Essays in nonlinear dynamics in economics and econometrics with applications to monetary policy and banking

Wolski, M.

Chapter 4

Exploring Nonlinearities in Financial Systemic Risk

4.1 Introduction

The 2007-2009 crisis shed new light on the complexity within the financial sector. The linkages and risk exposures between various institutions proved to be of great significance in transmitting distress across the whole financial system. Additionally, during systemic events the malaise spreads across the financial world rapidly through indirect channels, like price effects or liquidity spirals (Brunnermeier, 2009). In effect, market values of various financial assets tend to move closer together, drifting away from their fundamentals. In particular, one observes high regularities in their tail co-movements (Adrian and Brunnermeier, 2011).

Because of its strong adverse effects on the real economy, great attention has been paid to measuring and monitoring systemic risk, i.e. risk of disruption in the entire financial system, and individual risk exposures. The majority of econometric approaches in these fields focus on co-risk measures, where the risk of the financial system is assessed in relation to the risk of individual institutions. The intuition behind these models lies in negative externalities which one institution imposes on the others and on the system as a whole. As argued by Adrian

and Brunnermeier (2011), these externalities are a consequence of excessive risk taking and leverage. Given, for instance, that one institution is facing a liquidity shock, it liquidates its assets at fire-sale prices as given, affecting borrowing constraints of others and actually causing the fire-sale prices. A wonderful summary of research in this field can be found in Acharya (2009), Acharya et al. (2010) or Adrian and Brunnermeier (2011).

A commonly used econometric approach, in the growing body of literature on this topic, is Conditional Value-at-Risk (CoVaR), attributed to Adrian and Brunnermeier (2011). It is built around the concept of Value-at-Risk (VaR), which determines the maximum loss on returns within the γ-percentile confidence interval (Kupiec, 2002). CoVaR assesses VaR_γ of one institution conditional on distress in the other. In particular, if the former represents the system, one may associate CoVaR with a systemic risk measure.

A clear shortcoming of such an approach lies in its susceptibility to model misspecification. Imagine that returns come from an unknown probability distribution F, with density f. Assume now that f is steeper or nonlinear around its VaR_γ. Clearly, standard parametric approaches oversee this irregularity so that even a small variation in VaR_γ might affect co-risk results. In this chapter we develop a methodology which corrects for this shortcoming, contributing to the discussion on nonlinear economic dynamics in systemic risk.

The existence of nonlinearities in the field has been already recognized. Huang et al. (2010) suggest that "a bank's contribution to the systemic risk is roughly linear in its default probability and highly nonlinear with respect to institution size and asset correlation". This is supported by empirical observations of the financial markets described by He and Krishnamurthy (2012). In fact, He and Krishnamurthy (2012) built a theoretical model which matches nonlinear dynamics across different economic variables, including systemic risk. XiaoHua and Shiying (2012) investigated the topic from the neural network perspective and designed an early warning mechanism accordingly. This chapter aims to propose a formal approach to assess the relevance of nonlinearities in driving systemic events.

We build our approach around the intuition of CoVaR. In particular, we focus on the Granger

causal effect that distress in one institution may lead to distress in the other or in the whole system, where distress is defined by being near VaR_γ.

There are two main novelties in our methodology. The first one is the notion of causality. The basic CoVaR notion does not distinguish between direct causal and common factor effects. Adrian and Brunnermeier (2011) treat this as a virtue rather than a problem, arguing that common factor effects are of more importance when dealing with systemic risk, which can be expected to be particularly true for the herding behavior (Brunnermeier et al., 2009). One may, however, want to study the causal relations explicitly. Imagine for instance a group of the biggest financial institutions. Since they do not only trade with each other but also serve as clearing houses or liquidity backstops for smaller parties, they are central to the financial system. Now, imagine that one of them is in trouble. It affects all the banks that are exposed to its risk, but since it is relatively large, its distress might alone translate into problems in the entire financial system. The causal kind of reasoning seems therefore particularly appealing for policy makers and central bankers, who in fact might want to focus on preventing this individual causal relation.

Another justification for considering causality in individual and systemic risk lies in its possible applications to networks and contagion analysis (see for instance Chinazzi and Fagiolo (2013)). Looking at any pair of institutions, the possible risk effects of one on another do not have to be bilaterally equal (as they are assumed to be in a non-causal setting). For instance, a lender has a different kind of risk exposure to a creditor than the other way around. Causality captures that phenomenon explicitly, allowing for a more detailed analysis on network spillovers, cascades and shock propagation.

In our study we employ the general causality of Granger type, i.e. a nonparametric version of the concept originally proposed by Granger (1969), as it is intuitive and does not bring many model restrictions. It has been also successfully applied as a network mapping tool in financial analysis (Gao and Ren, 2013).

The second novelty lies in the definition of financial distress. In our study we assume that an

institution is in trouble when it is around its VaR_γ. Practically speaking, our definition captures the majority of events which fall below VaR_γ together with some of the events above it. The reason why we allow for some variation around VaR_γ lies in its possible nonlinear structure, whose role we want to study explicitly. We recognize that our definition might not capture some of the extreme values from the left tail of the distribution, being potentially susceptible to *black swans* (Taleb, 2010). Our analysis shows, however, that the optimal region around VaR_γ is very slowly decreasing with the sample size, somehow hampering the risk of neglecting the extreme events. Additionally, our setup might be naturally extended to a more general setting, including all the events below VaR_γ. This, however, is behind the scope of this chapter and we leave it for further investigation.

In our analysis we consider two scenarios of potential Granger causality. In the first setting we investigate the role of individual institutions in blocking the recovery of the system which is already under distress. In the second scenario we measure the contribution of individual institutions to the systemic problems. The second setting is more similar to the standard understanding of systemic risk (Acharya, 2009) and might be useful in *ex ante* applications. The first scenario might be perceived either as a kind of a robustness check or a policy relevant tool for *ex post* actions. Indeed, if the system is already in trouble one may want to determine which of its parts are hampering its recovery. In fact, we could think of these two scenarios from a perspective of a doctor who either prescribes precautionary drugs or is trying to heal an already sick patient.

This chapter is organized as follows. In Section 4.2 we explain the methodology of Conditional Value-at-Risk-Nonlinear Granger Causality (or NCoVaR for simplicity). We evaluate the asymptotic properties of the test statistic and we confirm them numerically in Section 4.3. In Section 4.4 we apply our approach to the euro zone financial sector and evaluate which institutions got the most significant impact on the systemic risk in years 2000-2012. Section 4.5 concludes.

4.2 Methodology of NCoVaR

Let us first bring some intuition behind the Conditional Value-at-Risk and Granger causality separately and then use this to build CoVaR-NGraCo (Conditional Value-at-Risk-Nonlinear Granger Causality) or NCoVaR for simplicity. In the standard setting we consider two institutions, i and j, whose returns on assets are given by X^i and X^j, respectively. Talking about systemic risk, we set j to be some aggregate variable so that we investigate the relationship between institution i and the system as a whole. Following the original CoVaR literature, let us define VaR_γ as the left γ-quantile of the unconditional returns of a given institution. (In practice γ is chosen from $\{0.01, 0.05, 0.1\}$.) For institution i we have therefore

$$P(X^i \leq \mathrm{VaR}^i_\gamma) = \gamma, \tag{4.1}$$

or equivalently

$$\mathrm{VaR}^i_\gamma = \inf\{x^i : F_{X^i}(x^i) \geq \gamma\}, \tag{4.2}$$

where F_{X^i} is the cumulative distribution function of X^i. (For institution j, the notation is analogous throughout the chapter.) The intuition behind CoVaR is to evaluate VaR_γ of institution j conditional on some event associated with institution i. In particular, Adrian and Brunnermeier (2011) consider two conditioning events, i.e. institution i is at its VaR^i_γ or at its median ($\mathrm{VaR}^i_{\gamma=0.5} = Median^i$). By comparing the difference between the two, it is possible to estimate the risk contribution of institution i onto j, denoted by ΔCoVaR.

In our study we follow a similar reasoning as Adrian and Brunnermeier (2011), however, we add a (discrete) time dimension. For any period t, let us define the future returns' information set by $\mathcal{G}X^i_t$, and the past and/or current returns' information set by $\mathcal{F}X^i_t$. Following Granger (1969), we say that returns of institution i are Granger causing those of institution j if $\mathcal{F}X^i_t$ contains additional information on $\mathcal{G}X^j_t$ which is not already contained in $\mathcal{F}X^j_t$ alone. We formulate the definition of conditional Granger causality analogously, i.e. we say that returns of institution i are Granger causing those of institution j if, conditional on some past or current

events of those institutions (denoted by $\mathcal{A}(\mathcal{F}X_t^i)$ and $\mathcal{B}(\mathcal{F}X_t^j)$, respectively), $\mathcal{F}X_t^i$ contains additional information on $\mathcal{G}X_t^j$ which is not already contained in $\mathcal{F}X_t^j$ alone.

Given the intuition behind the CoVaR and general Granger causality, we may now turn to NCoVaR. Similarly to ΔCoVaR, we test the difference in Granger causal risk effects from institution i on j, between two conditioning events, i.e. when institution i is and/or was in trouble (or around its VaR_γ^i) and when it is and/or was around the median of its returns. An advantage of allowing institutions to be around (and not exactly at) their VaR_γ or median levels is that we could thereof account for possible nonlinearities in corresponding distributions - something the original methodology could not capture. In particular, we consider a μ-radius ball ($\mu > 0$) centered at VaR_γ or the median. (The following reasoning holds for $\mathcal{G}$ and $\mathcal{F}$ being multivariate, provided that VaR_γ and the medians are taken over the marginals.) We also allow for conditioning on the past and/or current realizations of X_t^j. To formalize this we give the following definition of NCoVaR.

Definition 4.2.1. *Given any stationary bivariate process $\{(X_t^i, X_t^j)\}$, we say that $\{X_t^i\}$ is a nonlinear CoVaR Granger cause of $\{X_t^j\}$ if*

$$P\left(\|\mathcal{G}X_t^j - \mathrm{VaR}_\gamma^j\| \leq \mu \big| \|\mathcal{F}X_t^i - \mathrm{VaR}_\gamma^i\| \leq \mu, \mathcal{B}(\mathcal{F}X_t^j)\right) \neq$$
$$P\left(\|\mathcal{G}X_t^j - \mathrm{VaR}_\gamma^j\| \leq \mu \big| \|\mathcal{F}X_t^i - Median^i\| \leq \mu, \mathcal{B}(\mathcal{F}X_t^j)\right),$$

where $\mu > 0$, $\|.\|$ is the Euclidian distance measure, $\mathcal{G}$ denotes a set of future realizations and $\mathcal{F}$ denotes a set of past and/or current realizations of the corresponding variables and $\mathcal{B}(.)$ reflects some event over the argument.

In this study, we consider two possible scenarios. In the first, we assume that institution j is already in distress, so that potential Granger causal risk effects from institution i do not only induce even higher losses on j but also can clog its recovery. The second scenario is more similar to the traditional risk analysis, where future troubles in institution j come directly from the past problems of institution j. One may thereof reformulate Def. 4.2.1 in the form of two possible scenarios, which we investigate in detail below.

Scenario 1. *Given any stationary bivariate process* $\{(X_t^i, X_t^j)\}$*, we say that* $\{X_t^i\}$ *is a nonlinear CoVaR Granger cause of* $\{X_t^j\}$ *in tail if*

$$P\left(\|\mathcal{G}X_t^j - \mathrm{VaR}_\gamma^j\| \leq \mu \big| \|\mathcal{F}X_t^i - \mathrm{VaR}_\gamma^i\| \leq \mu, \|\mathcal{F}X_t^j - \mathrm{VaR}_\gamma^j\| \leq \mu\right) \neq$$
$$P\left(\|\mathcal{G}X_t^j - \mathrm{VaR}_\gamma^j\| \leq \mu \big| \|\mathcal{F}X_t^i - Median^i\| \leq \mu, \|\mathcal{F}X_t^j - \mathrm{VaR}_\gamma^j\| \leq \mu\right),$$

where $\mu > 0$*,* $\|.\|$ *is the Euclidian distance measure,* $\mathcal{G}$ *denotes a set of future realizations and* $\mathcal{F}$ *denotes a set of past and/or current realizations of the corresponding variables.*

Scenario 2. *Given any stationary bivariate process* $\{(X_t^i, X_t^j)\}$*, we say that* $\{X_t^i\}$ *is a nonlinear CoVaR Granger cause of* $\{X_t^j\}$ *in median if*

$$P\left(\|\mathcal{G}X_t^j - \mathrm{VaR}_\gamma^j\| \leq \mu | \|\mathcal{F}X_t^i - \mathrm{VaR}_\gamma^i\| \leq \mu, \|\mathcal{F}X_t^j - Median^j\| \leq \mu\right) \neq$$
$$P\left(\|\mathcal{G}X_t^j - \mathrm{VaR}_\gamma^j\| \leq \mu | \|\mathcal{F}X_t^i - Median^i\| \leq \mu, \|\mathcal{F}X_t^j - Median^j\| \leq \mu\right),$$

where $\mu > 0$*,* $\|.\|$ *is the Euclidian distance measure,* $\mathcal{G}$ *denotes a set of future realizations and* $\mathcal{F}$ *denotes a set of past and/or current realizations of the corresponding variables.*

In practice it is impossible to condition on the infinite sets of future or past realizations of variables of interest. Therefore, we reformulate $\mathcal{G}$ and $\mathcal{F}$ as finite sets of future periods or lags, respectively. We limit ourselves to the canonical setting where $\mathcal{G}X_t^j = X_{t+1}^j$, as it is most commonly used in practical Granger causality testing, however, our reasoning holds for any $\mathcal{G}X_t^j = X_{t+k}^j$, $1 \leq k < \infty$. Similarly, we replace $\mathcal{F}X_t^i$ and $\mathcal{F}X_t^j$ by $X_{t,l_i}^i = \{X_{t-l_i+1}^i, \ldots, X_t^i\}$ and $X_{t,l_j}^j = \{X_{t-l_j+1}^j, \ldots, X_t^j\}$, where $l_i \geq 1$ and $l_j \geq 1$ denote the number of lags of a corresponding variable.

In Granger causality testing, the goal is to find evidence against the null hypothesis of no causality, which according to Def. 4.2.1 is represented by equivalence in conditional probability. We assume that process $\{(X_t^i, X_t^j)\}$ is strictly stationary. In that case, the null hypothesis is a statement about the invariant distribution evaluated at conditional VaR_γ levels of the $(l_i + l_j + 1)$ -dimensional vector $W_t = (Z_t, X_{t,l_i}^i, X_{t,l_j}^j)$, where we substitute $Z_t = X_{t+1}^j$. (For clarity

purposes and to bring forward the fact that we consider the invariant distribution of W_t, we drop the time index, so that $W = (Z, X^i, X^j)$.) Formally, the null hypothesis from Scenarios 1 and 2 can be rewritten as

$$f_{Z,X^i,X^j}\left(z_\gamma | x^i_\gamma, x^j_*\right) = f_{Z,X^i,X^j}\left(z_\gamma | x^i_m, x^j_*\right), \tag{4.3}$$

where $z_\gamma = \mathrm{VaR}^Z_\gamma$, $x^i_\gamma = \mathrm{VaR}^i_\gamma$, $x^i_m = Median^i$ and $*$ distinguishes between Scenario 1 and 2 as $x^j_\gamma = \mathrm{VaR}^j_\gamma$ or $x^j_m = Median^j$, respectively. It is helpful to restate the problem in terms of ratios of joint densities evaluated at given quantiles, as under the null the density of Z evaluated around its VaR_γ level and conditional on specific events in X^i and X^j is equal to the same density conditional on the different set of events in X^i and X^j. Therefore, the joint probability density function, together with its marginals must satisfy

$$\frac{f_{Z,X^i,X^j}\left(z_\gamma, x^i_\gamma, x^j_*\right)}{f_{X^i,X^j}\left(x^i_\gamma, x^j_*\right)} = \frac{f_{Z,X^i,X^j}\left(z_\gamma, x^i_m, x^j_*\right)}{f_{X^i,X^j}\left(x^i_m, x^j_*\right)}. \tag{4.4}$$

Since Eq. (4.4) holds for any quantile of the vector (Z, X^i, X^j) in the support of Z, X^i, X^j, Eq. (4.4) might be equivalently rewritten as

$$\frac{f_{Z,X^i,X^j}\left(z_\gamma, x^i_\gamma, x^j_*\right)}{f_{X^i,X^j}\left(x^i_m, x^j_*\right)} = \frac{f_{X^i,X^j}\left(x^i_\gamma, x^j_*\right)}{f_{X^i,X^j}\left(x^i_m, x^j_*\right)} \frac{f_{Z,X^i,X^j}\left(z_\gamma, x^i_m, x^j_*\right)}{f_{X^i,X^j}\left(x^i_m, x^j_*\right)}. \tag{4.5}$$

Analogously to Baeck and Brock (1992) or Hiemstra and Jones (1994), a natural methodology to assess Eq. (4.5) comes from the test for conditional independence. However, as showed by Diks and Panchenko (2005) and Diks and Panchenko (2006), these tests can severely over-reject in Granger causal setting, because its dependence on the conditional variance. Diks and Panchenko (2006) propose to add a positive weight function $g(z_\gamma, x^i_m, x^j_*)$ and, given that the

null should hold in the support of the joint densities, it might be equivalently written as

$$\tau_g \equiv \left(\frac{f_{Z,X^i,X^j}\left(z_\gamma, x^i_\gamma, x^j_*\right)}{f_{X^i,X^j}\left(x^i_m, x^j_*\right)} - \frac{f_{X^i,X^j}\left(x^i_\gamma, x^j_*\right)}{f_{X^i,X^j}\left(x^i_m, x^j_*\right)} \frac{f_{Z,X^i,X^j}\left(z_\gamma, x^i_m, x^j_*\right)}{f_{X^i,X^j}\left(x^i_m, x^j_*\right)} \right) g(z_\gamma, x^i_m, x^j_*) = 0. \tag{4.6}$$

Diks and Panchenko (2006) discuss several possibilities of choosing $g(z_\gamma, x^i_m, x^j_*)$. In this study we focus on $g(z_\gamma, x^i_m, x^j_*) = f_{X^i,X^j}\left(x^i_m, x^j_*\right)^2$, as the estimator of τ_g has a corresponding U-statistic representation, bringing the desired asymptotic normality properties for weakly dependent data. Substituting into Eq. (4.6), one finds that

$$\tau = f_{Z,X^i,X^j}\left(z_\gamma, x^i_\gamma, x^j_*\right) f_{X^i,X^j}\left(x^i_m, x^j_*\right) - f_{X^i,X^j}\left(x^i_\gamma, x^j_*\right) f_{Z,X^i,X^j}\left(z_\gamma, x^i_m, x^j_*\right). \tag{4.7}$$

To evaluate the data-driven representation of τ, we rely on kernel methods. In particular, we consider the local density estimator

$$\hat{f}_W(w) = \frac{\varepsilon^{-d_W}}{n} \sum_{k=1}^{n} K\left(\frac{w - w_k}{\varepsilon}\right), \tag{4.8}$$

where n is the sample size, ε is the bandwidth parameter (similar to μ from the Def. 4.2.1), d reflects the dimensionality of a given vector W and $K(.)$ is a bounded Borel function $\mathbb{R}^{d_W} \to \mathbb{R}$ satisfying

$$\int |K(t)|dt < \infty, \qquad \int K(t)dt = 1 \qquad \text{and} \qquad |tK(t)| \to 0 \quad \text{as} \quad |t| \to \infty. \tag{4.9}$$

In practice, $K(.)$ is often chosen to be a probability density function (Wand and Jones, 1995). In order to guarantee the consistency of the pointwise density estimators, we assume that the bandwidth parameter ε comes from the sequence ε_n, which is slowly decreasing with the sample size, i.e.

$$\varepsilon_n \to 0 \qquad \text{and} \qquad n\varepsilon_n \to \infty \quad \text{as} \quad n \to \infty. \tag{4.10}$$

Parzen (1962) shows that under conditions (4.9) and (4.10) and provided that f is continuous at w, the estimate of density f at a given point w is consistent.

Given a given bandwidth ε, a natural estimator for τ is

$$T_n(\varepsilon) = C \sum_{k=1}^{n} \sum_{p=1}^{n} \left[K\left(\frac{(z_\gamma, x^i_\gamma, x^j_*)^T - (z_k, x^i_k, x^j_k)^T}{\varepsilon} \right) K\left(\frac{(x^i_m, x^j_*)^T - (x^i_p, x^j_p)^T}{\varepsilon} \right) - K\left(\frac{(x^i_\gamma, x^j_*)^T - (x^i_k, x^j_k)^T}{\varepsilon} \right) K\left(\frac{(z_\gamma, x^i_m, x^j_*)^T - (z_p, x^i_p, x^j_p)^T}{\varepsilon} \right) \right], \tag{4.11}$$

where ε is the bandwidth and

$$C = \frac{\varepsilon^{-d_Z - 2d_{X^i} - 2d_{X^j}}}{n^2}. \tag{4.12}$$

(We sum over two indices as it allows to calculate the variance of $T_n(\varepsilon)$ explicitly.) The asymptotic distribution of the test statistic can be derived from the behavior of the properties of the second order U-statistic, as described by Serfling (1980) and van der Vaart (1998).

Theorem 4.2.1. *Under the conditions described by Eqs. (4.9) and (4.10), for a given set of* VaR_γ *levels and given bandwidth parameter sequence* ε_n*, test statistic* $T_n(\varepsilon_n)$ *satisfies:*

$$\sqrt{n} \frac{(T_n(\varepsilon_n) - \tau)}{S_n} \xrightarrow{d} \mathcal{N}(0, 1),$$

where S_n *is a heteroskedasticity and autocorrelation consistent estimator of the asymptotic standard deviation of* $\sqrt{n}(T_n(\varepsilon_n) - \tau)$.

The proof of Theorem 4.2.1 can be found in Appendix 4.A. As argued by Diks and Panchenko (2006), although the test statistic is not positive definite, the one-sided version of the test, i.e. rejecting on larger values, turns out to yield better performance.

In this study we choose γ to be 0.05 as it is the most commonly applied VaR significance level. We calculate VaR_γ from the empirical quantile function (Jones, 1992). Following the literature on nonparametric Granger causality testing (Hiemstra and Jones, 1994; Diks and Panchenko, 2006) we take the square kernel function.[1] The square kernel form of the estimator

[1]The asymptotic properties of the test statistic are, however, robust to any kernel specification, provided that it

in Eq. (4.8), can be rewritten as

$$\hat{f}_W^{SQ}(w) = \frac{(2\varepsilon)^{-d_W}}{n-1} \sum_{k=1}^{n} I(\|w - w_k\| < \varepsilon), \tag{4.13}$$

where $I(\|w - w_k\| < \varepsilon)$ is the indicator function taking values 1 for any $\|w - w_k\| < \varepsilon$ and zero otherwise, and $\|.\|$ is the supremum norm over all the dimensions.

4.2.1 Optimal bandwidth

Although the asymptotic normality of the test statistic holds for an arbitrary decreasing sequence of bandwidths as long as it satisfies condition from Eq. (4.10), it influences the power of the test to a great extent (Silverman, 1998). Therefore, in order to improve the performance of the test, we calculate the optimal size of the bandwidth explicitly. Following Wand and Jones (1995) and Silverman (1998), the optimal bandwidth minimizes the Mean Squared Error (MSE) of $T_n(\varepsilon_n)$, which may be decomposed into the sum of variance and squared bias of $T_n(\varepsilon_n)$. In our inference it is worthwhile to point out that the optimal bandwidth values of $T_n(\varepsilon_n)$ do not violate the consistency properties of any of the density estimators.

Corollary 4.2.1. *Under the conditions given by Eqs. (4.9) and (4.10), the MSE-optimal sequence of bandwidths of $T_n(\varepsilon_n)$ guarantees consistency of any of the pointwise density estimators contributing to $T_n(\varepsilon_n)$.*

The proof of Corollary 4.2.1 is given in Appendix 4.B. In fact, the MSE optimum rate of convergence of the bandwidth of $T_n(\varepsilon_n)$ is slightly faster than that of individual density estimators, but still much slower than n^{-1}. This is caused by the increased variance of a product of two estimators compared to their individual variances. Therefore, in order to balance this effect in the MSE, the sequence of optimal bandwidths of $T_n(\varepsilon_n)$ should decrease at a slightly faster rate as $n \to \infty$, but never as fast as n^{-1}. In testing for systemic risk this proves to be of

satisfies conditions (4.9) and (4.10).

large importance as with a bandwidth parameter decreasing just slightly with the sample size we are still able to capture the majority of returns which are left to VaR_γ.

In evaluating the optimal bandwidth value we rely on Monte Carlo methods. Correcting for the weak dependency, we apply the autocorrelation consistent estimator for the variance of $T_n(\varepsilon)$, as proposed in Newey and West (1987). It might be verified that for a given bandwidth ε, the bias of $T_n(\varepsilon)$ may be calculated from the Taylor expansion around any point as

$$\begin{aligned} E[T_n(\varepsilon)] - \tau = \frac{1}{2}\kappa_2\varepsilon^2 \big[& f_{Z,X^i,X^j}\left(z_r, x_r^i, x_*^j\right) \nabla^2 f_{X^i,X^j}\left(x_s^i, x_*^j\right) \\ & + f_{X^i,X^j}\left(x_s^i, x_*^j\right) \nabla^2 f_{Z,X^i,X^j}\left(z_r, x_r^i, x_*^j\right) \\ & - f_{X^i,X^j}\left(x_r^i, x_*^j\right) \nabla^2 f_{Z,X^i,X^j}\left(z_r, x_s^i, x_*^j\right) \\ & - f_{Z,X^i,X^j}\left(z_r, x_s^i, x_*^j\right) \nabla^2 f_{X^i,X^j}\left(x_r^i, x_*^j\right) \big] + o(\varepsilon^2), \end{aligned} \tag{4.14}$$

where κ_2 is the second moment of the kernel and $\nabla^2 f_W(w)$ is the trace of the second derivative of density evaluated at point w. Up to the error of order $o(\varepsilon^2)$, Eq. (4.14) has a plug-in estimator, which can be easily calculated using kernel methods (Wand and Jones, 1995).

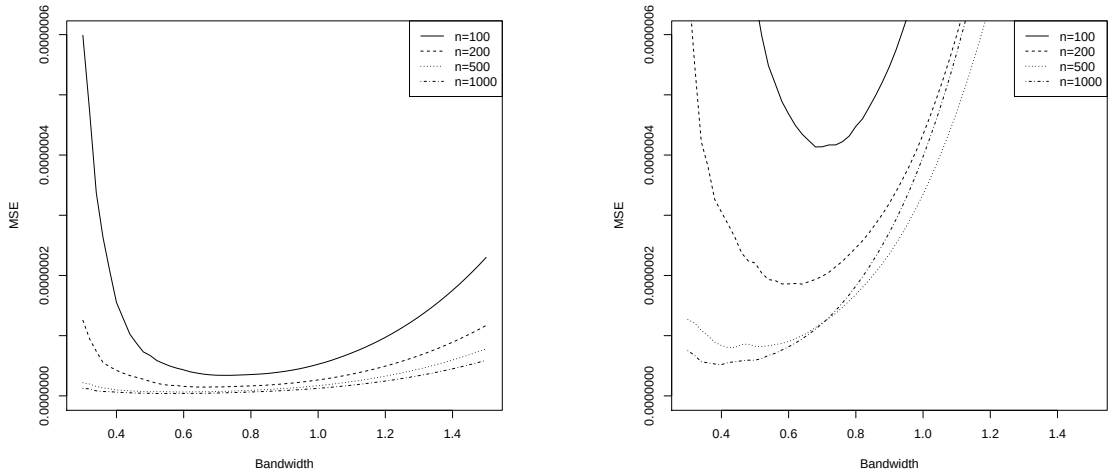
4.3 Numerical simulations

To give an example of the optimal bandwidth value, we perform a numerical experiment on the same bivariate process as considered by Jeong et al. (2012), i.e.

$$\begin{aligned} x_t^i &= 1 + \frac{1}{2}x_{t-1}^i + r_{1,t} \\ x_t^j &= \frac{1}{2}x_{t-1}^j + c\left(x_{t-1}^i\right)^2 + r_{2,t}, \end{aligned} \tag{4.15}$$

where $r_{1,t}$ and $r_{2,t}$ independent standard normal variables. The biggest advantage of the process in Eq. (4.15) is its tuning parameter on Granger causality, c. Clearly, if $c = 0$ the model corresponds to the hypothetical scenario of no Granger causality from X_t^i to X_t^j. The larger the parameter c becomes, the stronger the Granger causal effect, which we thus may control for

Figure 4.1: MSE of the test statistic for bandwidth values in the range $[0.3, 1.5]$ and for different sample sizes, aggregated over 1000 simulations.



(a) Null hypothesis as in Sc. 1

(b) Null hypothesis as in Sc. 2

explicitly.

We perform 1000 simulations of normalized data of process given by Eq. (4.15) for different sample sizes and evaluate the MSE of the test statistic for different bandwidth values within the range $[0.3, 1.5]$.[2] For practical reasons, we take lags of order 1 for both variables. The results for two scenarios of Granger causality are presented in Fig. 4.1 and the optimal bandwidths are reported in Table 4.1.

It is straightforward to notice the differences of the MSE curves between two settings. Firstly, for the same sample size and ε, Scenario 2 demonstrates larger MSE than in Scenario 1. Secondly, in Scenario 1 the MSE curve becomes flatter, whereas in Scenario 2 the visible U-shape is preserved as the sample size increases. These, in fact, are direct consequences of the curvature of the true distribution around particular quantiles. Scenario 1 is driven by the tail dependence, where the curvature is relatively flat. On the contrary, Scenario 2 represents the relation between the tail and the median, where the distribution is typically more bell-shaped or simply steeper. This, in fact, shows up in the steepness and in the relative size of the MSE curve. As expected, the minimum of the MSE curves is decreasing with the sample size in both

[2]We apply the standard score normalization.

scenarios (see Table 4.1).

Table 4.1: Optimal bandwidth values for test statistic evaluated for the process given by Eq. (4.15) for different sample sizes and for two scenarios. The values represent means over 1000 simulations.

	$n = 100$	$n = 200$	$n = 500$	$n = 1000$
ε^* (Sc.1)	0.74	0.66	0.6	0.52
ε^* (Sc.2)	0.68	0.64	0.48	0.44

The reported optimal bandwidth values represent the radius around the VaR_γ which is being considered in the NCoVaR. One may readily observe that Scenario 1 has slightly larger optimal bandwidths than Scenario 2. We may view this as a result of scarcity of data in tails compared with that around the median. Extracting information from tails requires, on average, slightly larger windows in comparison to the region near the median (Caers and Maes, 1998).

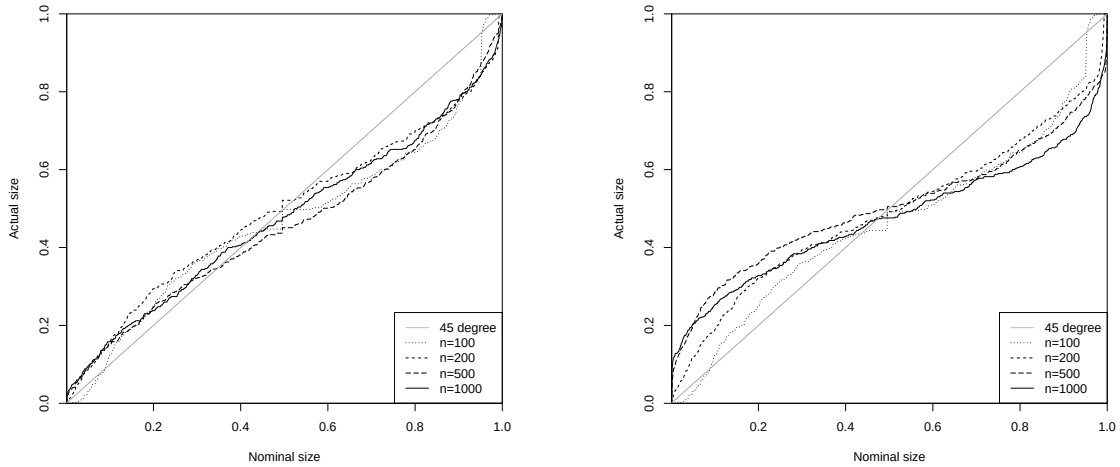
Because the MSE of the test statistics might be calculated explicitly, bootstrapping optimal bandwidths is a powerful technique which might be applied to any data set without assuming an underlying process structure. We recognize, however, that it might take a lot of computational time. For very large samples we suggest taking bins of 0.02 or 0.05 in order to make it computationally less demanding. Our simulations confirm that the power of the test is preserved in the range $[\varepsilon^* - 0.05, \varepsilon^* + 0.05]$.

4.3.1 Performance of the NCoVaR test

We perform two experiments to evaluate the practical side of the test. In both we rely on Monte Carlo methods on the example of the process in Eq. (4.15).[3] In the first one, we assess the distribution of the test statistic under the null, evaluated for different sample sizes for 500 runs. In the second experiment, we estimate the power of the test. Given that the null hypothesis is violated ($c > 0$), we estimate rejection rates for different nominal significance levels. We

[3]One may expect that the numerical size distortions and power of the NCoVaR test would depend on the exact process specification. Eq. (4.15) offers a simple testing environment, which has been already applied in the quantile testing literature (Jeong et al., 2012). We therefore leave the assessment of the NCoVaR numerical performance on other processes for future investigation.

Figure 4.2: Size-size diagram of the NCoVaR test for the process from Eq. (4.15) for different sample sizes over 500 simulations.



(a) Null hypothesis as in Sc. 1

(b) Null hypothesis as in Sc. 2

summarize the results from both experiments in the size-size plots and size-adjusted power diagrams. The former plots the actual against nominal cumulative rejection rates under the null, showing the size distortions. The latter shows the power of the test corrected for the possible size bias, plotting the observed cumulative rejection rates under the alternative (actual power) against observed rejection rates under the null (actual size). Ideally, the power function should be 1 for any significance level larger than 0, however, in practice we would like to observe an increase in the slope at the origin as the sample size grows. Fig. 4.2 shows the size-size diagrams whereas the size-adjusted power plots are presented in Figs 4.3-4.5.

Fig. 4.2 suggests that the nominal size distortions are larger in Scenario 2 than in Scenario 1. Additionally, the size-size curves are flatter in Scenario 1 whereas they are more wavy in Scenario 2. In fact, this is similar to the pattern observed in the MSE (see Fig. 4.1) and might be largely attributed to the curvature of the true probability density function around particular quantiles.

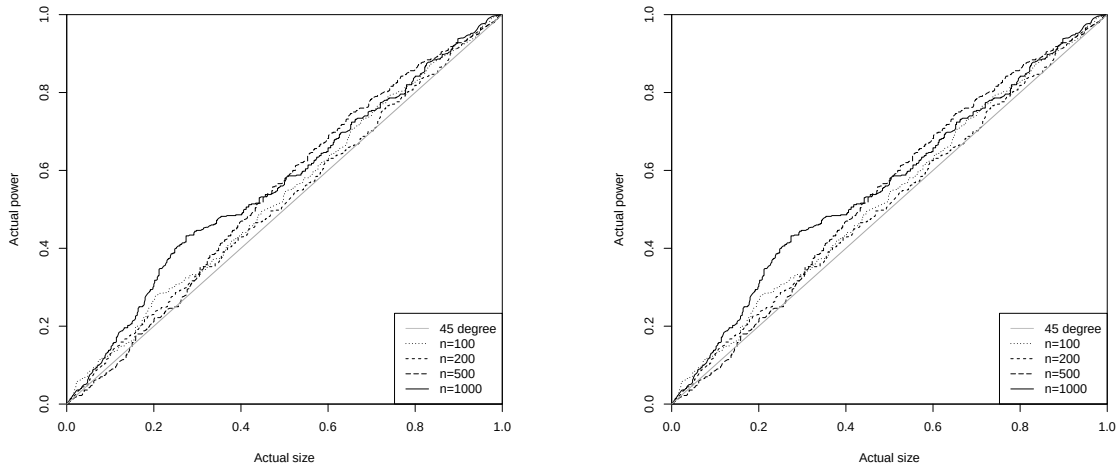
One can readily observe from Figs 4.3-4.5 that the size-adjusted power of the test increases with the sample size and with the strength of Granger causality. Nevertheless, there are two

main patterns emerging from the numerical analysis which deserve to be pointed out.

Firstly, for relatively smaller size the power of the test is higher for Scenario 1 than for Scenario 2. This is again the result of model dynamics, where the underlying relation on variable j, i.e. $(X^j_{t+1} \approx \mathrm{VaR}^j_\gamma | X^j_t \approx Median^j)$ is more rare to observe on the process given by Eq. (4.15). Practically speaking, as the sample size gets larger this effect is hampered.

Secondly, the size-adjusted power is almost negligible for very small Granger causality and short time series. Clearly, one should blame the relative scarcity of observations around quantiles for this discomfort. In order to apply the test to shorter data sets, we propose two solutions to overcome this issue. The first comprises different kernel specifications. The square kernel takes into account only observations which are ε-close to the quantile, leaving out many possibly informative data points. Replacing the kernel by a smoother one, like Gaussian or logistic, should therefore correct for this effect. The second possible solution lies in improving the precision of the density estimators. In the standard kernel estimators (like square kernel estimators applied here) the bias is of order ε^2 (Wand and Jones, 1995). Making the bias smaller should decrease the disinformative effect of the observations around a given quantile so that keeping the sample size fixed we get relatively better representation of the true Granger causal relation, which translates into improved test performance. One may consider Data Sharpening (DS) as being potentially attractive bias reduction method in our setting. Following Hall and Minnotte (2002), the idea behind DS is to slightly perturb the original data set in order to obtain desirable estimator properties (here it is the reduced bias). Diks and Wolski (2013) show that, besides reducing the estimator bias, DS does not affect other asymptotic properties of the test statistic in a similar Granger causality setting. Therefore, it seems to be a straightforward extension to NCoVaR for shorter samples.

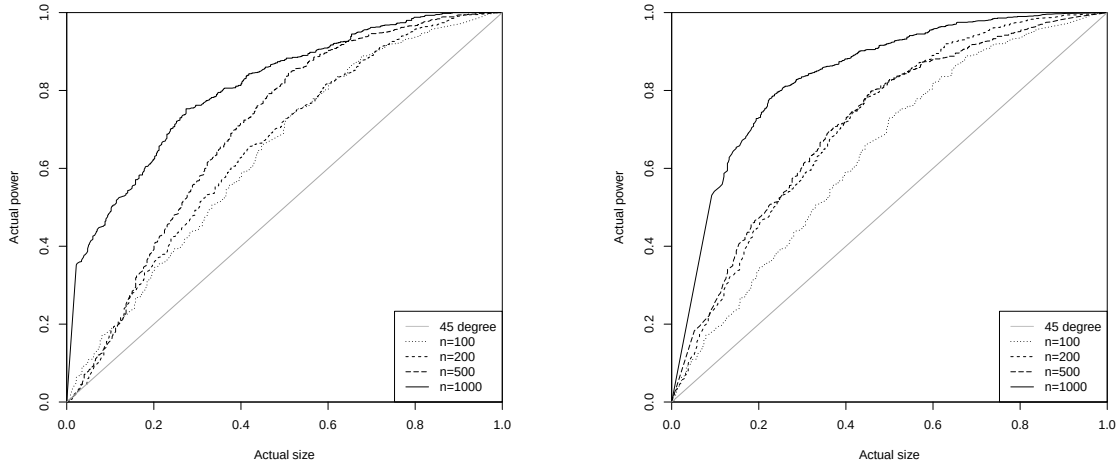
Figure 4.3: Size-adjusted power for the NCoVaR test for the process given by Eq. (4.15) for $c = 0.05$ for different sample sizes over 500 simulations.



(a) Null hypothesis as in Sc. 1

(b) Null hypothesis as in Sc. 2

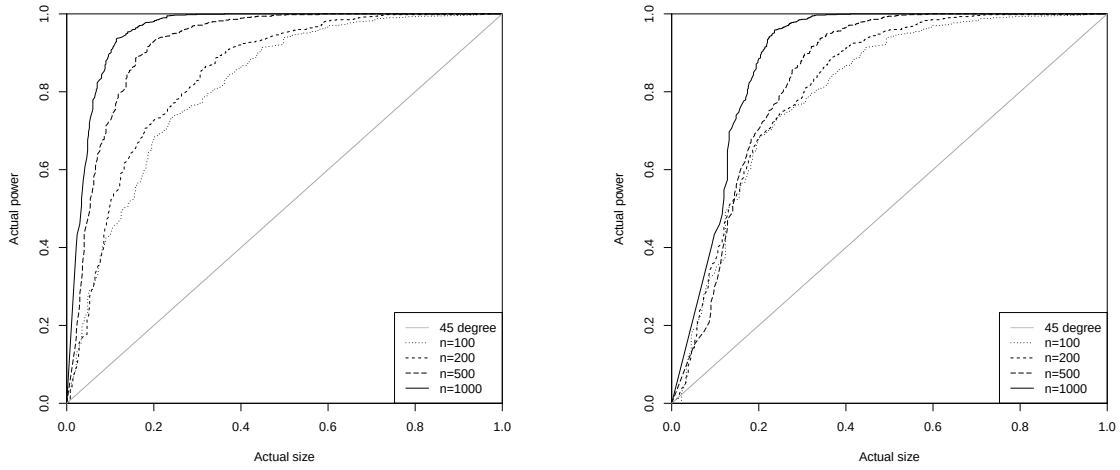
Figure 4.4: Size-adjusted power for the NCoVaR test for the process given by Eq. (4.15) for $c = 0.25$ for different sample sizes over 500 simulations.



(a) Null hypothesis as in Sc. 1

(b) Null hypothesis as in Sc. 2

Figure 4.5: Size-adjusted power for the NCoVaR test for the process given by Eq. (4.15) for $c = 0.4$ for different sample sizes over 500 simulations.



(a) Null hypothesis as in Sc. 1

(b) Null hypothesis as in Sc. 2

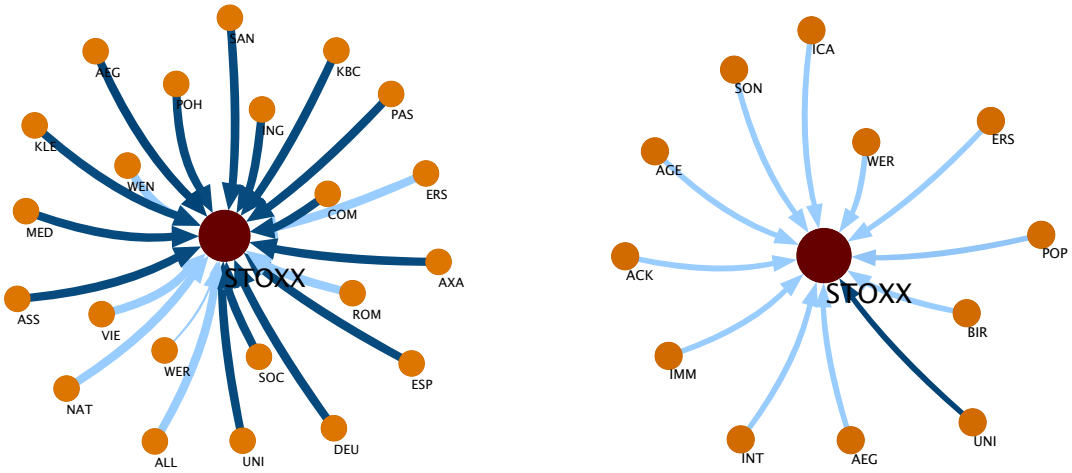
4.4 Assessing financial systemic risk

In our analysis we focus on the NCoVaR of individual institutions on the overall systemic risk. Therefore, we set j to represent the system variable and i individual financial institutions.

We approximate the returns on assets by equity returns and take into account financial institutions publicly traded within the euro zone. In order to make the analysis more transparent we focus on companies which constitute the Euro STOXX Financial Index in years 2000-2012. Our sample thus covers the Great Recession in Europe (2008-12), the financial crisis (2007-2009) and the sovereign debt crisis (2010-2012). In total we collect daily equity returns for 48 companies (3 financial, 13 insurance, 23 banks and 9 real estate) and one aggregate index. For each variable we have 3390 observations. The list of companies, together with the country of origin and their sector can be found in Appendix 4.C. The data have been obtained from the DataStream.

All time series are stationary at the 1% significance level, according to both the Phillips-Perron and Augmented Dickey-Fuller specifications (Phillips and Perron, 1988; Fuller, 1995). We run the pairwise tests against the null of no NCoVaR between each company and system

Figure 4.6: NCoVaR between euro area individual financial companies and system variable for raw data.



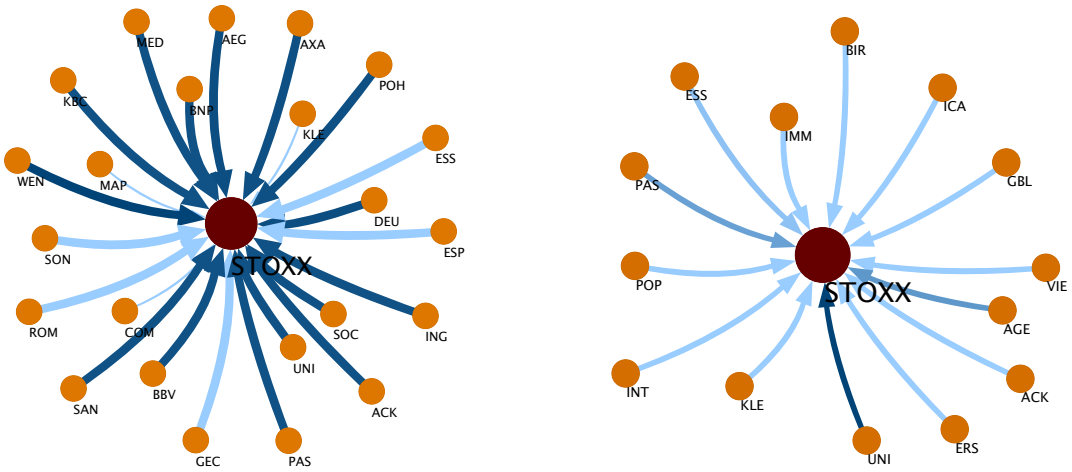
(a) Euro area NCoVaR in Sc. 1

(b) Euro area NCoVaR in Sc. 2

variable. In order to make sure that all the Granger causal relations are nonlinear, we run the same test specification on VAR-filtered residuals also. In each run the number of lags is taken according to the Schwarz-Bayes Information Criterion of the VAR specification, and the optimal bandwidth value is approximated by bootstrap. As a robustness check, we also correct for possible causality in second moments, as suggested in Francis et al. (2010), by running NCoVaR test on residuals from Dynamics Conditional Correlation GARCH model (Engle, 2002).

The detailed results can be found in Appendix 4.C (Tables 4.C.2, 4.C.3 and 4.C.3), however, for presentational clarity we refer to the star-graphs, which show the NCoVaR between each company and the system as a whole. The center of the star-graph represents the system variable and the satellite nodes correspond to individual institutions. The width of the arrows represents the inverse of the statistical significance level of NCoVaR (the stronger the NCoVaR effect, the wider (and darker) the arrow). Fig. 4.6 shows the results for the raw returns, Fig. 4.7 depicts the VAR-filtered returns and Fig. 4.8 refers to the GARCH residuals. Considering that at least one NCoVaR relation denotes a systemically important institution, our analysis suggests that out of 48 companies 33 might be so described. The group consists of 3 financial services companies,

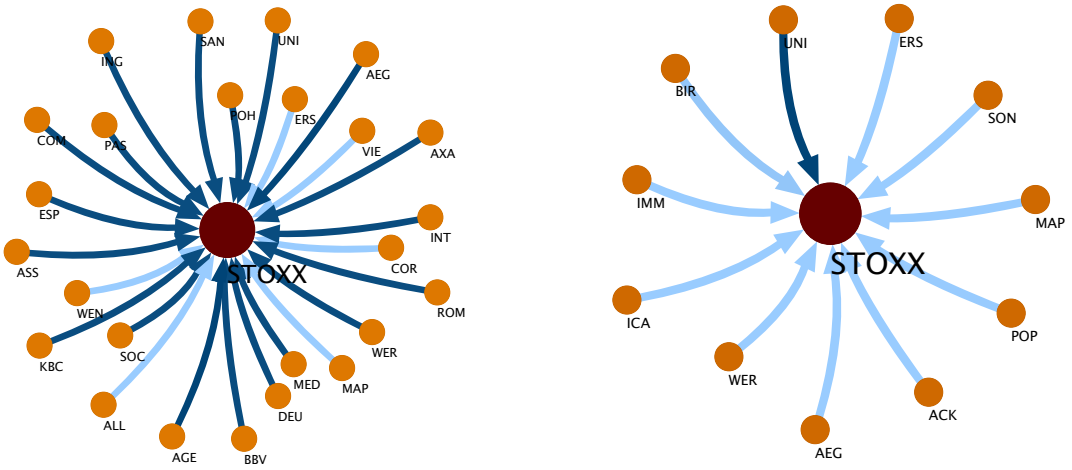
Figure 4.7: NCoVaR between euro area individual financial companies and system variable for VAR-filtered data.



(a) Euro area NCoVaR in Sc. 1

(b) Euro area NCoVaR in Sc. 2

Figure 4.8: NCoVaR between euro area individual financial companies and system variable for GARCH-filtered data.



(a) Euro area NCoVaR in Sc. 1

(b) Euro area NCoVaR in Sc. 2

6 insurance firms, 19 banks and 5 real estate companies. In fact, all of the financial services companies in our sample prove to be systemically important.

There are two main patterns emerging from our analysis. Firstly, there are fewer systemically risky institutions in Scenario 2. Secondly, NCoVaR in Scenario 1 is on average stronger than in Scenario 2. These findings hold for the original as well as the VAR- and GARCH-filtered data. Interestingly, our study suggests that only a few financial institutions pose a serious *ex ante* threat to the systemic risk in the euro area, whereas, given that the system is already in trouble, there are more institutions which hamper its recovery. This result confirms a common view in the literature on macroprudential supervision (Acharya, 2009) that the relative preventive costs are smaller than those after the crisis has already erupted.

The analysis confirms the nonlinear structure of the institutional contribution to the systemic risk. Filtering out the linear relations and second moment spillover effects does not remove the co-risk relations among individual companies and system as a whole. Interestingly, after filtering we observe some new co-risk relations emerging. To illustrate this better let us consider ACK (Ackermans & Van Haaren). The raw returns do not show any NCoVaR, however, after linear filtering it poses a very strong threat to the system's recovery (see Table 4.C.3 in Appendix 4.C which shows a test statistic of order 6.351 in Scenario 1) and after GARCH filtering it has a weak *ex ante* effect on the system's risk (test statistic of order 1.329 in Scenario 2). One may speculate that there are some strong purely nonlinear and second moment co-risk effects from ACK on the system variable, which are being partly offset by their linear equivalents. In other words, under normal circumstances ACK does not seem to be an important systemic risk contributor. However, in abnormal times, like a crisis, it reveals its systemic importance.

There is one more finding which we believe is worth pointing out. We confront our results with the official list of Global Systemically Important Banks (G-SIBs), published by the Financial Stability Board (FSB) in 2011.[4] The FSB recognizes 11 G-SIBs in the euro area. Our sample covers 8 of them, i.e. Banco Bilbao Vizcaya Argentaria (BBV), Banco Santander

[4]The G-SIBs list is being frequently updated. In our comparison we consider the most recent version of the list, published on November 11th, 2013.

(SAN), BNP Paribas (BNP), Commerzbank (COM), Deutsche Bank (DEU), Societe Generale (SOC), UniCredit (Uni) and ING Bank (ING), as a part of the ING Groep. For all of them we confirm their G-SIB status in at least one NCoVaR setting.

4.5 Conclusions and discussion

Conditional Value-at-Risk-Nonlinear Granger Causality, or NCoVaR, is a new methodology of assessing co-risk relations, designed to capture the possible nonlinear Granger causal effects. Our approach distinguishes between two possible scenarios. In the first one, we test what is the role of individual institutions in hampering the recovery of others, given that they are already in distress. In the second scenario, we assess the contribution of individuals to the others' troubles. We derive the regular asymptotic properties of the NCoVaR test for both scenarios and we confirm them numerically.

We apply our methodology to assess the systemic importance of financial institutions in the euro area. Our findings suggest that (i) only a few financial institutions pose a serious *ex ante* threat to the systemic risk, whereas, given that the system is already in trouble, there are more institutions which hamper its recovery and (ii) there are intriguing nonlinear structures in its systemic risk profile.

Our study suggests that the most systemically risky institution in our sample is UNI (UniCredit), an Italian bank. In all settings it demonstrates a very strong NCoVaR relation to the system. In 2011 it was recognized by the FSB as G-SIB. This analysis confirms its systemical importance, also revealing its nonlinear nature. Interestingly, there are two more companies which demonstrate very strong NCoVaR in 5 out of 6 settings, i.e. ERS (Erste Group Bank), an Austrian bank, and AEG (Aegon), a Dutch insurer. Only the latter was recognized by the FSB to be potentially systemically important, with no official view on the former. However, the former was recognized as a systemically important bank for the Austrian financial sector (von Kruechten et al., 2009). Our results point to potential systemic importance of Erste Group Bank

in the entire euro area.

NCoVaR might be of great use for macroprudential policy, however, it has to be tested on other samples and in other periods. It reveals some intriguing phenomena in the co-risk relations. In order to understand these better, a tempting idea is to investigate the underlying nonlinear structures analytically in models of the aggregate economy. Such settings would allow to capture not only the contribution of individual institutions to systemic risk but also how individual companies are affected by aggregate disturbances. One may also apply NCoVaR as a mapping tool and bring the risk analysis to the network level.

Appendix 4.A Asymptotic properties of test statistic (Theorem 4.2.1)

We first deal with the properties for the independent sample and consider the dependency later. By symmetrization with respect to two indices, the test statistic in Eq. (4.11) has a corresponding U-statistic representation of the form

$$T_n(\varepsilon_n) \equiv T_n(\varepsilon) = \frac{1}{\binom{n}{2}} \sum_{k=1}^{n} \sum_{p \leq k} \tilde{K}(W_k, W_p), \tag{4.16}$$

with $W_k = (Z_k, X^i_{k,l_i}, X^j_{k,l_j})$, $k = 1, \ldots, n$ and kernel given by

$$\begin{aligned} \tilde{K}(W_k, W_p) = \frac{\varepsilon^{-d_Z - 2d_{X^i} - 2d_{X^j}}(n-1)}{2n} \big[& K_k(z_\gamma, x^i_\gamma, x^j_*) K_p(x^i_m, x^j_*) \\ - & K_k(x^i_\gamma, x^j_*) K_p(z_\gamma, x^i_m, x^j_*) + K_p(z_\gamma, x^i_\gamma, x^j_*) K_k(x^i_m, x^j_*) \\ - & K_p(x^i_\gamma, x^j_*) K_k(z_\gamma, x^i_m, x^j_*) \big], \end{aligned}$$

where for clarity we denote $K_k(w) = K((w - w_k)/\varepsilon)$ and d_Z, d_{X^i} and d_{X^j} are general representations of the dimensionality of $\mathcal{G}$ and $\mathcal{F}$ operators for particular variables. It is worth to remind here that subscript n in the test statistic refers to its sequence.

The asymptotic properties of the sequence of test statistic can be derived by the projection method (van der Vaart, 1998). From the Háyek's projection lemma we know that the projection of $T_n(\varepsilon) - \tau$ on the set of all function of the form $\sum_{k=1}^{n} \kappa_k(W_k)$ is given by

$$\hat{T}_n(\varepsilon) = \sum_{k=1}^{n} E[(T_n(\varepsilon) - \tau)|W_k] = \frac{2}{n} \sum_{k=1}^{n} \tilde{K}_1(w_k), \tag{4.17}$$

where

$$\tilde{K}_1(w_k) = E_{W_p}\left[\tilde{K}(w_k, W_p))\right] - \tau. \tag{4.18}$$

Projection $\hat{T}_n(\varepsilon)$ is mean zero sequence with variance $4/n\text{Var}(\tilde{K}_1(W_1))$. By the Central Limit Theorem, one may verify that $\sqrt{n}\hat{T}_n(\varepsilon)$ converges in distribution to the normal law with

mean 0 and variance given by $4\text{Var}(\tilde{K}_1(W_1))$.

Provided that $\text{Var}(\hat{T}_n(\varepsilon)) \to \text{Var}(T_n(\varepsilon))$ as $n \to \infty$, by Slutsky's lemma, we now observe that for a given ε and given quantiles of any independent finite-variance process $(Z_t, X^i_{t,l_i}, X^j_{t,l_j})$, the sequence $\sqrt{n}\left(T_n(\varepsilon) - \tau - \hat{T}_n(\varepsilon)\right)$ converges in probability to zero as $n \to \infty$. What follows, the sequence $\sqrt{n}\left(T_n(\varepsilon) - \tau\right)$ converges in distribution to $\mathcal{N}\left(0, \sigma^2\right)$, where

$$\sigma^2 = 4\zeta_1, \tag{4.19}$$

with $\zeta_1 = \text{Cov}\left(\tilde{K}(W_1, W_2), \tilde{K}(W_1, W_2')\right) = \text{Var}(\tilde{K}_1(W_1))$.

Appendix 4.A.1 Dependence

Following the reasoning from Denker and Keller (1983), the above asymptotic normality properties of the test statistic, $T_n(\varepsilon)$, hold for a weakly dependent process if we take into account the covariance between estimators of particular vectors in the asymptotic variance σ^2,

$$\sigma^2 = 4\left[\zeta_1 + 2\sum_{t=2}^{n} \text{Cov}\left(\tilde{K}_1(W_1), \tilde{K}_1(W_t)\right)\right]. \tag{4.20}$$

According to the kernel specification, the estimator for $\tilde{K}_1(W_k)$ is given by

$$\hat{K}_1(W_k) = \frac{(2\varepsilon)^{-d_Z - 2d_{X^i} - 2d_{X^j}}}{n} \sum_{p=1}^{n} \tilde{K}(W_k, W_p).$$

The Newey and West (1987) heteroskedasticity and autocorrelation consistent estimator of σ^2 is

$$S_n^2 = \sum_{b=1}^{B} R_b \omega_b, \tag{4.21}$$

where B is equal to the floor of $n^{1/4}$, R_b is the sample covariance function of $\hat{K}_1(W_b)$ given by

$$R_b = \frac{1}{n-b} \sum_{a=1}^{n-b} (\hat{K}_1(W_a) - T_n(\varepsilon))(\hat{K}_1(W_{a+b}) - T_n(\varepsilon)), \tag{4.22}$$

and ω_b is the weight function of the form

$$\omega_b = \begin{cases} 1, & \text{if} \quad b = 1 \\ 2 - \frac{2(b-1)}{\tau}, & \text{if} \quad b > 1. \end{cases} \tag{4.23}$$

For any finite-variance process $(Z_t, X^i_{t,l_i}, X^j_{t,l_j})$, it follows from Denker and Keller (1983) that

$$\sqrt{n}\frac{(T_n(\varepsilon) - \tau)}{S_n} \xrightarrow{d} \mathcal{N}(0,1), \tag{4.24}$$

which completes the proof of Theorem 4.2.1.

Appendix 4.B Optimal bandwidth sequence (Corollary 4.2.1)

For a given bandwidth ε, the MSE of the test statistic might be rewritten as as sum of variance and squared bias (Wand and Jones, 1995), i.e.

$$\text{MSE}[T_n(\varepsilon)] = \text{Var}(T_n(\varepsilon)) + \text{Bias}(T_n(\varepsilon))^2, \tag{4.25}$$

where $\text{Bias}(T_n(\varepsilon))$ can be calculated explicitly from the Taylor expansion as in Eq. (4.14) and variance of the test statistic might be represented as $4S_n^2/n$ from Appendix 4.A.1. Asymptotic covariance terms tend to zero as $n \to \infty$ so that under the null one might find that the asymptotic

variance of $T_n(\varepsilon)$ might be decomposed into

$$\begin{aligned}
\text{Var}(T_n(\varepsilon)) = &\text{Var}(\hat{f}_{Z,X^i,X^j}\left(z_\gamma, x^i_\gamma, x^j_*\right))\text{Var}(\hat{f}_{X^i,X^j}\left(x^i_m, x^j_*\right)) \\
&+ \text{Var}(\hat{f}_{X^i,X^j}\left(x^i_\gamma, x^j_*\right))\text{Var}(\hat{f}_{Z,X^i,X^j}\left(z_\gamma, x^i_m, x^j_*\right)) \\
&+ \text{Var}(\hat{f}_{X^i,X^j}\left(x^i_m, x^j_*\right))E[\hat{f}_{Z,X^i,X^j}\left(z_\gamma, x^i_\gamma, x^j_*\right)]^2 \\
&+ \text{Var}(\hat{f}_{Z,X^i,X^j}\left(z_\gamma, x^i_\gamma, x^j_*\right))E[\hat{f}_{X^i,X^j}\left(x^i_m, x^j_*\right)]^2 \\
&+ \text{Var}(\hat{f}_{X^i,X^j}\left(x^i_\gamma, x^j_*\right))E[\hat{f}_{Z,X^i,X^j}\left(z_\gamma, x^i_m, x^j_*\right)]^2 \\
&+ \text{Var}(\hat{f}_{Z,X^i,X^j}\left(z_\gamma, x^i_m, x^j_*\right))E[\hat{f}_{X^i,X^j}\left(x^i_\gamma, x^j_*\right)]^2 + o(1).
\end{aligned}$$

One may find that the variance and bias of the individual density estimators are $o(n^{-1}\varepsilon^{-d_W})$ and $o(\varepsilon^{-2})$, respectively (Silverman, 1998). Therefore, the dominant terms in the asymptotic variance are of order $o(n^{-1}\varepsilon^{-d_Z-d_{X^i}-d_{X^j}-4})$.

Taking the first order conditions of the MSE of of individual density estimators, one finds that the optimum rate of convergence of bandwidth parameter is $n^{-1/(d_W+4)}$. Doing the same for our test statistic, we find that this rate is $n^{-1/(d_Z+d_{X^i}+d_{X^j})}$. Therefore, for any finite dimension, the MSE-optimal rate of convergence of the test statistic's bandwidth is slightly faster than those of individual density estimators but never as fast as n^{-1} which would violate condition imposed by Eq. (4.10). Provided that the optimum rate of convergence of the individual estimators is sufficient for the consistency (Silverman, 1998), the optimum rate of $T_n(\varepsilon_n)$ guarantees consistency as well.

Appendix 4.C Data description and results

The Euro STOXX Financials Index consists originally of 61 entities. However, only 48 of them cover years 2000-2012 (see Table 4.C.1). For all of them we collect daily equity prices and calculate their log returns accordingly. Data comes from the DataStream and covers period 01/01/2000 till 12/31/2012.

Table 4.C.1: List of all entities used in the empirical analysis.

	Company name/Index	Symbol	Sector	Country
1	Euro STOXX Financials	STOXX	Aggregate	Aggregate
2	Ackermans & Van Haaren	ACK	Financial Services	BE
3	Aegon	AEG	Insurance	NL
4	Ageas	AGE	Insurance	NL
5	Allianz	ALL	Insurance	DE
6	Assicurazioni Generali	ASS	Insurance	IT
7	AXA	AXA	Insurance	FR
8	Bank Of Ireland	BIR	Banks	IR
9	Bankinter	BAN	Banks	ES
10	Banca Monte Dei Paschi	PAS	Banks	IT
11	Banca Popolare Di Milano	MIL	Banks	IT
12	Banca Popolare Di Sondrio	SON	Banks	IT
13	Banca Popolare Emilia Romagna	ROM	Banks	IT
14	BBV Argentaria	BBV	Banks	ES
15	Banco Comr. Portugues	POR	Banks	PT
16	Banco Espirito Santo	ESS	Banks	PT
17	Banco Popolare	POP	Banks	IT
18	Banco Popular Espanol	ESP	Banks	ES
19	Banco Santander	SAN	Banks	ES
20	BNP Paribas	BNP	Banks	FR
21	CNP Assurances	CNP	Insurance	FR
22	Cofinimmo	COF	Real Estate	BE
23	Commerzbank	COM	Banks	DE
24	Corio	COR	Real Estate	NL
25	Deutsche Bank	DEU	Banks	DE
26	Erste Group Bank	ERS	Banks	AT
27	Fonciere Des Regions	FON	Real Estate	FR
28	Gecina	GEC	Real Estate	FR
29	GBL New	GBL	Financial Services	BE
30	Societe Generale	SOC	Banks	FR
31	Hannover Ruck.	HAN	Insurance	DE
32	ICADE	ICA	Real Estate	FR
33	Immofinanz	IMM	Real Estate	AT
34	ING Groep	ING	Insurance	NL
35	Intesa Sanpaolo	INT	Banks	IT
36	KBC Group	KBC	Banks	BE
37	Klepierre	KLE	Real Estate	FR
38	Mapfre	MAP	Insurance	ES
39	Mediobanca	MED	Banks	IT
40	Muenchener Ruck.	MUE	Insurance	DE
41	Natixis	NAT	Banks	FR
42	Pohjola Pankki	POH	Banks	FI
43	Sampo	SAM	Insurance	FI
44	SCOR	SCO	Insurance	FR
45	Unibail-Rodamco	ROD	Real Estate	FR
46	UniCredit	UNI	Banks	IT
47	Vienna Insurance Group	VIE	Insurance	AT
48	Wendel	WEN	Financial Services	FR
49	Wereldhave	WER	Real Estate	NL

Table 4.C.2: NCoVaR from institution i on the system risk in two scenarios in period 01/01/2000 till 12/31/2012 for raw returns. Lags determines the optimal number of lags from the VAR specification using the Schwarz-Bayes Information Criterion. Optimal epsilon values calculated from bootstrap. T-val represents the test statistic of NCoVaR from Eq.(4.11). (*),(**), (***) denotes one-sided p-value statistical significance at 10%, 5% and 1%, respectively.

			Scenario 1			Scenario 2		
	Institution i	System variable	Lags	Opt. ε	T-val	Lags	Opt. ε	T-val
1	ACK	STOXX	1	0.54	0.913	1	0.44	1.329*
2	AEG	STOXX	1	0.28	6.359***	1	0.4	2.855***
3	AGE	STOXX	4	0.6	0.213	4	0.6	3.516***
4	ALL	STOXX	1	0.38	1.45	1	0.2	-6.365
5	ASS	STOXX	1	0.24	6.351***	1	0.22	-6.359
6	AXA	STOXX	1	0.26	6.351***	1	0.2	-3.035
7	BIR	STOXX	2	0.6	0.585	2	0.6	2.611***
8	BAN	STOXX	1	0.38	-1.262	1	0.22	-3.446
9	PAS	STOXX	1	0.32	6.351***	1	0.2	-4.382
10	MIL	STOXX	1	0.24	-6.369	1	0.2	-4.477
11	SON	STOXX	1	0.3	-3.854	1	0.3	4.448***
12	ROM	STOXX	1	0.48	1.571*	1	0.26	-5.833
13	BBV	STOXX	1	0.38	-1.081	1	0.2	-6.361
14	POR	STOXX	1	0.26	-6.351	1	0.3	-5.378
15	ESS	STOXX	1	0.26	-6.357	1	0.2	-5.44
16	POP	STOXX	1	0.32	-1.551	1	0.26	5.359***
17	ESP	STOXX	1	0.26	6.351***	1	0.2	-6.364
18	SAN	STOXX	1	0.2	6.351***	1	0.42	0.673
19	BNP	STOXX	1	0.34	-1.947	1	0.2	-3.121
20	CNP	STOXX	1	0.3	0.78	1	0.2	-6.367
21	COF	STOXX	1	0.2	-6.351	1	0.2	-6.365
22	COM	STOXX	1	0.26	6.351***	1	0.2	-6.36
23	COR	STOXX	1	0.58	0.411	1	0.22	-6.375
24	DEU	STOXX	1	0.24	6.352***	1	0.24	-6.371
25	ERS	STOXX	1	0.38	1.576*	1	0.26	3.295***
26	FON	STOXX	1	0.26	-6.357	1	0.32	0.052
27	GEC	STOXX	1	0.24	-6.362	1	0.32	-0.389
28	GBL	STOXX	1	0.3	-6.359	1	0.42	0.088
29	SOC	STOXX	1	0.24	6.351***	1	0.22	-6.364
30	HAN	STOXX	1	0.32	-0.645	1	0.2	-3.911
31	ICA	STOXX	1	0.24	-1.474	1	0.34	3.846***
32	IMM	STOXX	2	0.6	-0.316	2	0.6	2.684***
33	ING	STOXX	1	0.26	6.351***	1	0.22	-3.734
34	INT	STOXX	1	0.28	-2.919	1	0.28	4.994***
35	KBC	STOXX	1	0.26	6.352***	1	0.2	-6.359
36	KLE	STOXX	1	0.24	6.351***	1	0.52	-0.315
37	MAP	STOXX	1	0.26	-0.444	1	0.22	-6.371
38	MED	STOXX	1	0.26	6.351***	1	0.2	-6.361
39	MUE	STOXX	1	0.28	-6.351	1	0.2	-6.357
40	NAT	STOXX	1	0.26	1.417*	1	0.28	-4.937
41	POH	STOXX	1	0.28	6.352***	1	0.24	-3.925
42	SAM	STOXX	1	0.6	-0.026	1	0.28	-4.666
43	SCO	STOXX	1	0.32	-2.487	1	0.2	-6.365
44	ROD	STOXX	1	0.52	0.239	1	0.24	-6.377
45	UNI	STOXX	1	0.28	6.352***	1	0.3	5.886***
46	VIE	STOXX	1	0.32	3.747***	1	0.2	-6.362
47	WEN	STOXX	1	0.24	4.733***	1	0.22	-6.364
48	WER	STOXX	1	0.3	3.34***	1	0.44	1.572*

Table 4.C.3: NCoVaR from institution i on the system risk in two scenarios in period 01/01/2000 till 12/31/2012 for VAR-filtered returns. Lags determines the optimal number of lags from the VAR specification using the Schwarz-Bayes Information Criterion. Optimal epsilon values calculated from bootstrap. T-val represents the test statistic of NCoVaR from Eq.(4.11). (*),(**), (***) denotes one-sided p-value statistical significance at 10%, 5% and 1%, respectively.

			Scenario 1			Scenario 2		
	Institution i	System variable	Lags	Opt. ε	T-val	Lags	Opt. ε	T-val
1	ACK	STOXX	1	0.26	6.351***	1	0.44	1.417*
2	AEG	STOXX	1	0.24	6.355***	1	0.2	-6.36
3	AGE	STOXX	1	0.38	0.983	1	0.24	5.506***
4	ALL	STOXX	1	0.36	0.247	1	0.2	-6.363
5	ASS	STOXX	1	0.3	-2.589	1	0.22	-6.361
6	AXA	STOXX	1	0.32	6.353***	1	0.2	-6.358
7	BIR	STOXX	1	0.22	-6.362	1	0.32	4.329***
8	BAN	STOXX	1	0.28	-2.625	1	0.2	-6.36
9	PAS	STOXX	1	0.34	6.35***	1	0.3	5.469***
10	MIL	STOXX	1	0.34	-5.756	1	0.2	-3.407
11	SON	STOXX	1	0.36	4.458***	1	0.2	-5.44
12	ROM	STOXX	1	0.34	2.491***	1	0.22	-5.627
13	BBV	STOXX	1	0.2	6.35***	1	0.2	-6.36
14	POR	STOXX	1	0.32	-6.36	1	0.24	-6.336
15	ESS	STOXX	1	0.36	1.998***	1	0.24	5.202***
16	POP	STOXX	1	0.34	-3.085	1	0.3	4.667***
17	ESP	STOXX	1	0.36	2.102**	1	0.26	-6.375
18	SAN	STOXX	1	0.2	6.351***	1	0.44	-0.077
19	BNP	STOXX	1	0.22	6.35***	1	0.2	-6.356
20	CNP	STOXX	1	0.34	-0.525	1	0.2	-6.374
21	COF	STOXX	1	0.34	-2.344	1	0.2	-6.363
22	COM	STOXX	1	0.4	3.693***	1	0.2	-6.354
23	COR	STOXX	1	0.6	0.627	1	0.24	-6.38
24	DEU	STOXX	1	0.32	6.355***	1	0.22	-6.366
25	ERS	STOXX	1	0.36	0.493	1	0.34	1.903**
26	FON	STOXX	1	0.3	-6.355	1	0.3	0.633
27	GEC	STOXX	1	0.34	4.518***	1	0.46	-0.53
28	GBL	STOXX	1	0.26	-2.699	1	0.38	3.204***
29	SOC	STOXX	1	0.2	6.351***	1	0.2	-6.36
30	HAN	STOXX	1	0.32	-6.35	1	0.2	-2.984
31	ICA	STOXX	1	0.32	0.502	1	0.32	4.231***
32	IMM	STOXX	1	0.58	0.962	1	0.32	4.231***
33	ING	STOXX	1	0.26	6.35***	1	0.24	-6.363
34	INT	STOXX	1	0.34	-4.345	1	0.28	3.23***
35	KBC	STOXX	1	0.24	6.351***	1	0.2	-6.358
36	KLE	STOXX	1	0.36	2.593***	1	0.36	3.947***
37	MAP	STOXX	1	0.48	1.519*	1	0.2	-6.364
38	MED	STOXX	1	0.2	6.35***	1	0.2	-6.364
39	MUE	STOXX	1	0.24	-6.35	1	0.2	-6.354
40	NAT	STOXX	1	0.28	1.093	1	0.3	-4.97
41	POH	STOXX	1	0.32	6.353***	1	0.44	0.314
42	SAM	STOXX	1	0.52	-0.857	1	0.2	-6.362
43	SCO	STOXX	1	0.32	0.759	1	0.2	-6.364
44	ROD	STOXX	1	0.5	-0.401	1	0.24	-6.38
45	UNI	STOXX	1	0.3	6.351***	1	0.28	5.673***
46	VIE	STOXX	1	0.34	0.611	1	0.3	2.273***
47	WEN	STOXX	1	0.26	6.368***	1	0.2	-6.359
48	WER	STOXX	1	0.32	-2.424	1	0.22	-6.37

Table 4.C.3: NCoVaR from institution i on the system risk in two scenarios in period 01/01/2000 till 12/31/2012 for GARCH-filtered returns. Lags determines the number of lags used in the test. Optimal epsilon values calculated from bootstrap. T-val represents the test statistic of NCoVaR from Eq.(4.11). (*),(**), (***) denotes one-sided p-value statistical significance at 10%, 5% and 1%, respectively.

			Scenario 1			Scenario 2		
	Institution i	System variable	Lags	Opt. ε	T-val	Lags	Opt. ε	T-val
1	ACK	STOXX	1	0.54	0.913	1	0.44	1.329*
2	AEG	STOXX	1	0.28	6.359***	1	0.4	2.855***
3	AGE	STOXX	1	0.25	6.362***	1	0.23	-4.646
4	ALL	STOXX	1	0.38	1.45*	1	0.21	-6.366
5	ASS	STOXX	1	0.23	6.351***	1	0.21	-6.359
6	AXA	STOXX	1	0.26	6.351***	1	0.2	-3.035
7	BIR	STOXX	1	0.25	-4.733	1	0.21	5.319***
8	BAN	STOXX	1	0.38	-1.262	1	0.23	-3.61
9	PAS	STOXX	1	0.32	6.351***	1	0.2	-4.382
10	MIL	STOXX	1	0.24	-6.369	1	0.21	-4.477
11	SON	STOXX	1	0.3	-3.854	1	0.31	4.083***
12	ROM	STOXX	1	0.27	6.351***	1	0.27	-5.883
13	BBV	STOXX	1	0.25	6.352***	1	0.21	-6.362
14	POR	STOXX	1	0.26	-6.351	1	0.3	-5.378
15	ESS	STOXX	1	0.26	-6.357	1	0.21	-5.771
16	POP	STOXX	1	0.29	-4.344	1	0.27	4.958***
17	ESP	STOXX	1	0.27	6.351***	1	0.21	-6.364
18	SAN	STOXX	1	0.21	6.351***	1	0.43	0.418
19	BNP	STOXX	1	0.34	-1.947	1	0.2	-3.121
20	CNP	STOXX	1	0.31	0.531	1	0.21	-6.37
21	COF	STOXX	1	0.2	-6.351	1	0.21	-6.365
22	COM	STOXX	1	0.26	6.351***	1	0.2	-6.36
23	COR	STOXX	1	0.25	2.84***	1	0.21	-6.371
24	DEU	STOXX	1	0.23	6.352***	1	0.23	-6.37
25	ERS	STOXX	1	0.38	1.576*	1	0.27	3.141***
26	FON	STOXX	1	0.26	-6.357	1	0.32	0.052
27	GEC	STOXX	1	0.24	-6.362	1	0.33	-0.46
28	GBL	STOXX	1	0.3	-6.359	1	0.42	0.088
29	SOC	STOXX	1	0.24	6.351***	1	0.22	-6.364
30	HAN	STOXX	1	0.27	-3.068	1	0.2	-3.911
31	ICA	STOXX	1	0.25	-3.335	1	0.34	3.846***
32	IMM	STOXX	1	0.22	-6.361	1	0.35	4.183***
33	ING	STOXX	1	0.26	6.351***	1	0.22	-3.734
34	INT	STOXX	1	0.26	6.352***	1	0.2	-6.359
35	KBC	STOXX	1	0.24	6.351***	1	0.52	-0.315
36	KLE	STOXX	1	0.26	-0.444	1	0.22	-6.371
37	MAP	STOXX	1	0.25	1.432*	1	0.28	4.994***
38	MED	STOXX	1	0.27	6.351***	1	0.2	-6.361
39	MUE	STOXX	1	0.29	-6.351	1	0.2	-6.357
40	NAT	STOXX	1	0.27	0.34	1	0.28	-4.937
41	POH	STOXX	1	0.29	6.352***	1	0.24	-3.925
42	SAM	STOXX	1	0.59	-0.147	1	0.2	-6.362
43	SCO	STOXX	1	0.32	-2.487	1	0.2	-6.365
44	ROD	STOXX	1	0.52	0.239	1	0.23	-6.375
45	UNI	STOXX	1	0.28	6.352***	1	0.3	5.886***
46	VIE	STOXX	1	0.32	3.747***	1	0.2	-6.362
47	WEN	STOXX	1	0.25	4.869***	1	0.22	-6.364
48	WER	STOXX	1	0.25	6.352***	1	0.45	1.411*