Lemma A.2.1. *Let* I_n *denote the n-dimensional identity matrix and* W *the m-dimensional stack of spatial matrices, i.e.* $W' = (W_1', \ldots, W_m')$ *with* $W_i \in R^{n \times n}$ *for* $i = 1, ..., m$. *Under Assumption 1 and given that* $\{\boldsymbol{\varepsilon}_t\}_{t \in \{1,\ldots,T\}}$ *is serially independent the following holds for* $\boldsymbol{\rho} := (\rho_1, \ldots, \rho_m)$ *and* $\hat{\boldsymbol{\rho}} = (\hat{\rho}_1, \ldots, \hat{\rho}_m)$

$$\sqrt{T}\ \hat{Cov}[\hat{\boldsymbol{\varepsilon}}_t] = \frac{1}{\sqrt{T}} \sum \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' + \frac{1}{T} \sum \Delta_T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' + \frac{1}{T} \sum \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \Delta_T' + \frac{1}{T} \sum \Delta_T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \frac{\Delta_T'}{\sqrt{T}}$$

with $\Delta_T := \sqrt{T}((\boldsymbol{\rho} - \hat{\boldsymbol{\rho}}) \otimes I_n) W (I_n - (\boldsymbol{\rho} \otimes I_n) W)^{-1}$, *where* $\otimes$ *represents the Kronecker product.*

Proof. It holds:

$$
\begin{aligned}
&\sqrt{T}\hat{\text{Cov}}[\hat{\boldsymbol{\varepsilon}}_t] = \sqrt{T}\hat{\mathbb{E}}[\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t'] \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t' = \frac{1}{\sqrt{T}}\sum(I_n - (\hat{\boldsymbol{\rho}}\otimes I_n)W)\boldsymbol{y}_t\boldsymbol{y}_t'(I_n - (\hat{\boldsymbol{\rho}}\otimes I_n)W)' \\
&= \frac{1}{\sqrt{T}}\sum(I_n - (\hat{\boldsymbol{\rho}}\otimes I_n)W)(I_n - (\boldsymbol{\rho}\otimes I_n)W)^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' \\
&\quad (I_n - (\hat{\boldsymbol{\rho}}\otimes I_n)W)(I_n - (\boldsymbol{\rho}\otimes I_n)W)^{-1}]' \\
&= \frac{1}{\sqrt{T}}\sum(I_n - (\boldsymbol{\rho}\otimes I_n)W + (\boldsymbol{\rho}\otimes I_n)W - (\hat{\boldsymbol{\rho}}\otimes I_n)W)(I_n - (\boldsymbol{\rho}\otimes I_n)W)^{-1}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' \\
&\quad [(I_n - (\boldsymbol{\rho}\otimes I_n)W + (\boldsymbol{\rho}\otimes I_n)W - (\hat{\boldsymbol{\rho}}\otimes I_n)W)(I_n - (\boldsymbol{\rho}\otimes I_n)W)^{-1}]' \\
&= \frac{1}{T}\sum[\sqrt{T}I_n + \sqrt{T}((\boldsymbol{\rho}-\hat{\boldsymbol{\rho}})\otimes I_n)W(I_n - (\boldsymbol{\rho}\otimes I_n)W)^{-1}]\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' \\
&\quad [\sqrt{T}I_n + \sqrt{T}((\boldsymbol{\rho}-\hat{\boldsymbol{\rho}})\otimes I_n)W(I_n - (\boldsymbol{\rho}\otimes I_n)W)^{-1}]' \\
&= \frac{1}{T}\sum[\sqrt{T}I_n + \Delta_T]\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'[I_n + \frac{\Delta_T}{\sqrt{T}}]' \\
&= \frac{1}{T}\sum[\sqrt{T}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' + \Delta_T\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'][\frac{\Delta_T'}{\sqrt{T}} + I_n] \\
&= \frac{1}{T}\sum[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\Delta_T' + \Delta_T\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\frac{\Delta_T'}{\sqrt{T}} + \sqrt{T}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' + \Delta_T\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] \\
&= \underbrace{\frac{1}{\sqrt{T}}\sum\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'}_{\overset{\lim_{T\to\infty}}{=}A} + \underbrace{\frac{1}{T}\sum\Delta_T\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'}_{\overset{\lim_{T\to\infty}}{=}B} + \underbrace{\frac{1}{T}\sum\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\Delta_T'}_{\overset{\lim_{T\to\infty}}{=}B'} + \underbrace{\frac{1}{T}\sum\Delta_T\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\frac{\Delta_T'}{\sqrt{T}}}_{o_p(\sqrt{T})}
\end{aligned}
$$

□

The claim in Theorem 4.4.1 is achieved by standard arguments and an adjustment of Theorem 2.1. in Arnold et al. (2013).

Lemma A.2.2. *Let the assumptions from Lemma A.2.1 hold, then* $\boldsymbol{\alpha} = (A)_{i<j,i\neq j} = \left(\frac{1}{\sqrt{T}}\sum\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)_{i<j,i\neq j}$ *has expectation zero and the following covariance matrix*

$$
Cov[\boldsymbol{\alpha}] = \begin{pmatrix} \lim_{T\to\infty}\mathsf{Var}\ [\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{1t}\boldsymbol{\varepsilon}_{2t}] & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & \lim_{T\to\infty}\mathsf{Var}\ [\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_{(n-1)t}\boldsymbol{\varepsilon}_{nt}] \end{pmatrix}.
$$

Proof. The zero mean statement follows directly from the cross-sectional uncorrelatedness for every $t = 1, ..., T$. Furthermore, we observe

$$
\begin{aligned}
\mathrm{Cov}[\boldsymbol{\alpha}] &= \lim_{T\to\infty} \begin{pmatrix} \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\epsilon_{1t}\epsilon_{2t}] & \cdots & \frac{1}{T}\mathrm{Cov}[\sum\epsilon_{1t}\epsilon_{2t},\sum\epsilon_{(n-1)s}\epsilon_{ns}] \\ \vdots & \cdots & \vdots \\ \frac{1}{T}\mathrm{Cov}[\sum\epsilon_{(n-1)t}\epsilon_{nt},\sum\epsilon_{1s}\epsilon_{2s}] & \cdots & \mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\epsilon_{(n-1)t}\epsilon_{nt}] \end{pmatrix} \\
&= \begin{pmatrix} \lim\limits_{T\to\infty}\mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\epsilon_{1t}\epsilon_{2t}] & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & \lim\limits_{T\to\infty}\mathsf{Var}\,[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\epsilon_{(n-1)t}\epsilon_{nt}] \end{pmatrix} \\
&\in \mathbb{R}^{\frac{n(n-1)}{2}\times\frac{n(n-1)}{2}}.
\end{aligned}
$$

□

Lemma A.2.3. *Let the assumptions from Lemma A.2.1 hold and let $\{\boldsymbol{\varepsilon}_t\}_{t\in\{1,\ldots,T\}}$ be serially independent. Then $\boldsymbol{\alpha} = (A)_{i<j,i\neq j} = \underset{T\to\infty}{dlim}\left(\frac{1}{\sqrt{T}}\sum\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)_{i<j,i\neq j}$ is multivariate normally distributed with expectation zero and*

$$
Cov[(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t')_{i<j,i\neq j}] \;=\; Cov[\boldsymbol{\alpha}] \;=\; diag\left(\sigma_1^2\sigma_2^2,...,\sigma_{n-1}^2\sigma_n^2\right).
$$

Proof. The vector α can be rewritten as $\underset{T\to\infty}{\mathrm{dlim}}\sqrt{T}\left(\frac{1}{T}\sum\limits_{t=1}^{T}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)_{i<j,i\neq j}$. By Assumption 1.5 and the multivariate central limit theorem we obtain that $\boldsymbol{\alpha}$ is normally distributed with expectation zero. Since we assume uncorrelatedness in the cross-section for every $t = 1,...,T$, we have for $i \neq j \neq k \neq i$

$$
\begin{aligned}
\mathrm{Cov}[\varepsilon_{it}\varepsilon_{jt},\varepsilon_{it}\varepsilon_{jt}] &= \mathbb{E}[\varepsilon_{it}^2\varepsilon_{jt}^2] - 0 = \sigma_i^2\sigma_j^2, && \text{(A.8)} \\
\mathrm{Cov}[\varepsilon_{it}\varepsilon_{jt},\varepsilon_{it}\varepsilon_{kt}] &= \mathbb{E}[\varepsilon_{it}^2\varepsilon_{jt}\varepsilon_{kt}] - 0 \;=\; \mathbb{E}[\varepsilon_{it}^2]\mathbb{E}[\varepsilon_{jt}\varepsilon_{kt}] \;=\; 0. && \text{(A.9)}
\end{aligned}
$$

Thus, the covariance matrix for the limiting normal distribution is given by

$$
\begin{aligned}
\mathrm{Cov}[\boldsymbol{\alpha}] &= \begin{pmatrix} \mathrm{Cov}[\varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{1t}\varepsilon_{2t}] & \cdots & \mathrm{Cov}[\varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{(n-1)t}\varepsilon_{nt}] \\ \mathrm{Cov}[\varepsilon_{1t}\varepsilon_{3t}, \varepsilon_{1t}\varepsilon_{2t}] & \cdots & \mathrm{Cov}[\varepsilon_{1t}\varepsilon_{3t}, \varepsilon_{(n-1)t}\varepsilon_{nt}] \\ \vdots & \cdots & \vdots \\ \mathrm{Cov}[\varepsilon_{(n-1)t}\varepsilon_{nt}, \varepsilon_{1t}\varepsilon_{2t}] & \cdots & \mathrm{Cov}[\varepsilon_{(n-1)t}\varepsilon_{nt}, \varepsilon_{(n-1)t}\varepsilon_{nt}] \end{pmatrix} \\
&= \begin{pmatrix} \sigma_1^2\sigma_2^2 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & \sigma_{(n-1)}^2\sigma_n^2 \end{pmatrix} \in \mathbb{R}^{\frac{n(n-1)}{2} \times \frac{n(n-1)}{2}}.
\end{aligned}
$$

□

APPENDIX B

Testing for Relevant Dependence Change in Financial Data: A CUSUM Copula Approach

B.1 Assumptions

For the theoretical justification we need some slightly adjusted assumptions following Dette & Wied (2016):

A1) The marginals $F_i(\cdot)$ and its inverse $F_i^{-1}(\cdot)$ are assumed to be known for all $i \in \{1, ..., N\}$.

A2) Let $\{X_{T,1}, ..., X_{T,T}\}_{T\in\mathbb{N}}$ denote a triangular array of strong mixing random vectors and $\{\mathbf{U}_{T,1}, ..., \mathbf{U}_{T,T}\}_{T\in\mathbb{N}}$ its corresponding probability transform such that

$$\mathbf{U}_{T,1}, ..., \mathbf{U}_{T,\lfloor sT\rfloor} \sim C_1(\mathbf{u}) \; ; \quad \mathbf{U}_{T,\lfloor sT\rfloor+1}, ..., \mathbf{U}_{T,T} \sim C_2(\mathbf{u}).$$

A3) Consider the triangular array $\{\mathbf{U}_{T,j} \,|\, j = 1, ..., T\}_{T\in\mathbb{N}}$ and define for $1 \leq s \leq t$ the corresponding σ-field $\mathcal{F}_s^t(T) := \sigma(\{X_{T,j} \,|\, s \leq j \leq t\})$ generate by the random variable $\{\mathbf{U}_{T,j} \,|\, s \leq j \leq t\}$. For $m \in \mathbb{N}$ we denote by

$$\alpha(m) := \sup_{T\in\mathbb{N}} \sup_{1\leq k\leq T-m} \sup\{|P(A \cap B) - P(A)P(B)| \,|\, A \in \mathcal{F}_{m+k}^T(T), B \in \mathcal{F}_1^k(T)\},$$

the strong mixing coefficients of the triangular array $\{\mathbf{U}_{T,1}, ..., \mathbf{U}_{T,T}\}$ and assume that for some $\eta > 0$

$$\alpha(T) = \mathcal{O}(T^{-(1+\eta)})$$

as $T \to \infty$.

A4) For $l = 1, 2$ let $\{W_t(l)\}_{t\in\mathbb{Z}}$ denote sequences of strictly stationary processes, such that for each $T \in \mathbb{N}$

$$(\mathbf{U}_{T,1}, ..., \mathbf{U}_{T,\lfloor sT\rfloor}) \stackrel{d}{=} (W_1(1), ..., W_{\lfloor sT\rfloor}(1))$$
$$(\mathbf{U}_{T,\lfloor sT\rfloor+1}, ..., \mathbf{U}_{T,T}) \stackrel{d}{=} (W_1(2), ..., W_{T-\lfloor sT\rfloor}, (2))$$

where $\stackrel{d}{=}$ means equality in distribution. That means, there are two regimes $\{W_t(1)\}_{t\in\mathbb{Z}}$ and $\{W_t(2)\}_{t\in\mathbb{Z}}$ and the considered process switches from one regime to the other.

B.2 Derivation and Asymptotic Distribution of the Test Statistic

We impose the Assumptions given in Appendix B.1 to be valid. Then, the testing problem of no relevant change in the copula can be defined as follow:

$$H_0 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} \leq \Delta$$

versus the alternative

$$H_1 : \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} > \Delta,$$

where $\|.\|_{L^2}$ is the L^2-norm and $\Delta > 0$ fixed. For every $\mathbf{u} := (u_1, ..., u_N) \in [0, 1]^N$ and $t \in (0, 1)$ the CUSUM approach for detecting changes in the copula is then

$$\hat{\mathbb{U}}_T(t, \mathbf{u}) := t(1-t)\left(\frac{1}{\lfloor tT\rfloor}\sum_{i=1}^{\lfloor tT\rfloor} Z_i(\mathbf{u}) - \frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor tT\rfloor+1}^{T} Z_i(\mathbf{u})\right), \tag{B.1}$$

where $Z_i(\mathbf{u}) := \mathbb{1}\{F_1(X_{i1}) \leq u_1, ..., F_N(X_{iN}) \leq u_N\}$, $i = 1, ..., T$ is the vector of marginal distributions at time i where $F_j(\cdot)$ is the known j-th marginal cumulative distribution function for all $j = 1, ..., N$. Before we start the calculation we compute the expected value of some showing up sums. Since Z_i is Bernoulli distributed for

$i = 1, ..., N$ we have

$$\begin{aligned}
\mathbb{E}[Z_i(\mathbf{u})] &= P(F_1(X_{i1}) \leq u_1, ..., F_N(X_{iN}) \leq u_N) \\
&= P(X_{i1} \leq F_1^{-1}(u_1), ..., X_{iN} \leq F_N^{-1}(u_N)) \\
&= C(u_1, .., u_N).
\end{aligned}$$

Furthermore, we obtain from A3) and A4) $\mathbb{E}[Z_i(\mathbf{u})Z_j(\mathbf{u})] = C_i(\mathbf{u})C_j(\mathbf{u}) + o(1) \quad \forall i \neq j$. Due to readability we introduce the following abbreviations $C_i := C_i(\mathbf{u})$ and $Z_i := Z_i(\mathbf{u})$ for $i = 1, 2$. For fixed $s \in (0,1)$, we compute $\lim\limits_{T\to\infty} \mathbb{E}[\hat{\mathbb{U}}_T(t,\mathbf{u})]$. We first consider the case $t > s$

$$\begin{aligned}
\mathbb{E}[\hat{\mathbb{U}}_T(t,\mathbf{u})] &= t(1-t)\mathbb{E}\left[\frac{s}{t}\frac{1}{\lfloor sT\rfloor}\sum_{i=1}^{\lfloor sT\rfloor} Z_i + \frac{1}{\lfloor tT\rfloor}\sum_{i=\lfloor sT\rfloor+1}^{\lfloor tT\rfloor} Z_i - \frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor tT\rfloor+1}^{T} Z_i\right] \\
&= t(1-t)\left(\frac{s}{t}C_1 + \mathbb{E}\left[\frac{\lfloor tT\rfloor - (\lfloor sT\rfloor)}{\lfloor tT\rfloor}\frac{1}{\lfloor tT\rfloor - (\lfloor sT\rfloor)}\sum_{i=\lfloor sT\rfloor+1}^{\lfloor tT\rfloor} Z_i\right] - C_2\right) \\
&= t(1-t)\left(\frac{s}{t}C_1 + \frac{t-s}{t}C_2 - C_2\right) = s(1-t)\left(C_1 - C_2\right).
\end{aligned}$$

For $t \leq s$ we obtain

$$\begin{aligned}
\mathbb{E}[\hat{\mathbb{U}}_T(t,\mathbf{u})] &= t(1-t)\mathbb{E}\left[\frac{1}{\lfloor tT\rfloor}\sum_{i=1}^{\lfloor tT\rfloor} Z_i - \frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor tT\rfloor+1}^{\lfloor sT\rfloor} Z_i - \frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor sT\rfloor+1}^{T} Z_i\right] \\
&= t(1-t)\left(C_1 - \mathbb{E}\left[\frac{\lfloor sT\rfloor - (\lfloor tT\rfloor)}{T-\lfloor tT\rfloor}\frac{1}{\lfloor sT\rfloor - (\lfloor tT\rfloor)}\sum_{i=\lfloor sT\rfloor+1}^{\lfloor tT\rfloor} Z_i\right] - \frac{1-s}{1-t}\right) \\
&= t(1-t)\left(\frac{1-s}{1-t}C_1 - \frac{1-s}{1-t}C_2\right) = t(1-s)\left(C_1 - C_2\right).
\end{aligned}$$

Considering both cases yields:

$$\mathbb{E}[\hat{\mathbb{U}}_T(t,\mathbf{u})] = \begin{cases} s(1-t)\left(C_1 - C_2\right) & \text{for } t > s \\ t(1-s)\left(C_1 - C_2\right) & \text{for } t \leq s. \end{cases} \tag{B.2}$$

The aim is to lose the quantile and time dimension $\mathbf{u}$ and t, respectively. As an intermediate step we consider $\mathbb{E}[(\hat{\mathbb{U}}_T(t,\mathbf{u}))^2]$ that can be decomposed into three partial sums A, B, C with

$$A := \left(\frac{1}{\lfloor tT \rfloor} \sum_{i=1}^{\lfloor tT \rfloor} Z_i \right)^2 \tag{B.3}$$

$$B := \frac{1}{\lfloor tT \rfloor} \frac{1}{T - \lfloor tT \rfloor} \sum_{i=1}^{\lfloor tT \rfloor} Z_i \sum_{j=\lfloor tT \rfloor+1}^{T} Z_j \tag{B.4}$$

$$C := \left(\frac{1}{T - \lfloor tT \rfloor} \sum_{i=\lfloor tT \rfloor+1}^{T} Z_i \right)^2 . \tag{B.5}$$

Considering the case $t > s$ yields:

$$\begin{aligned}
A^{t>s}: \quad & \mathbb{E}\left[\left(\frac{1}{\lfloor tT \rfloor} \sum_{i=1}^{\lfloor tT \rfloor} Z_i \right)^2 \right] = \mathbb{E}\left[\left(\frac{1}{\lfloor tT \rfloor} \sum_{i=1}^{\lfloor sT \rfloor} Z_i + \frac{1}{\lfloor tT \rfloor} \sum_{i=\lfloor sT \rfloor+1}^{\lfloor tT \rfloor} Z_i \right)^2 \right] \\
=& \mathbb{E}\left[\left(\frac{1}{\lfloor tT \rfloor} \sum_{i=1}^{\lfloor sT \rfloor} Z_i \right)^2 \right] + 2\mathbb{E}\left[\left(\frac{1}{\lfloor tT \rfloor} \sum_{i=\lfloor sT \rfloor+1}^{\lfloor tT \rfloor} Z_i \right) \left(\frac{1}{\lfloor tT \rfloor} \sum_{i=1}^{\lfloor sT \rfloor} Z_i \right) \right] \\
& + \mathbb{E}\left[\left(\frac{1}{\lfloor tT \rfloor} \sum_{i=\lfloor sT \rfloor+1}^{\lfloor tT \rfloor} Z_i \right)^2 \right] \\
=& \frac{1}{\lfloor tT \rfloor^2} \left[\lfloor sT \rfloor (\frac{C_1(1-C_1)}{\lfloor sT \rfloor}) + (\lfloor sT \rfloor C_1)^2 \right] + 2\frac{s(t-s)}{t^2} C_1 C_2 \\
& + \frac{1}{\lfloor tT \rfloor^2} \left[\frac{\lfloor tT \rfloor - \lfloor sT \rfloor}{\lfloor tT \rfloor - \lfloor sT \rfloor} C_2(1 - C_2 + (\lfloor tT \rfloor - \lfloor sT \rfloor)^2 C_2^2 \right] + o(1) \\
=& \frac{s^2}{t^2} \left[\frac{C_1(1-C_1)}{T^2} + C_1^2 \right] + 2\frac{s(t-s)}{t^2} C_1 C_2 + \frac{C_2(1-C_2)}{\lfloor tT \rfloor^2} + \frac{(t-s)^2}{t^2} C_2^2 \\
& + o(1) \\
=& \frac{s^2}{t^2} C_1^2 + 2\frac{s}{t^2} C_1 (t-s) C_2 + \frac{(t-s)^2}{t^2} C_2^2 + o(1)
\end{aligned}$$

$$
\begin{aligned}
B^{t>s}: \quad & \mathbb{E}\left[\frac{1}{\lfloor tT \rfloor}\frac{1}{T-\lfloor tT \rfloor}\sum_{i=1}^{\lfloor tT \rfloor} Z_i \sum_{j=\lfloor tT \rfloor+1}^{T} Z_j\right] \\
&= \mathbb{E}\left[\frac{1}{\lfloor tT \rfloor}\frac{1}{T-\lfloor tT \rfloor}\sum_{i=1}^{\lfloor sT \rfloor} Z_i \sum_{j=\lfloor tT \rfloor+1}^{T} Z_j\right] \\
&\qquad + \mathbb{E}\left[\frac{1}{\lfloor tT \rfloor}\frac{1}{T-\lfloor tT \rfloor}\sum_{i=\lfloor sT \rfloor+1}^{\lfloor tT \rfloor} Z_i \sum_{j=\lfloor tT \rfloor+1}^{T} Z_j\right] \\
&= \frac{s}{t}C_1C_2 + \frac{t-s}{t}C_2^2 + o(1) \\
C^{t>s}: \quad & \mathbb{E}\left[\left(\frac{1}{T-\lfloor tT \rfloor}\sum_{i=\lfloor tT \rfloor+1}^{T} Z_i\right)^2\right] \\
&= \left(\frac{1}{T-\lfloor tT \rfloor}\right)^2 \left[\frac{T-\lfloor tT \rfloor}{T-\lfloor tT \rfloor}C_2(1-C_2) + [(T-\lfloor tT \rfloor)C_2]^2\right] \\
&= C_2^2 + o(1).
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
&\frac{1}{t^2(1-t)^2}\mathbb{E}[\hat{\mathbb{U}}_T(t,\mathbf{u})^2] \\
&= \mathbb{E}\left[\underbrace{\left(\frac{1}{\lfloor tT \rfloor}\sum_{i=1}^{\lfloor tT \rfloor} Z_i\right)^2}_{A^{t>s}} - \underbrace{\frac{2}{\lfloor tT \rfloor}\frac{1}{T-\lfloor tT \rfloor}\sum_{i=1}^{\lfloor tT \rfloor} Z_i \sum_{j=\lfloor tT \rfloor+1}^{T} Z_j}_{B^{t>s}} + \underbrace{\left(\frac{1}{T-\lfloor tT \rfloor}\sum_{i=\lfloor tT \rfloor+1}^{T} Z_i\right)^2}_{C^{t>s}}\right] \\
&= \frac{s^2}{t^2}C_1^2 + 2\frac{s}{t^2}C_1(t-s)C_2 + \frac{(t-s)^2}{t^2}C_2^2 - 2\left[\frac{s}{t}C_1 + \frac{t-s}{t}C_2\right]C_2^2 + C_2^2 + o(1) \\
&= \frac{s^2}{t^2}(C_1-C_2)^2 + o(1)
\end{aligned}
$$

Considering the $t \le s$ yields:

$$
A^{t\le s}: \quad \mathbb{E}\left[\left(\frac{1}{T}\sum_{i=1}^{\lfloor tT \rfloor} Z_i\right)^2\right] = \frac{1}{\lfloor tT \rfloor^2}\left(\frac{\lfloor tT \rfloor}{\lfloor tT \rfloor}[C_1(1-C_1)] + (\lfloor tT \rfloor C_1)^2\right) = C_1^2 + o(1)
$$

$$
\begin{aligned}
B^{t\leq s}: \quad & \mathbb{E}\left[\frac{1}{\lfloor tT\rfloor}\frac{1}{T-\lfloor tT\rfloor}\sum_{i=1}^{\lfloor tT\rfloor} Z_i \sum_{j=\lfloor tT\rfloor+1}^{T} Z_j\right] \\
=&\frac{1}{\lfloor tT\rfloor}\frac{1}{T-\lfloor tT\rfloor}\left(\mathbb{E}\left[\sum_{i=1}^{\lfloor tT\rfloor} Z_i \sum_{j=\lfloor tT\rfloor+1}^{\lfloor sT\rfloor} Z_j\right] + \mathbb{E}\left[\sum_{i=1}^{\lfloor tT\rfloor} Z_i \sum_{j=\lfloor sT\rfloor+1}^{T} Z_j\right]\right) \\
=&\left[\frac{s-t}{1-t}C_1 + \frac{1-s}{1-t}C_2\right] C_1 + o(1) \\
C^{t\leq s}: \quad & \mathbb{E}\left[\left(\frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor tT\rfloor}^{T} Z_i\right)^2\right] \\
=&\mathbb{E}\left[\left(\frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor tT\rfloor}^{\lfloor sT\rfloor} Z_i + \frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor sT\rfloor+1}^{T} Z_i\right)^2\right] \\
=&\mathbb{E}\left[\left(\frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor tT\rfloor}^{\lfloor sT\rfloor} Z_i\right)^2\right] \\
&+ 2\mathbb{E}\left[\left(\frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor tT\rfloor}^{\lfloor sT\rfloor} Z_i\right)\left(\frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor sT\rfloor+1}^{T} Z_i\right)\right] \\
&+ \mathbb{E}\left[\left(\frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor sT\rfloor+1}^{T} Z_i\right)^2\right] \\
=&\frac{1}{(T-\lfloor tT\rfloor)^2}\left(\frac{\lfloor sT\rfloor-\lfloor tT\rfloor}{\lfloor sT\rfloor-\lfloor tT\rfloor}[C_1(1-C_1)] + [(\lfloor sT\rfloor-\lfloor tT\rfloor)C_1]^2\right) \\
&+ 2\frac{s-t}{1-t}C_1\frac{1-s}{1-t}C_2 \\
&+ \frac{1}{(T-\lfloor tT\rfloor)^2}\left(\frac{T-\lfloor sT\rfloor}{T-\lfloor sT\rfloor}[C_2(1-C_2)] + [(T-\lfloor sT\rfloor)C_2]^2\right) + o(1) \\
=&\frac{(s-t)^2}{(1-t)^2}C_1^2 + +2\frac{s-t}{1-t}C_1\frac{1-s}{1-t}C_2 + \frac{(1-s)^2}{(1-t)^2}C_2^2 + o(1)
\end{aligned}
$$

For the expression $E[\hat{\mathbb{U}}_T(t,\mathbf{u})^2]$ we have

$$
\begin{aligned}
&\frac{1}{t^2(1-t)^2}\mathbb{E}[\hat{\mathbb{U}}_T(t,\mathbf{u})^2] \\
&= \mathbb{E}\left[\underbrace{\left(\frac{1}{\lfloor tT\rfloor}\sum_{i=1}^{\lfloor tT\rfloor} Z_i\right)^2}_{A^{t\leq s}} -2\underbrace{\frac{1}{\lfloor tT\rfloor}\frac{1}{T-\lfloor tT\rfloor}\sum_{i=1}^{\lfloor tT\rfloor} Z_i \sum_{j=\lfloor tT\rfloor+1}^{T} Z_j}_{B^{t\leq s}} + \underbrace{\left(\frac{1}{T-\lfloor tT\rfloor}\sum_{i=\lfloor tT\rfloor+1}^{T} Z_i\right)^2}_{C^{t\leq s}}\right]
\end{aligned}
$$

$$= \frac{(s-t)^2}{(1-t)^2}C_1^2 + 2\frac{s-t}{1-t}C_1\frac{1-s}{1-t}C_2 + \frac{(1-s)^2}{(1-t)^2}C_2^2 - 2\left[\frac{s-t}{1-t}C_1 + \frac{1-s}{1-t}C_2\right]C_1 + C_1^2 + o(1)$$
$$= \frac{(1-s)^2}{(1-t)^2}(C_1 - C_2)^2 + o(1)$$

Combining the previous calculations for $t > s$, $t \leq s$ and with the help of Fubini we obtain

$$L(t) := \lim_{T\to\infty} \mathbb{E}[\|\hat{\mathbb{U}}_T(t,\mathbf{u})\|_{L^2}^2] = \begin{cases} s^2(1-t)^2\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2, & t > s \\ (1-s)^2t^2\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2, & t \leq s. \end{cases}$$

By integrating out t a straightforward calculation yields

$$\int_0^1 L(t)dt = \frac{s^2(1-s)^2}{3}\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}^2. \tag{B.6}$$

The next theorem provides the limiting distribution of the empirical centred counterpart $\hat{L}_T(t) := \|\hat{\mathbb{U}}_T(t,\mathbf{u})\|_{L^2}^2$

Theorem B.2.1. *Under Assumptions A1)-A4)*

$$\sqrt{T}\left(\int_0^1 \hat{L}_T(t)dt - \frac{1}{3}s^2(1-s)^2||C_1(\mathbf{u}) - C_2(\mathbf{u})||_{L^2}^2\right) \xrightarrow{d} N(0, \sigma_{C_1,C_2,s}^2), \tag{B.7}$$

with $\sigma_{C_1,C_2,s}^2 = 4\int_0^1\int_0^1 \mathbb{E}\left[< \mathbb{U}(t_1,\mathbf{u}), A(t_1,\mathbf{u}) >_{L^2}< \mathbb{U}(t_2,\mathbf{u}), A(t_2,\mathbf{u}) >_{L^2}\right] dt_1dt_2$ *and* $< \cdot,\cdot >_{L^2}$ *the* L^2 *inner product.*

Proof. See Appendix B.3 $\square$

Due to the high computational effort in high dimensions using the L^2-norm it could be reasonable to only test for specific quantiles (points) $\mathbf{q}$ in the copula. So similar to the L^2-norm testing we can test on fixed points $\mathbf{q} = (q_1, \dots, q_N)'$ in the copula, using the previous notation and considering a constant functions $g := C(\mathbf{q})$, where $C(\mathbf{q})$ is the copula value at some fixed quantile $\mathbf{q}$.

Corollary B.2.2. *Under Assumptions A1)-A4)*

$$\sqrt{T}\left(\int_0^1 \hat{L}_T^{\mathbf{q}}(t)dt - \frac{1}{3}s^2(1-s)^2|C_1(\mathbf{q}) - C_2(\mathbf{q})|^2\right) \xrightarrow{d} N(0, \sigma_{C_1,C_2,s,\mathbf{q}}^2), \tag{B.8}$$

with $\hat{L}_T^{\mathbf{q}}(t) := (\hat{\mathbb{U}}_T(t,\mathbf{q}))^2$ *and*

$$\sigma_{C_1,C_2,s,\mathbf{q}}^2 := 4\int_0^1\int_0^1 \mathbb{E}\left[\mathbb{U}(t_1,\mathbf{q}) \cdot A(t_1,\mathbf{q}) \cdot \mathbb{U}(t_2,\mathbf{q}) \cdot A(t_2,\mathbf{q})\right] dt_1dt_2$$

for fixed $\mathbf{q} \in [0,1]^N$.

$\hat{L}_T^{\mathbf{q}}(t)$ and $\sigma^2_{C_1,C_2,s,\mathbf{q}}$ are called the quantile version of $\hat{L}_T(t)$ and $\sigma^2_{C_1,C_2,s}$, respectively. The next Lemma shows that the test holds the size level and has considerable power.

Lemma B.2.3. *The test*

$$\hat{\kappa}_T \geq \frac{1}{3}s^2(1-s)^2\Delta^2 + \frac{k_{1-\alpha}(s)}{\sqrt{T}} \tag{B.9}$$

is a consistent asymptotic α test for all $s > 0$, where $k_{1-\alpha}(s)$ is the $(1-\alpha)$-quantile of the limiting normal distribution given in (B.7) *and* $\hat{\kappa}_T = \int\limits_0^1 \hat{L}_T(t)dt$.

Proof. Suppose $\delta := \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} \leq \Delta$. Then

$$\begin{aligned}
&P_\delta(\hat{\kappa}_T \geq \frac{1}{3}s^2(1-s)^2\Delta^2 + \frac{k_{1-\alpha}(s)}{\sqrt{T}}) \\
&= P(\sqrt{T}(\hat{\kappa}_T - \frac{1}{3}s^2(1-s)^2\delta^2) \geq \sqrt{T}\frac{1}{3}s^2(1-s)^2(\Delta^2 - \delta^2) + k_{1-\alpha}(s)) \\
&\leq P(\sqrt{T}(\hat{\kappa}_T - \frac{1}{3}s^2(1-s)^2\delta^2) \geq k_{1-\alpha}(s)) \\
&\underset{T\to\infty}{\longrightarrow} 1 - (1-\alpha) = \alpha.
\end{aligned}$$

Otherwise, if $\delta > \Delta$

$$\begin{aligned}
&P_\delta(\hat{\kappa}_T \geq \frac{1}{3}s^2(1-s)^2\Delta^2 + \frac{k_{1-\alpha}(s)}{\sqrt{T}}) \\
&= P(\sqrt{T}((\hat{\kappa}_T - \frac{1}{3}s^2(1-s)^2\delta^2) \geq \underbrace{\sqrt{T}\frac{1}{3}s^2(1-s)^2(\Delta^2 - \delta^2)}_{<0} + k_{1-\alpha}(s)) \\
&= 1 - P(\sqrt{T}(\hat{\kappa}_T - \frac{1}{3}s^2(1-s)^2\delta^2) < \sqrt{T}\frac{1}{3}s^2(1-s)^2(\Delta^2 - \delta^2) + k_{1-\alpha}(s)) \\
&\underset{T\to\infty}{\longrightarrow} 1 - 0 = 1.
\end{aligned}$$

□

The test given in equation (B.9) is an exact level α test if Δ is chosen as the copula difference $\delta = ||C_1(\mathbf{u}) - C_2(\mathbf{u})||_{L^2}$. Otherwise the size is smaller than α.

B.3 Proof of Theorem 1

We execute the proof of Theorem 1 stepwise. First, we start to consider only one partial sum of the process $\hat{\mathbb{U}}_T(\cdot,\cdot)$, i.e.

$$\hat{\mathbb{C}}_T(t,\mathbf{u}) := \frac{1}{T}\sum_{i=1}^{\lfloor tT\rfloor} Z_i(\mathbf{u}). \tag{B.10}$$

Second, by means of the continuous mapping theorem we obtain the limiting distribution of the process $\hat{\mathbb{U}}_T(\cdot,\cdot)$ and can then finally derive the limiting distribution given in Theorem B.2.1. Again, for the computation of the expectation of $\hat{\mathbb{C}}_T(\cdot,\cdot)$ we have to distinguish two cases, i.e. either $t \le s$ or $t > s$. If $t \le s$, we have $\lim_{T\to\infty}\mathbb{E}[\hat{\mathbb{C}}_T(t,\mathbf{u})] = tC_1(\mathbf{u})$. For $t > s$ a straightforward calculation yields

$$\begin{aligned}\mathbb{E}[\hat{\mathbb{C}}_T(t,\mathbf{u})] &= \mathbb{E}\left[\frac{1}{T}\sum_{i=1}^{\lfloor sT\rfloor} Z_i(\mathbf{u}) + \frac{1}{T}\sum_{i=\lfloor sT\rfloor+1}^{\lfloor tT\rfloor} Z_i(\mathbf{u})\right]\\ &= sC_1(\mathbf{u}) + \frac{\lfloor tT\rfloor - \lfloor sT\rfloor}{T}C_2(\mathbf{u}) = sC_1(\mathbf{u}) + (t-s)C_2(\mathbf{u}) + o(1).\end{aligned}$$

Thus, the expectation of the partial sum $\hat{\mathbb{C}}_T(\cdot,\cdot)$ is given by

$$E_{C_1,C_2,s}(t,\mathbf{u}) := \lim_{T\to\infty}\mathbb{E}[\hat{\mathbb{C}}_T(t,\mathbf{u})] = (s\wedge t)C_1(\mathbf{u}) + (t-s)_+C_2(\mathbf{u}). \tag{B.11}$$

With the expectation (B.11) we derive the asymptotic distribution of the centred partial sum process (B.10), which leads to the following theorem.

Theorem B.3.1. *Let Assumptions A1)-A4) hold. Then, a standardized version of the process* $\{\hat{\mathbb{C}}_T(t,\mathbf{u})\}_{t\in(0,1),\mathbf{u}\in[0,1]^N}$ *converges weakly in* $\ell^\infty((0,1)\times[0,1]^N)$*, i.e.*

$$\sqrt{T}\left\{\hat{\mathbb{C}}_T(t,\mathbf{u}) - E_{C_1,C_2,s}(t,\mathbf{u})\right\}_{t\in(0,1),\mathbf{u}\in[0,1]^N} \overset{d}{\Rightarrow} \{\mathbb{G}_{C_1,C_2,s}(t,\mathbf{u})\}_{t\in(0,1),\mathbf{u}\in[0,1]^N}.$$

Here, $\mathbb{G}_{C_1,C_2,s}$ *denotes a centered Gaussian process with covariance kernel*

$$\mathbb{E}[\mathbb{G}_{C_1,C_2,s}(t_1,\mathbf{u}_1)\mathbb{G}_{C_1,C_2,s}(t_2,\mathbf{u}_2)] \tag{B.12}$$
$$= (t_1\wedge t_2\wedge s)k_1(\mathbf{u}_1,\mathbf{u}_2) + (t_1\wedge t_2 - s)_+k_2(\mathbf{u}_1,\mathbf{u}_2), \tag{B.13}$$

and the kernels k_1 *and* k_2 *are defined by*

$$k_l(\mathbf{u}_1,\mathbf{u}_2) = \sum_{i\in\mathbb{Z}} Cov[\mathbb{1}\{W_0(l)\le\mathbf{u}_1\},\mathbb{1}\{W_i(l)\le\mathbf{u}_2\}], \quad l=1,2. \tag{B.14}$$

Proof. Consider

$$
\begin{aligned}
&\hat{\mathbb{C}}_T(t,u) - \mathbb{E}_{C_1,C_2,s}[t,\mathbf{u}] \\
&= \frac{1}{T}\sum_{i=1}^{\lfloor tT\rfloor} Z_i(\mathbf{u}) - [(t\wedge s)C_1(\mathbf{u}) + (t-s)_+ C_2(\mathbf{u})] + o_P(\frac{1}{\sqrt{T}}) \\
&= \underbrace{\frac{1}{T}\sum_{i=1}^{\lfloor T(s\wedge t)\rfloor} [Z_i(\mathbf{u}) - C_1(\mathbf{u})]}_{X_T^{(1)}(t,\mathbf{u}) := \sum_{i=1}^{\lfloor T(s\wedge t)\rfloor} Y_{T,i}(\mathbf{u})} + \underbrace{\mathbb{1}\{t>s\}\frac{1}{T}\sum_{i=\lfloor T(s\wedge t)\rfloor+1}^{\lfloor tT\rfloor} [Z_i(\mathbf{u}) - C_2(\mathbf{u})]}_{X_T^{(2)}(t,\mathbf{u}) := \mathbb{1}\{t>s\}\sum_{i=\lfloor T(s\wedge t)\rfloor+1}^{\lfloor tT\rfloor} Y_{T,i}(\mathbf{u})} + o_P(\frac{1}{\sqrt{T}})
\end{aligned}
$$

with $Y_{T,i}(\mathbf{u}) := \mathbb{1}\{i \le \lfloor sT\rfloor\}\dfrac{Z_i(\mathbf{u}) - C_1(\mathbf{u})}{T} + \mathbb{1}\{i > \lfloor sT\rfloor\}\dfrac{Z_i(\mathbf{u}) - C_2(\mathbf{u})}{T}$

Then it follows by Bücher et al. (2014) for $T \to \infty$

1. $\{\sqrt{T}X_T^{(1)}(t,\mathbf{u})\}_{t\in[0,1],\mathbf{u}\in[0,1]^n} \overset{d}{\Longrightarrow} \mathbb{G}(t\wedge s,\mathbf{u})$
2. $\{\sqrt{T}X_T^{(2)}(t,\mathbf{u})\}_{t\in[0,1],\mathbf{u}\in[0,1]^n} \overset{d}{\Longrightarrow} \mathbb{G}(t,\mathbf{u}) - \mathbb{G}(t\wedge s,\mathbf{u})$

where $\mathbb{G}(\cdot,\cdot)$ are tight centred Gaussian processes with covariance function

$$\mathrm{Cov}[\mathbb{G}(t_1\wedge s,\mathbf{u}_1),\mathbb{G}(t_2\wedge s,\mathbf{u}_2)] = (t_1\wedge t_2\wedge s)k_1(\mathbf{u}_1,\mathbf{u}_2) \tag{B.15}$$

and

$$
\begin{aligned}
&\mathrm{Cov}[\mathbb{G}(t_1,\mathbf{u}_1) - \mathbb{G}(t_1\wedge s,\mathbf{u}_1),\mathbb{G}(t_2,\mathbf{u}_2) - \mathbb{G}(t_2\wedge s,\mathbf{u}_2)] \\
&= \mathrm{Cov}[\mathbb{G}(t_1,\mathbf{u}_1,\mathbb{G}(t_2,\mathbf{u}_2)] - \mathrm{Cov}[\mathbb{Z}(t_1,\mathbf{u}_1),\mathbb{G}(t_2\wedge s,\mathbf{u}_2)] - \\
&\qquad \mathrm{Cov}[\mathbb{G}(t_1\wedge s,\mathbf{u}_1),\mathbb{G}(t_2,\mathbf{u}_2)] + \mathrm{Cov}[\mathbb{G}(t_1\wedge s,\mathbf{u}_1),\mathbb{G}(t_2\wedge s,\mathbf{u}_2)] \\
&= (t_1\wedge t_2)k_2(\mathbf{u}_1,\mathbf{u}_2) - (t_1\wedge t_2\wedge s)k_2(\mathbf{u}_1,\mathbf{u}_2) \\
&\quad - (t_1\wedge t_2\wedge s)k_2(\mathbf{u}_1,\mathbf{u}_2) + (t_1\wedge t_2\wedge s)k_2(\mathbf{u}_1,\mathbf{u}_2) \\
&= (t_1\wedge t_2 - t_1\wedge t_2\wedge s)k_2(\mathbf{u}_1,\mathbf{u}_2) \\
&= (t_1\wedge t_2 - s)_+ k_2(\mathbf{u}_1,\mathbf{u}_2).
\end{aligned}
$$

Thus, the composition $\sqrt{T}X_T := \sqrt{T}\left(X_T^{(1)} + X_T^{(2)}\right)$ is asymptotically tight (cf. van der Vaart & Wellner, 1996, Section 1.5). In order to prove convergence in distribution of $\sqrt{T}X_T$ it remains to establish the weak convergence of the finite dimensional distributions. Therefore, we use the Cramér-Wold-device and show for all sequences

$(t_1, \mathbf{u}_1), ..., (t_n, \mathbf{u}_n) \in [0,1] \times [0,1]^n$

$$\sqrt{T}\{\sum_{i=1}^{k} a_j X_T(t_j, \mathbf{u}_j)\} \overset{d}{\Longrightarrow} \sum_{j=1}^{k} a_j \mathbb{G}_{C_1,C_2,s}(t_j, \mathbf{u}_j) \tag{B.16}$$

with $\alpha_1, ..., \alpha_k \in \mathbb{R}$ and $\mathbb{G}_{C_1,C_2,s}$ is the Gaussian process defined in Theorem B.3.1. Now, we restrict ourselves to the case $k = 2$ and begin with the calculation of the covariance of $X_T^{(1)}(t_1, u_1)$ and $X_T^{(2)}(t_2, u_2)$. Therefore, we consider four different cases.
$\underline{t_1 \leq t_2 \leq s}$:

$$T\mathrm{Cov}[X_T^{(l)}(t_1, \mathbf{u}_1), X_T^{(l)}(t_2, \mathbf{u}_2)] \overset{T\to\infty}{\to} \begin{cases} (t_1 \wedge t_2 \wedge s)k_1(\mathbf{u}_1, \mathbf{u}_2) & \text{if } l = 1 \\ 0 & \text{if } l = 2. \end{cases}$$

$\underline{s \leq t_1 \leq t_2}$:

$$T\mathrm{Cov}[X_T^{(l)}(t_1, \mathbf{u}_1), X_T^{(l)}(t_2, \mathbf{u}_2)] \overset{T\to\infty}{\to} \begin{cases} tk_1(\mathbf{u}_1, \mathbf{u}_2) & \text{if } l = 1 \\ (t_1 \wedge t_2 - s)_+ k_2(\mathbf{u}_1, \mathbf{u}_2) & \text{if } l = 2 \end{cases}$$

$\underline{t_1 < s \leq t_2}$:

$$\begin{aligned} T|\mathrm{Cov}[X_T^{(l)}(t_1, \mathbf{u}_1), X_T^{(l)}(t_2, \mathbf{u}_2)]| &= T|\mathrm{Cov}[\sum_{j=1}^{\lfloor t_1 T \rfloor} Y_{T,i}(t_2, \mathbf{u}_2), \sum_{j=\lfloor sT \rfloor+1}^{\lfloor t_2 T \rfloor} Y_{T,i}(t_2, \mathbf{u}_2)]| \\ &= \mathcal{O}(\frac{1}{T^{\eta+1}}) = \mathcal{O}(\frac{1}{T^{\eta}}) = o(1) \end{aligned}$$

for all $\eta > 0$.
In the case where $t_1 = s \leq t_2$ we use a sequence ϵ_T such that $\epsilon_T T \to \infty$ and $\epsilon_T^2 T \to 0$ and obtain by the same argument of strong mixing

$$\begin{aligned} &T|\mathrm{Cov}[X_T^{(l)}(t_1, \mathbf{u}_1), X_T^{(l)}(t_2, \mathbf{u}_2)]| \\ &= T|\mathrm{Cov}[\sum_{i=1}^{\lfloor T(s-\epsilon_T) \rfloor} \mathbb{Y}_{T,i}(t_1, \mathbf{u}_1) + \sum_{i=\lfloor T(s-\epsilon_T) \rfloor+1}^{\lfloor sT \rfloor} \mathbb{Y}_{T,i}(t_1, \mathbf{u}_1), \sum_{i=\lfloor sT \rfloor+1}^{\lfloor T(s+\epsilon_T) \rfloor} \mathbb{Y}_{T,i}(t_2, \mathbf{u}_2) \\ &\quad + \sum_{i=\lfloor T(s+\epsilon_T) \rfloor+1}^{\lfloor t_2 T \rfloor} \mathbb{Y}_{T,i}(t_2, \mathbf{u}_2)]| \\ &= \mathcal{O}(\frac{1}{(\epsilon_T)T^{\eta}}) + \mathcal{O}(T\epsilon_T^2) = o(1) \end{aligned}$$

$$
\begin{aligned}
\sigma^2 &= \lim_{T\to\infty} \mathbb{V}[\sqrt{T}\sum_{j=1}^{2} \alpha_j X_T(t_j, \mathbf{u}_j)] \\
&= \lim_{T\to\infty} \mathbb{V}[\alpha_1(X_T^{(1)}(t_1,\mathbf{u}_1) + X_T^{(2)}(t_1,\mathbf{u}_1)) + \alpha_2(X_T^{(1)}(t_2,\mathbf{u}_2) + X_T^{(2)}(t_2,\mathbf{u}_2))] \\
&= \lim_{T\to\infty} T\{\alpha_1^2 \mathrm{Cov}[X_T^{(1)}(t_1,\mathbf{u}_1), X_T^{(1)}(t_1,\mathbf{u}_1)] \\
&\quad + 2\alpha_1\alpha_2 \mathrm{Cov}[X_T^{(1)}(t_1,\mathbf{u}_1), X_T^{(1)}(t_2,\mathbf{u}_2)] \\
&\quad + \alpha_1^2 \mathrm{Cov}[X_T^{(2)}(t_1,\mathbf{u}_1), X_T^{(2)}(t_1,\mathbf{u}_1)] + 2\alpha_1\alpha_2 \mathrm{Cov}[X_T^{(2)}(t_1,\mathbf{u}_1), X_T^{(2)}(t_2,\mathbf{u}_2)] \\
&\quad + \alpha_2^2 \mathrm{Cov}[X_T^{(1)}(t_2,\mathbf{u}_2), X_T^{(1)}(t_2,\mathbf{u}_2)] + \alpha_2^2 \mathrm{Cov}[X_T^{(2)}(t_2,\mathbf{u}_2), X_T^{(2)}(t_2,\mathbf{u}_2)]\} \\
&= \alpha_1^2\left((t_1\wedge s)k_1(\mathbf{u}_1,\mathbf{u}_1) + (t_1 - s)_+ k_2(\mathbf{u}_1,\mathbf{u}_1)\right) \\
&\quad + a_2^2\left((t_2\wedge s)k_2(\mathbf{u}_2,\mathbf{u}_2) + (t_2 - s)_+ k_2(\mathbf{u}_2,\mathbf{u}_2)\right) \\
&\quad + 2\alpha_1\alpha_2\left((t_1\wedge t_2\wedge s)k_1(\mathbf{u}_1,\mathbf{u}_2) + (t_1\wedge t_2 - s)_+ k_2(\mathbf{u}_1,\mathbf{u}_2)\right) \\
&= \mathbb{V}[\alpha_1\mathbb{G}_{C_1,C_2,s}(t_1,\mathbf{u}_1) + \alpha_2\mathbb{G}_{C_1,C_2,s}(t_2,\mathbf{u}_2)]
\end{aligned}
\tag{B.17}
$$

with $\mathbb{E}[\mathbb{G}_{C_1,C_2,s}(t_1,\mathbf{u}_1)\mathbb{G}_{C_1,C_2,s}(t_2,\mathbf{u}_2)] = (t_1\wedge t_2\wedge s)k_1(\mathbf{u}_1,\mathbf{u}_2) + (t_1\wedge t_2 - s)_+ k_2(\mathbf{u}_1,\mathbf{u}_2)$ where the kernels for $i = 1, 2$ are given by

$$
k_i(\mathbf{u}_1,\mathbf{u}_2) = \sum_{k\in\mathbb{Z}} \mathrm{Cov}[\mathbb{1}\{W_0(i)\le \mathbf{u}_1\}, \mathbb{1}\{W_k(i)\le \mathbf{u}_2\}]
$$

In order to prove asymptotic normality of $\sqrt{T}\sum_{j=1}^{2}\alpha_j X_T(t_j,\mathbf{u}_j)$ we introduce the notation

$$
\mathcal{T}_T := \frac{\sqrt{T}}{\sigma}\sum_{j=1}^{2}\alpha_j X_T(t_j,\mathbf{u}_j) = \sum_{j=1}^{T} S_{T,j} + o_P(1)
$$

with

$$
\begin{aligned}
S_{T,j} = &\frac{\alpha_1\mathbb{1}\{j\le\lfloor t_1 T\rfloor\}}{\sigma\sqrt{T}}\left(\mathbb{1}\{\mathbf{U}_j\le\mathbf{u}_1\} - \mathbb{E}_{C_1,C_2,t}(t_1,\mathbf{u}_1)\right) \\
&+\frac{\alpha_2\mathbb{1}\{j\le\lfloor t_2 T\rfloor\}}{\sigma\sqrt{T}}\left(\mathbb{1}\{\mathbf{U}_j\le\mathbf{u}_2\} - \mathbb{E}_{C_1,C_2,t}(t_2,\mathbf{u}_2)\right)
\end{aligned}
$$

and we use a central limit theorem for triangular arrays of strong mixing random variables (see Liebscher, 1996, Theorem 2.1 with $p = \infty$). From the previous discussion it follows that $\lim_{T\to\infty}\mathbb{E}[\mathcal{T}_T^2]=1$ and thus, we have

$$\lim_{T\to\infty}\sum_{j=1}^{T}(\operatorname*{ess\,sup}_{\omega\in\Omega}[\,|\,S_{T,j}\,|\,\mathbb{1}\{|S_{T,j}|>\epsilon\}])^2=0 \quad \text{a.s..}$$

Similarly, it follows that the condition

$$\lim_{T\to\infty}\sum_{j=1}^{T}(\operatorname*{ess\,sup}_{\omega\in\Omega}\,|\,S_{T,j}\,|\,)^2\leq \text{const} \quad \text{a.s..}$$

of Theorem 2.1 in Liebscher (1996) is also satisfied. Therefore this result shows that

$$\sqrt{T}\sum_{j=1}^{2}\alpha_j X_n(t_j,\mathbf{u}_j)=\frac{\sigma\mathcal{T}_T}{\sqrt{\mathbb{E}[\mathcal{T}_T^2]}}\overset{D}{\Longrightarrow}N(0,\sigma^2)$$

where the asymptotic variance σ^2 is defined in (B.17). This proves the convergence of the finite dimensional distributions and completes the proof of the theorem. $\square$

Now, we can follow the asymptotic distribution of the centered $\hat{\mathbb{U}}_T(t,\mathbf{u})$, by using the continuous mapping theorem with $\hat{\mathbb{U}}_T(t,\mathbf{u})=\hat{\mathbb{C}}_T(t,\mathbf{u})-t\hat{\mathbb{C}}_T(1,\mathbf{u})$.

Corollary B.3.2. *Under assumptions A1)-A4) we receive for $t\in(0,1)$ and $\mathbf{u}\in[0,1]^N$*

$$\sqrt{T}\left(\hat{\mathbb{U}}_T(t,\mathbf{u})-\mathbb{U}(t,\mathbf{u})\right)\overset{d}{\Longrightarrow}\{A(t,\mathbf{u})\}_{t\in(0,1),\mathbf{u}\in[0,1]^N}, \tag{B.18}$$

where $\hat{\mathbb{U}}_T(t,\mathbf{u})=\hat{\mathbb{C}}_T(t,\mathbf{u})-t\hat{\mathbb{C}}_T(1,\mathbf{u})$, $\mathbb{U}(t,\mathbf{u})=\mathbb{E}_{C_1,C_2,s}(t,\mathbf{u})-t\mathbb{E}_{C_1,C_2,s}(1,\mathbf{u})$ and $A(t,\mathbf{u})=\mathbb{G}_{C_1,C_2,s}(t,\mathbf{u})-t\mathbb{G}_{C_1,C_2,s}(1,\mathbf{u})$ with covariance kernel

$$a_{C_1,C_2,s}(t_1,\mathbf{u}_1,t_2,\mathbf{u}_2)=\mathbb{E}\left[A(t_1,\mathbf{u}_1)A(t_2,\mathbf{u}_2)\right]. \tag{B.19}$$

Now, we can complete the proof for Theorem B.2.1. By Corollary B.3.2 we have for $t\in(0,1)$ and $\mathbf{u}\in[0,1]^N$

$$\sqrt{T}\left(\hat{\mathbb{U}}_T(t,\mathbf{u})-\mathbb{U}(t,\mathbf{u})\right)\overset{d}{\Longrightarrow}A(t,\mathbf{u}).$$

Thus, for every inner product space we have we can rewrite $\hat{L}_T(t)-L(t)$ for $t\in(0,1)$

as

$$\hat{L}_T(t) - L(t) = ||\hat{\mathbb{U}}_T(t,\mathbf{u}) - \mathbb{U}(t,\mathbf{u})||^2 + 2 < \mathbb{U}(t,\mathbf{u}), \hat{\mathbb{U}}_T(t,\mathbf{u}) - \mathbb{U}(t,\mathbf{u}) >_{L^2} .$$

Then, by Corollary B.3.2 and the consistency of $\hat{\mathbb{U}}_T(\cdot)$ in (B.2) we get

$$\sqrt{T}\left(\hat{L}_T(t) - L(t)\right) \overset{\mathrm{d}}{\Longrightarrow} 2 < \mathbb{U}(t,\mathbf{u}), A(t,\mathbf{u}) >_{L^2} .$$

Thus, with the help of the continuous mapping theorem we receive

$$\begin{aligned} &\sqrt{T}\left(\int_0^1 \hat{L}_T(t)dt - \int_0^1 L(t)dt\right) \overset{d}{\to} \int_0^1 2 < \mathbb{U}(t,\mathbf{u}), A(t,\mathbf{u}) >_{L^2} dt =: Q \\ \Leftrightarrow &\sqrt{T}\left(\int_0^1 \hat{L}_T(t)dt - \frac{1}{3}s^2(1-s)^2||C_1(\mathbf{u}) - C_2(\mathbf{u})||^2\right) \overset{d}{\longrightarrow} Q, \end{aligned}$$

where the random variable Q is normally distributed $N(0, \sigma^2_{C_1,C_2,s})$ with variance term

$$\sigma^2_{C_1,C_2,s} = 4\int_0^1 \int_0^1 \mathbb{E}\left[< \mathbb{U}(t_1,\mathbf{u}), A(t_1,\mathbf{u}) >_{L^2} < \mathbb{U}(t_2,\mathbf{u}), A(t_2,\mathbf{u}) >_{L^2}\right] dt_1 dt_2.$$

B.4 Covariance Bootstrap

Another approach next to the full bootstrap is to estimate the variance term of the limiting normal distribution[1]. Therefore, we have to estimate the covariance of the centred Gaussian process $d_{C_1,C_2,s}(t_1,t_2) = \mathbb{E}[D_{C_1,C_2,s}(t_1), D_{C_1,C_2,s}(t_2)]$ by using re-sampling, cf. Theorem B.2.1. We also assume that our sample $\{X_i\}_{i=1}^T$ is compounded of $\{X_i\}_{i=1}^{\lfloor sT \rfloor}$ and $\{X_i\}_{i=\lfloor sT \rfloor+1}^T$, such that there is only one breakpoint location in $\lfloor sT \rfloor$ with $s \in (0,1)$, i.e. $\{X_i\}_{i=1}^{\lfloor sT \rfloor} \sim C_1(F(X))$ and $\{X_i\}_{i=\lfloor sT \rfloor+1}^T \sim C_2(F(X))$. Then, the covariance bootstrap procedure suggests the following course of action:

i) Estimate the breakpoint location $\lfloor sT \rfloor$ with $\lfloor \hat{s}T \rfloor$, where $\hat{s}$ is determined by $\hat{s} := \underset{s\in(0,1)}{\operatorname{argmax}} \|\hat{\mathbb{U}}_T(s,\mathbf{u})\|_{L^2}$. Sample separately with replacement from $\{X_i\}_{i=1}^{\lfloor \hat{s}T \rfloor}$ and $\{X_i\}_{i=\lfloor \hat{s}T \rfloor+1}^T$ to obtain B bootstrap samples $\{X_i^{(b)}\}_{i=1}^T$, for $b = 1,\dots,B$.

ii) Estimate B versions of the copula difference $\Delta_C^b = \|\hat{C}^{1:\hat{s}T}(\mathbf{u}) - \hat{C}^{\hat{s}T+1:T}(\mathbf{u})\|_{L^2}$, using the estimated break point location $\hat{s}T$ and re-sampled data $\{X_i^{(b)}\}_{i=1}^T$, for $b = 1,\dots,B$.

iii) For $t_1, t_2 \in [0,1]$ compute separately

$$D_i^b(t_i) :=< \sqrt{T}\left(\hat{\mathbb{U}}_T^b(t_i,\mathbf{u}) - \mathbb{U}^b(t_i,\mathbf{u})\right), \mathbb{U}^b(t_i,\mathbf{u}) >_{L^2}$$

for $i = 1,2$ using $\{X_i^{(b)}\}_{i=1}^T$ for $b = 1,\dots,B$, where

$$\mathbb{U}^b(t_i,\mathbf{u}) = (\mathsf{min}\{\hat{\mathsf{s}},\mathsf{t_i}\} - \hat{s}t_i)\,\Delta_C^b.$$

iv) Estimate the expected value given covariance of Theorem B.2.1 for $t_1, t_2 \in (0,1)$ by the mean

$$\hat{d}_{C_1,C_2,\hat{s}}(t_1,t_2) := \frac{1}{B}\sum_{b=1}^{B} D_1^b(t_1) D_2^b(t_2).$$

v) Estimate the variance $\sigma^2_{C_1,C_2,s}$ from Theorem B.2.1 by integrating out over t_1 and t_2, i.e

$$\hat{\sigma}^2_{C_1,C_2,\hat{s}} = 4\int_0^1\int_0^1 \hat{d}_{C_1,C_2,\hat{s}}(t_1,t_2)\,dt_1 dt_2$$

and compute the q-quantile z_q of $N(0,\hat{\sigma}^2_{C_1,C_2,\hat{s}})$ where $q \in (0,1)$.

[1]Since we are only able to derive the limiting distribution in the case of known marginals, there is no theoretical evidence that the covariance bootstrap is applicable for sequentially estimated case.

The testing procedure is as follows: We reject the null of no relevant change $\|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2} \leq \Delta$ if

$$\int_0^1 \hat{L}_T(t)dt > \frac{\hat{s}^2(1-\hat{s})^2}{3}\Delta^2 + \frac{z_q}{\sqrt{T}}. \tag{B.20}$$

The bootstrap and testing procedure can be easily adapted to the quantile case, i.e. $\mathbf{u}$ is fixed, by adapting step i) - iii). Note, the test given in equation (B.20) is an exact level α test if Δ is chosen as the copula difference $||C_1(\mathbf{u}) - C_2(\mathbf{u})||_{L^2}$ or $|C_1(\mathbf{q}) - C_2(\mathbf{q})|^2$. Otherwise the size is smaller than α. By the continuous mapping theorem we obtain that the left hand side of (B.20) converges weakly to a degenerated random variable if the copula difference is equal to zero (no break point). Consequently, the level of the proposed tests have practically size zero, whereas classical stationarity tests hold the asymptotic α-level. Thus, the power of the classical tests is usually larger than the power of the relevant change tests cosndiered here. For practitioners we suggest to run a classical test first, e.g. Bücher (2013) for the case of known marginals and Bücher et al. (2014) in the case of sequentially estimated marginals. If the test rejects the null of stationarity, i.e. the copula difference is significantly larger than zero, estimate the break fraction and apply the proposed relevant change test.

[2]For a detailed description of the quantile version of the test statistic we refer to the main paper.

B.5 Simulations for the Covariance Bootstrap

The data generating process (DGP) is similar to the DGP used in the main paper. We recap the description of the DGP since we want the Supplement Appendix to be autonomous readable. Let

$$X_t = [X_{1t}, X_{2t}]' = N_2(\mathbf{0}, \Sigma_t(\rho)), \tag{B.21}$$

where $N_2(\mathbf{0}, \Sigma_t(\rho))$ with $t = 1, ..., T$ describes the bivariate normal distribution with expectation vector zero and covariance matrix $\Sigma_t(\rho) = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$ and $\rho \in [-1, 1]$. We set ρ equal to -0.3 for $t = 1, ..., \frac{T}{2}$ and $\rho = 0.8$ for $t = \frac{T}{2} + 1, ..., T$. Thus, the breakpoint sT is chosen at $\frac{T}{2}$. We restrict the size analysis in this subsection to the two dimensional case $N = 2$. The following size study presents both L^2-norm based results and an analysis where we consider the specific point $\mathbf{u} = (0.6, 0.6)$. Note, the closer the quantile is to its boundaries, i.e. 0 or 1, the more observations are needed. Critical values of our tests are computed using the bootstrap algorithms from Sections B.4 with $B = 300$ bootstrap replications. The tests are performed at the $\alpha = 0.05, 0.1$ significance level using 301 Monte Carlo replications. The computations were implemented in Matlab, parallelized and performed using CHEOPS, a scientific High Performance Computer at the Regional Computing Center of the University of Cologne (RRZK).

Table B.1 presents the results of the relevant change tests under the null with Δ chosen as the estimated copula difference $|C_1(\mathbf{u}) - C_2(\mathbf{u})|$, where C_1 and C_2 are estimated by the consistent copula estimator

$$\hat{C}(\mathbf{u}) = \frac{1}{t_2 - t_1} \sum_{i=t_1}^{t_2} \mathbb{1}\{F_1(X_{i1}) \leq u_1, \ldots, F_N(X_{iN}) \leq u_N\}, \tag{B.22}$$

using realizations $\{X_1, \ldots X_{\lfloor \hat{s}T \rfloor}\}$ and $\{X_{\lfloor \hat{s}T \rfloor + 1}, \ldots X_T\}$. The breakpoint $\lfloor \hat{s}T \rfloor$ is estimated by

$$\hat{s} := \underset{s \in (0,1)}{\operatorname{argmax}} |\hat{\mathbb{U}}_T(s, \mathbf{u})|. \tag{B.23}$$

Table B.1 reports the results of the relevant change tests under the null, where the functional difference between the copulas is determined by the L^2-norm. Similar to the quantile case we consider for the size analysis $\Delta := \|C_1(\mathbf{u}) - C_2(\mathbf{u})\|_{L^2}$ and accordingly $\hat{s} := \underset{s \in (0,1)}{\operatorname{argmax}} \|\hat{\mathbb{U}}_T(s, \mathbf{u})\|_{L^2}$. Collectively, the tests show good size properties and converges to the predetermined rejection level α if T gets larger.

Overall, the covariance bootstrap shows good size properties for both, the quantile

Table B.1: Size analysis with known marginals

	$T = 300$	$T = 500$	$T = 750$	$T = 1000$
q_{95}	0.099	0.083	0.059	0.046
q_{90}	0.142	0.106	0.109	0.109

Table B.1 reports the rejection rate of the relevant change test for data generated with the DGP described in (B.21) for known marginal distributions and sequential estimated marginals using the two distribution estimation methods with $B = 300$ bootstrap replications. The copula difference is evaluated at $\mathbf{u} = (0.6, 0.6)$. In total, we conducted 301 Monte Carlo replications.

Table B.2: Size analysis with known marginals using the L^2-norm

	$T = 1000$	$T = 2000$	$T = 3000$	$T = 4000$
q_{95}	0.085	0.063	0.046	0.066
q_{90}	0.156	0.122	0.113	0.102

Table B.2 reports the rejection rate of the relevant change test for data generated with the DGP described in (B.21) for known marginal distributions and sequential estimated marginals using the two distribution estimation methods with $B = 300$ bootstrap replications. The copula difference is determined using the L^2-norm. In total, we conducted 301 Monte Carlo replications.

version of the test and the test given in (B.20) by a moderate rate of bootstrap replications.

APPENDIX C

Specification Testing in Functional Quantile Regression Models

C.1 Proofs

In order to maintain readability we omit the index $Y|X$ for the conditional cdf F. To prove Theorem 4.4.1, we first derive and prove three auxiliary results. Therefore, we define the following three processes for $(y,x) \in \mathbb{R}^{K+1}$ and $(\theta,\tau) \in \Theta \times \mathcal{T}$:

$$\nu_n(y,x) := \sqrt{n}\left(\hat{F}_n(y,x) - F(y,x)\right) \tag{C.1}$$

$$\gamma_n(\theta,\tau) := \sqrt{n}\left(\hat{G}_n(\hat{\theta}_n,\tau) - G(\theta,\tau)\right) \tag{C.2}$$

$$\nu_n^0(y,x) := \sqrt{n}\left(\hat{F}_n(y,x,\hat{\theta}_n) - F(y,x,\theta_0)\right). \tag{C.3}$$

Lemma C.1.1. *Let Assumptions* (2) *be true. For the processes* (C.1) *and* (C.2) *it holds under the null, that*

$$(\nu_n,\gamma_n) \Rightarrow \tilde{\mathbb{G}} := (\mathbb{G}_1,\tilde{\mathbb{G}}_2) \text{ in } \ell^\infty(\mathcal{S}\times\Theta\times\mathcal{T}), \tag{C.4}$$

where $\tilde{\mathbb{G}}$ is a tight bivariate mean zero Gaussian process.

Proof. First, we notice that the Donsker property is conserved under the union of Donsker classes. Hence, ν_n and $\gamma_n(\theta,\tau)$ are F_{YX}- Donsker for all $\theta \in \mathcal{B}(\mathcal{T},\Theta)$ and $\tau \in \mathcal{T}$ with limiting process $\mathbb{G}_1(y,x)$ and $\tilde{\mathbb{G}}_2$, respectively. Since arbitrary linear combinations of ν_n and γ_n are Lipschitz and thus Donsker (see Vaart, 1998, Example 29.20), we conclude by the Cramér-Wold theorem that (ν_n,γ_n) converge in distribution to $\tilde{\mathbb{G}}$. □

Before we prove the next lemma we slightly generalize Lemma E.3 from Chernozhukov et al. (2013) for our purposes. This modification summarized in the following corollary states conditions under which a Z-estimation process satisfies the functional delta method for Gaussian processes.

Corollary C.1.2. *Let Assumption 2 i.) − iv.) be satisfied and* $\sqrt{n}\left(\hat{G}_n - G\right) \Rightarrow \tilde{\mathbb{G}}_2$ *in* $\ell^\infty(\Theta \times I_l)$ *for all* $l = 1, ..., L$*, where* $\tilde{\mathbb{G}}_2$ *is a Gaussian process with a.s. uniformly continuous paths on* $\Theta \times I_l$*,* $l = 1, ..., L$*. Further, we assume that the estimator* $\hat{\theta}_n(\tau)$ *is an approximate* Z*-estimator ((4.12)) for all* $\tau \in I_l$ *with* $l = 1, ..., L$*. Then*

$$\sqrt{n}\left(\hat{\theta}_n(\cdot) - \theta_0(\cdot)\right) = -\dot{G}^{-1}_{\theta_0(\cdot),\cdot}\left[\sqrt{n}(\hat{G}_n - G)(\theta_0(\cdot), \cdot)\right] + o_P(1) \tag{C.5}$$

$$\Rightarrow -\dot{G}^{-1}_{\theta_0(\cdot),\cdot}\left[\tilde{\mathbb{G}}_2(\theta_0(\cdot), \cdot)\right] \in \ell^\infty(\mathcal{T}). \tag{C.6}$$

If Assumption 2 v.) also holds, then the paths $\tau \mapsto -\dot{G}^{-1}_{\theta_0(\tau),\tau}\left[\tilde{\mathbb{G}}_2(\theta_0(\tau), \tau)\right]$ *are a.s. uniformly continuous on* $\mathcal{T}$*.*

Proof. The intersection of I_{l_1} and I_{l_2} is a singleton by assumption for $l_1 \neq l_2$. Thus, the set of possible discontinuities is a null set with respect to the Lebesgue measure. Hence, the limiting process $\tilde{\mathbb{G}}_2$ is a.s. continuous on $\Theta \times \mathcal{T}$ with respect to the Euclidean metric. Further we notice, that by assumption the decomposition of the unit interval is finite. Consequently, the property of uniformity is also applicable to the finite union of compact sets. Hence, the conditions of Lemma E.3 in Chernozhukov et al. (2013) are fulfilled. □

Lemma C.1.3. *Let either the null hypothesis or a fixed alternative and Assumptions 2 be true. Then it holds that*

$$(\nu_n, \nu_n^0) \Rightarrow \mathbb{G} := (\mathbb{G}_1, \mathbb{G}_2) \text{ in } \ell^\infty(\mathcal{S} \times \mathcal{S}), \tag{C.7}$$

where $\mathbb{G}_1$ *is the limiting tight bivariate mean zero Gaussian process of* ν_n *and*

$$G_2 := \int F(y|x^*)\mathbb{1}_{\{x^* \leq x\}} d\mathbb{G}_1(\infty, x^*) + \int \mathbb{G}_2^*(y, x^*)\mathbb{1}_{\{x^* \leq x\}} dF_X(x^*). \tag{C.8}$$

Proof. Under either the null hypothesis or a fixed alternative, it follows by standard arguments from Lemma C.1.1 and Corollary C.1.2 that in $\ell^\infty(\mathcal{S}) \times \ell^\infty(\mathcal{T})$

$$\sqrt{n}\left(\hat{F}_n(\cdot, \cdot) - F(\cdot, \cdot), \hat{\theta}_n(\cdot) - \theta_0(\cdot)\right) \Rightarrow \left(\mathbb{G}_1(\cdot, \cdot), -\dot{G}^{-1}_{\theta_0(\cdot),\cdot}(\tilde{\mathbb{G}}_2(\theta_0(\cdot), \cdot)\right). \tag{C.9}$$

Next, it follows from the Hadamard differentiability (cf. Assumption 2 *vii.*)) that

$$\begin{aligned}\sqrt{n}\left(\hat{F}_n(y|x, \hat{\theta}_n) - F(y|x, \theta_0)\right) \Rightarrow &- \dot{F}(y|x, \theta_0)\left[\dot{G}^{-1}_{\theta_0(\cdot),\cdot}(\tilde{\mathbb{G}}_2(\theta_0(\cdot), \cdot)\right] \\ &=: \mathbb{G}_2^+(y, x).\end{aligned} \tag{C.10}$$

The statement of the lemma then follows directly from the Hadamard derivative $\dot{\phi}$ of the mapping

$$\phi((A,B))[x^*] := \int A(\cdot,x^*)\mathbb{1}_{\{x^*\leq\cdot\}}dB(x^*) \tag{C.11}$$

given by

$$\dot{\phi}_{A,B}(\alpha,\beta)[x^*] = \int A(\cdot,x^*)\mathbb{1}_{\{x^*\leq\cdot\}}d\beta(x^*) + \int \alpha(\cdot,x^*)\mathbb{1}_{\{x^*\leq\cdot\}}dB(\cdot,x^*) \tag{C.12}$$

and the functional delta method. In particular, for the second component $\mathbb{G}_2$ of the joint limiting process, we have

$$\mathbb{G}_2(y,x) = \int \mathbb{G}_2^+(y,x^*)\mathbb{1}_{\{x^*\leq x\}}dF_X(x^*) + \int F(y|x^*)\mathbb{1}_{\{x^*\leq x\}}d\mathbb{G}_1(\infty,x^*). \tag{C.13}$$

□

of Theorem 4.4.1. We start with the first statement of Theorem 4.4.1. Under the null hypothesis it holds that $\hat{F}_n(y,x) = F(y,x,\theta_0)$ for all $(y,x)\in\mathcal{S}$. By linearity, we have

$$S_n^{CM} = \sqrt{n}\int\left(\hat{F}_n(y,x) - \hat{F}_n(y,x,\hat{\theta})\right)d\hat{F}_n(y,x) \tag{C.14}$$

$$\begin{aligned}&= \int\left(\nu_n(y,x) - \nu_n^0(y,x)\right)^2 dF(y,x)\\ &\quad + \int\left(\nu_n(y,x) - \nu_n^0(y,x)\right)^2 d\left(\hat{F}_n(y,x) - F(y,x)\right).\end{aligned} \tag{C.15}$$

From Lemma C.1.3 we know that $(\nu,\nu_0)\Rightarrow(\mathbb{G}_1,\mathbb{G}_2) = \mathbb{G}$, where $\mathbb{G}$ is a tight bivariate mean zero Gaussian process. Applying the continuous mapping theorem and the Donsker class property yield

$$S_n^{CM} = \int(\mathbb{G}_1(y,x) - \mathbb{G}_2(y,x)^2 dF(y,x) + o_p(1) \tag{C.16}$$

which claims the statement.
To show part *ii.*), we use the fact that under any fixed alternative $P(F(y,x)\neq F(y,x,\theta_0) > 0$ due to construction of the alternative hypothesis in (4.9). Thus,

$$\begin{aligned}S_n^{CM} = &\int\left(\nu_n(y,x) - \nu_n^0(y,x) + \sqrt{n}(F(y,x) - F(y,x,\theta_0)\right)^2 dF(y,x)\\ &+ o_P(1) = \mathcal{O}_P(n),\end{aligned} \tag{C.17}$$

which implies that S_n^{CM} is greater than any fixed constant $\varepsilon > 0$ and hence, the probability that S_n^{CM} is greater than any $\varepsilon > 0$ tends to 1. □

In order to prove Theorem 4.4.2 we present the bootstrap version of Lemma C.1.1 as an auxiliary result.

Lemma C.1.4. *Let Assumption 2 be true. We define the bootstrap version of the empirical processes* (C.1) *and* (C.3)

$$\nu_{n,B}(y,x) := \sqrt{n}\left(\hat{F}_{n,B}(y,x) - \hat{F}_n(y,x,\hat{\theta}_n)\right) \tag{C.18}$$

$$\nu^0_{n,B}(y,x) := \sqrt{n}\left(\hat{F}_{n,B}(y,x,\hat{\theta}_n) - \hat{F}_n(y,x,\hat{\theta}_n)\right). \tag{C.19}$$

Then it holds under either the null or a fixed alternative hypothesis that

$$\left(\nu_{n,B}, \nu^0_{n,B}\right) \Rightarrow \mathbb{G}_b, \tag{C.20}$$

where $\mathbb{G}_b := (\mathbb{G}_{b1}, \mathbb{G}_{b2})$ *is a tight bivariate mean zero Gaussian process whose distribution function coincides with that of the process* $\mathbb{G}$ *in Lemma C.1.1.*

Proof. This follows from Lemma C.1.1 and the functional delta method for the bootstrap (Rothe & Wied, 2013). □

Finally, we can prove the statements of Theorem 4.4.3.

Theorem 4.4.3. To prove part *i.*), let $c(\alpha)$ be the true critical value satisfying $P(S_n^{CM} > c(\alpha)) = \alpha + o_P(1)$. Then it follows from Lemma C.1.4 that $\hat{c}_n(\alpha) = c(\alpha) + o_P(1)$. This implies that S_n^{CM} and $\tilde{S}_n := S_n^{CM} - (\hat{c}_n(\alpha) - c(\alpha))$ converge to the same limiting distribution as n tends to infinity. Hence, $P(S_n^{CM} > \hat{c}_n(\alpha)) = \alpha + o_P(1)$ as claimed. To prove part *ii.*), we deduce from Lemma C.1.4 that the bootstrap critical values are bounded in probability under fixed alternatives. Thus, for any $\varepsilon > 0$, there is an $N(\epsilon)$ such that $P(\hat{c}_n(\alpha) > N(\varepsilon)) < \varepsilon + o_P(1)$. By Kolmogorv axioms we obtain

$$\begin{aligned}
&P(S_n^{CM} \leq \hat{c}_n(\alpha)) && \text{(C.21)}\\
&\quad = P(S_n^{CM} \leq \hat{c}_n(\alpha), S_n^{CM} \leq N(\epsilon)) + P(S_n^{CM} \leq \hat{c}_n(\alpha), S_n^{CM} > N(\epsilon)) && \text{(C.22)}\\
&\quad \leq P(S_n^{CM} \leq N(\epsilon)) + P(S_n^{CM} > N(\epsilon)) && \text{(C.23)}\\
&\quad \leq \varepsilon + o_P(1), && \text{(C.24)}
\end{aligned}$$

where the last inequality can be deduced from Theorem 4.4.1 *ii.*).
Statements *iii.*) and *iv.*) follow in addition with assumption 3 immediately from *i.*) and *ii.*). □

Bibliography

Aloui, R., Aïssa, M. S. B., & Nguyen, D. K. (2011). Global financial crisis, extreme interdependences, and contagion effects: The role of economic structure? *Journal of Banking & Finance*, *35*(1), 130–141.

Angrist, J., Chernozhukov, V., & Fernández-Val, I. (2006). Quantile regression under misspecification, with an application to the us wage structure. *Econometrica*, *74*(2), 539–563.

Arnold, M., Stahlberg, S., & Wied, D. (2013). Modelling different kinds of spatial dependence in stock returns. *Empirical Economics*, *44(2)*, 761–774.

Asgharian, H., Hess, W., & Liu, L. (2013). A spatial analysis of international stock market linkages. *Journal of Banking and Finance*, *37(12)*, 4738–4754.

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

Bach, S., Corneo, G., & Steiner, V. (2009). From bottom to top: The entire income distribution in germany, 1992 - 2003. *Review of Income and Wealth*, *55*(2), 303–330.

Belloni, A., Chernozhukov, V., Chetverikov, D., & Fernández-Val, I. (2019). Conditional quantile processes based on series or many regressors. *Journal of Econometrics*, *213*(1), 4–29.

Berkson, J. (1938). Some difficulties of interpretation encountered in the application of the chi-square test. *J. Am. Statist. Ass.*, *33*, 526–536.

Bierens, H. J. (1990). A consistent conditional moment test of functional form. *Econometrica*, 1443–1458.

Biewen, M. (2000). Income inequality in germany during the 1980s and 1990s. *Review of Income and Wealth*, *46*(1), 1–19.

Billingsley, P. (1995). *Probability and Measure*. Wiley Series in Probability and Statistics. Wiley.

Blasques, F., Koopman, S., Lucas, A., & Schaumburg, J. (2016). Spillover dynamics for systemic risk measurement using spatial financial time series models. *Journal of Econometrics*, *195(2)*, 211–223.

Bondell, H. D., Reich, B. J., & Wang, H. (2010). Noncrossing quantile regression curve estimation. *Biometrika*, *97*(4), 825–838.

Born, B. & Breitung, J. (2011). Simple regression based tests for spatial dependence. *Econometrics Journal*, *14(2)*, 330–342.

Breitung, J. & Wigger, C. (2017). Alternative gmm estimators for spatial regression models. *Spatial Economic Analysis*, *13:2*, 148–170.

Breunig, C. (2019). Specification testing in nonparametric instrumental quantile regression. *arXiv preprint arXiv:1909.10129*.

Brodsky, B., Penikas, H., Safaryan, I., et al. (2009). Detection of structural breaks in copula models. *Applied Econometrics*, *16*(4), 3–15.

Bücher, A., Kojadinovic, I., Rohmer, T., & Segers, J. (2014). Detecting changes in cross-sectional dependence in multivariate time series. *Journal of Multivariate Analysis*, *132*, 111–128.

Bücher, R. (2013). Consistent testing for a constant copula under strong mixing based on the tapered block multiplier technique. *Journal of Multivariate Analysis*, *116*, 208–229.

Busetti F., H. A. (2011). When is a copula constant? a test for changing relationships. *Journal of Financial Econometrics*, *9*, 106–131.

Card, D., Heining, J., & Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, *128*(3), 967–1015.

Cardot, H., Crambes, C., & Sarda, P. (2005). Quantile regression when the covariates are functions. *Nonparametric Statistics*, *17*(7), 841–856.

Catania, L. & Billé, A. (2017). Dynamic spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Applied Econometrics*, *32(6)*, 1178–1196.

Chernozhukov, V. (2002). Inference on quantile regression process, an alternative. *SSRN Electronic Journal*.

Chernozhukov, V., Fernández-Val, I., & Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, *81*(6), 2205–2268.

Chow, G. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, *28*, 591–605.

Christ, C. F. (1994). The Cowles Commission's Contributions to Econometrics at Chicago, 1939-1955. *Journal of Economic Literature*, *32*(1), 30–59.

Corradi, V. & Swanson, N. R. (2003). Bootstrap conditional distribution tests in the presence of dynamic misspecification. Working Paper 2003-11, New Brunswick, NJ.

Creedy, J. & Hart, P. E. (1979). Age and the distribution of earnings. *The Economic Journal*, *89*(354), 280–293.

De Finetti, B. (2017). *Theory of probability: A critical introductory treatment*, volume 6. John Wiley & Sons.

Dehling, H., Vogel, D., Wendler, M., & Wied, D. (2017). Testing for changes in kendall's tau. *Econometric Theory*, *33*(6), 1352–1386.

Demetrescu, M. & Wied, D. (2019). Testing for constant correlation of filtered series under structural change. *The Econometrics Journal*, *22*(1), 10–33.

Dette, H., Gösmann, J., et al. (2018). Relevant change points in high dimensional time series. *Electronic Journal of Statistics*, *12*(2), 2578–2636.

Dette, H. & Wied, D. (2016). Detecting changes in time series models. *Journal of the Royal Statistical Society, Series B*, *78*, 371–394.

Dette, H., Wu, W., & Zhou, Z. (2018). Change point analysis of second order characteristics in non-stationary time series. *Statistica Sinica, forthcoming.*

Dudley, R. M. (2014). *Uniform Central Limit Theorems* (2 ed.). Cambridge Studies in Advanced Mathematics. Cambridge University Press.

Dustmann, C., Ludsteck, J., & Schönberg, U. (2009). Revisiting the German Wage Structure. *The Quarterly Journal of Economics*, *124*(2), 843–881.

Efron, B. (2005). Bayesians, frequentists, and scientists. *Journal of the American Statistical Association*, *100*, 1–5.

Elhorst, J. P., Lacombe, D. J., & Piras, G. (2012). On model specification and parameter space definitions in higher order spatial econometric models. *Regional Science and Urban Economics*, *42*, 211–220.

Escanciano, J., Goh, C., et al. (2010). Specification analysis of structural quantile regression models. *Technical Report.*

Escanciano, J. C. & Velasco, C. (2010). Specification tests of parametric dynamic conditional quantiles. *Journal of Econometrics*, *159*(1), 209–221.

Fleming, M. (2004). Techniques for estimating spatially dependent discrete choice models. In L. Anselin, R. Florax, & S. Rey (Eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Springer, Berlin, Heidelberg.

Frisch, R. (1933). Editor's note. *Econometrica*, *1*(1), 1–4.

Gong, P. & Weng, Y. (2016). Value-at-Risk Forecasts by a Spatiotemporal Model in Chinese Stock Market. *Physica A: Statistical Mechanics and its Applications*, *441(C)*, 173–191.

Gozalo, P. L. (1993). A consistent model specification test for nonparametric estimation of regression function models. *Econometric Theory*, *9*(3), 451–477.

Granger, C. W., King, M. L., & White, H. (1995). Comments on testing economic theories and the use of model selection criteria. *Journal of Econometrics*, *67*(1), 173 – 187.

Guerre, E. & Sabbah, C. (2012). Uniform bias study and bahadur representation for local polynomial estimators of the conditional quantile function. *Econometric Theory*, *28*(1), 87–129.

Hallin, M., Lu, Z., Yu, K., et al. (2009). Local linear spatial quantile regression. *Bernoulli*, *15*(3), 659–686.

Hansen, E. B. (1994). Autoregressive density estimation. *International Economic Review*, *35*, 705–730.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, *50*, 1029–1054.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 1251–1271.

He, X. & Shi, P. (1997). Monotone b-spline smoothing. *Journal of the American Statistical Association*, *93*.

Härdle, W. & Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *The Annals of Statistics*, *21*(4), 1926–1947.

Kapoor, M., Kelejian, H. H., & Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of Econometrics, 140(1):97-130*.

Kelejian, H. & Piras, G. (2016). An extension of the j - test to a spatial panel data frameworkwork. *Journal of Applied Econometrics*, *31*(2), 387–402.

Kelejian, H. & Prucha, I. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, *40*, 509–533.

Kelejian, H. & Prucha, I. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, *157*, 53–67.

Keuzenkamp, H. A. (2000). *Probability, Econometrics and Truth: The Methodology of Econometrics*. Cambridge University Press.

Keuzenkamp, H. A. & Magnus, J. R. (1995). On tests and significance in econometrics. *Journal of Econometrics*, *67*(1), 5 – 24.

Keynes, J. (1921). *A Treatise on Probability*. Macmillan and Co.

Klein, N., Kneib, T., Lang, S., Sohn, A., et al. (2015). Bayesian structured additive distributional regression with an application to regional income inequality in germany. *The Annals of Applied Statistics*, *9*(2), 1024–1052.

Kluge, J. & Weber, M. (2018). Decomposing the german east–west wage gap. *Economics of Transition*, *26*(1), 91–125.

Koenker, R. (2005). *Quantile Regression*. Econometric Society Monographs. Cambridge University Press.

Koenker, R. & Bassett Jr, G. (1978). Regression quantiles. *Econometrica*, 33–50.

Koenker, R., Ng, P., & Portnoy, S. (1994). Quantile smoothing splines. *Biometrika*, *81*(4), 673–680.

Koenker, R. & Xiao, Z. (2002). Inference on the quantile regression process. *Econometrica*, *70*(4), 1583–1612.

Krämer, W. & van Kampen, M. (2011). A simple nonparametric test for structural change in joint tail probabilities. *Economics Letters*, *110*, 245–247.

Kutzker, T. & Davidyan, A. (2020). Unsupervised learning with non-uniform random forests. *Working Paper*.

Kutzker, T., Klein, N., & Wied, D. (2020). Specification testing in functional quantile regression models with an application to income differences in germany. *Working Paper*.

Kutzker, T. & Schreiter, M. (2020). The optimal capital structure under risks of illiquidity and over-indebtedness in a double barrier option framework. *Working Paper*.

Kutzker, T., Stark, F., & Wied, D. (2019). Testing for relevant dependence change in financial data: a cusum copula approach. *Empirical Economics*, 1–20.

Kutzker, T. & Wied, D. (2019). Testing the correct specification of a spatial dependence panel model for stock returns. *Working Paper*.

Lakatos, I. (1978). *The Methodology of Scientific Research Programmes: Philosophical Papers*, volume 1. Cambridge University Press.

Landmesser, J. M. (2016). Decomposition of differences in income distributions using quantile regression. *Statistics in Transition. New Series*, *17*(2), 331–349.

Li, D., Li, Q., & Li, Z. (2020). Nonparametric quantile regression estimation with mixed discrete and continuous data. *Journal of Business & Economic Statistics*, 1–16.

Li, H., Calder, C., & Cressie, N. (2007). Beyond Moran's I: Testing for Spatial Dependence Based on the Spatial Autoregressive Model. *Geographical Analysis*, *39(4)*, 357–375.

Li, Q. & Racine, J. S. (2008). Nonparametric estimation of conditional cdf and quantile functions with mixed categorical and continuous data. *Journal of Business & Economic Statistics*, *26*(4), 423–434.

Liebscher, E. (1996). Central limit theorems for the sums of α-mixing random variables. *Stochastics and Stochastic Reports*, *59*, 241–258.

Lu, L. (2017). Simultaneous spatial panel data models with common shocks. *FRB Boston Risk and Policy Analysis Unit Paper No. RPA*, 17–3.

Luong, M. & Hébert, B.-P. (2009). *Age and earnings*. Citeseer.

Machado, J. A. F. & Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, *20*(4), 445–465.

MacKinnon, J. G. (2002). Bootstrap inference in econometrics. *The Canadian Journal of Economicss*, *35*, 615–645.

Manner, H., Stark, F., & Wied, D. (2019). Testing for structural breaks in factor copula models. *Journal of econometrics*, *208*(2), 324–345.

Melly, B. (2005). Public-private sector wage differentials in Germany: Evidence from quantile regression. *Empirical Economics*, *30*, 505–520.

Moran, P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37(1)*, 17–23.

Oh, D. H. & Patton, A. J. (2017). Modelling dependence in high dimensions with factor copulas. *Journal of Business and Economic Statistics*, *35*, 139–154.

Page, E. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika*, *42*(3/4), 523–527.

Qu, Z. & Yoon, J. (2015). Nonparametric estimation and inference on conditional quantile processes. *Journal of Econometrics*, *185*(1), 1–19.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, *31*(2), 350–371.

Rothe, C. & Wied, D. (2013). Misspecification testing in a class of conditional distributional models. *Journal of the American Statistical Association*, *108*(501), 314–324.

Schmitt, T., Schäfer, R., Wied, D., & Guhr, T. (2016). Spatial Dependence in Stock Returns - Local Normalization and VaR Forecasts. *Empirical Economics*, *50(3)*, 1091–1109.

Selan, B. & Kalatzis, A. (2017). Peer effects of stock returns and financial characteristics: Spatial approach for an emerging market. *working paper*.

Spokoiny, V. & Zhilova, M. (2014). Bootstrap confidence sets under model misspecification.

Stute, W. (1997). Nonparametric model checks for regression. *The Annals of Statistics*, 613–641.

Stute, W., Manteiga, W. G., & Quindimil, M. P. (1998). Bootstrap approximations in model checks for regression. *Journal of the American Statistical Association*, *93*(441), 141–149.

Su, L. & Qu, X. (2017). Specification tests for spatial autoregressive models. *Journal of Business and Economic Statistics*, *35(4)*, 572–584.

Tam, P. (2014). A spatial-temporal analysis of east asian equity market linkages. *Journal of Comparative Economics*, *42(2)*, 304–327.

Vaart, A. W. v. d. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

van der Vaart, A. W. & Wellner, J. (1996). Weak convergence and empirical processes. *Springer Series in Statistics*.

Wagner, G., Frick, J., & Schupp, J. (2007). The german socio-economic panel study (soep) - scope, evolution and enhancements. *Schmollers Jahrbuch : Journal of Applied Social Science Studies / Zeitschrift für Wirtschafts- und Sozialwissenschaften*, *127*(1), 139 – 169.

White, H. (1980). Using least squares to approximate unknown regression functions. *International Economic Review*, *21*(1), 149–170.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, *50*(1), 1–25.

Wied, D. (2013). Cusum-type testing for changing parameters in an autoregressive model for stock returns. *Journal of Time Series Analysis*, *34*, 221–229.

Wied, D., Dehling, H., van Kampen, M., & Vogel, D. (2013). A fluctuation test for constant spearmans's rho with nuisance-free limit distribution. *Computational Statistics and Data Analysis*, *76*, 723–736.

Wu, D.-M. (1973). Alternative tests of independence between stochastic regressors and disturbances. *Econometrica*, 733–750.

Xue, L. & Wang, J. (2010). Distribution function estimation by constrained polynomial spline regression. *Journal of Nonparametric Statistics*, *22*(4), 443–457.

Yu, D., Kong, L., & Mizera, I. (2017). An alternative approach to functional linear partial quantile regression. *arXiv preprint arXiv:1709.02069*.

Zhang, W.-G., Mo, G.-L., Liu, F., & Liu, Y.-J. (2018). Value-at-risk forecasts by dynamic spatial panel gjr-garch model for international stock indices portfolio. *Soft Computing*, *22*(16), 5279–5297.

Ziggel, D., Berens, T., Weiß, G., & Wied, D. (2014). A new set of improved value-at-risk backtests. *Journal of Banking and Finance*, *48*, 29–41.