SFB 823

Spatial dependence in stock returns - Local normalization and VaR forecasts

Discussion Paper

Thilo A. Schmitt, Rudi Schäfer, Dominik Wied, Thomas Guhr

Nr. 18/2013

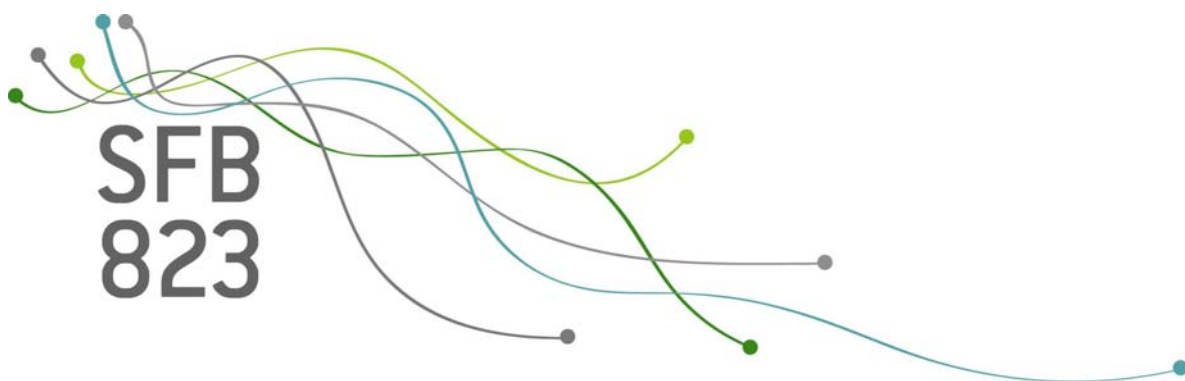SFB 823

Spatial dependence in stock returns - Local normalization and VaR forecasts

Thilo A. Schmitt[1], Rudi Schäfer[1], Dominik Wied[2], and Thomas Guhr[1]

[1]Fakultät für Physik, Universität Duisburg-Essen, 47048 Duisburg, Germany
[2]Fakultät Statistik, TU Dortmund, 44221 Dortmund, Germany

May 2, 2013

We analyze a recently proposed spatial autoregressive model for stock returns and compare it to a one-factor model and the sample covariance matrix. The influence of refinements to these covariance estimation methods is studied. We employ power mapping as a noise reduction technique for the correlations. Further, we address the empirically observed non-stationary behavior of stock returns. Local normalization strips the time series of changing trends and fluctuating volatilities. As an alternative method, we consider a GARCH fit. In the context of portfolio optimization, we find that the spatial model has the best match between the estimated and realized risk measures. (JEL: G17, C33)

1 Introduction

The covariance matrix of stock returns plays a crucial role for portfolio optimization, see Markowitz (1952). Finding the portfolio weights that yield minimum risk given a desired portfolio return requires the best possible estimation of the future covariance matrix. A natural approach is to estimate the sample covariance matrix using historical data, assuming that the recent past is a good predictor for the future. The length of the historical time series which can be used for covariance estimation is often rather short. In particular in emerging markets, the total length of the available time series may be the limiting factor. Another consideration is the non-stationarity of the financial markets, see, e.g., Longin and Solnik (1995), Bekaert and Harvey (1995) and Münnix et al. (2012): The correlation structure changes with time. Hence, to achieve a decent estimate of the

current or future covariance matrix, we should only take into account rather recent data. However, since the sample covariance matrix for n assets requires $n(n+1)/2$ parameters to be estimated, the finiteness of the time series leads to a considerable amount of measurement noise, see Laloux et al. (1999), Bouchaud and Potters (2009), Plerou et al. (1999) and Plerou et al. (2002). As pointed out by, *e.g.*, Pafka and Kondor (2002, 2003), this has dire consequences for portfolio optimization, but can be mitigated to a large extent by noise reduction techniques. Here we will concentrate on one such technique, the power mapping, which has been introduced in Guhr and Kälber (2003) and further studied in Schäfer et al. (2010).

An alternative approach to reducing the noise in sample covariance matrices is to consider a model for the correlation or covariance matrix which entails fewer parameters. Many models have been proposed to reduce the number of parameters which have to be estimated, see Pantaleo et al. (2011). Here we consider a simple one-factor model, see Sharpe (1963), where $2n+1$ parameters have to be estimated. In addition, we study the spatial autoregressive model for stock returns, which was recently introduced by Arnold et al. (2013). One particular feature of the spatial model is its ability to produce reliable Value-at-Risk (VaR) forecasts. This is partly due to the fact that the model captures a lot of dependence with a small number of parameters. It involves only $n+3$ parameters, 3 for the dependence and n parameters describing the individual volatilities.

A specific issue we want to address in this paper concerns the empirically observed non-stationarity of financial time series. We are going to study the influence of sudden changes in local trends and volatilities on the covariance estimation methods described above. In particular, we investigate how the estimation of the spatial parameters is affected. We suggest the following refinements to substantially improve covariance estimation methods: The well known GARCH(1,1) model (see Bollerslev (1986) and Bollerslev et al. (1988)) can be utilized to remove the fluctuating volatilities in the return time series and to predict future volatilities. This approach is compared to a local normalization method, recently introduced by Schäfer and Guhr (2010), and a short-term historical prediction of the individual volatilities. Moreover, we apply the above mentioned power mapping to reduce the noise in the correlation matrices. We study the influence of these refinements on each of the covariance estimation methods and the implications for portfolio optimization. To this end we consider the stocks of the Euro Stoxx 50 and use the realized portfolio variances as a risk measure. We compare the results to the predicted portfolio variances. We analyze the VaR forecast quality in more detail by comparing it to predicted and realized variances. Comparing different models with respect to their VaR forecast ability is a quite common approach in the literature, see e.g. Santos et al. (2013). While it were basically possible to include still other models like a multivariate DCC model, see Engle (2002), in the study or to account for possible structural breaks in the model parameters (in the spirit of Wied (2013)), we have decided to focus on the present models in order to keep the presentation clear. In fact, the current analysis is in our opinion sufficient for the main results: It is extremely important for covariance estimation to take into account non-stationarity of financial time series and the measurement noise. And while using GARCH residuals and volatility forecasts yields comparable results, the combination of local normalization and short term historical

volatilities requires much shorter time series.

The paper is structured as follows: Section 2 presents the above mentioned methods for estimating the covariance matrix of stock returns. In Section 3 we discuss the refinements to these methods, which aim at removing changes in local trends and volatilities, improving the quality of volatility predictions, and reducing estimation noise. Section 4 details the portfolio optimization technique and the data set under consideration. In Section 5 we discuss the results of the covariance estimation and VaR forecasts, and we summarize our findings in Section 6.

2 Covariance estimation

The covariance matrix is a crucial input parameter for many risk assessing methods in finance, such as portfolio optimization. Due to the non-stationarity of financial markets the estimation of the covariance matrix is a non-trivial task. The natural way is to calculate the sample covariance matrix from the time series. As usual, we use the definition $\widehat{\text{Cov}}(y_t) = \frac{1}{T-1} \sum_{t=1}^{T} (y_t - \bar{y})(y_t - \bar{y})'$, where $\bar{y}$ is the mean of the time series. The sample covariance matrix requires the estimation of $n(n+1)/2$ parameters. If the covariance matrix is calculated on a short time horizon it contains a great degree of noise. For larger time horizons the predictive power of the covariance matrix decreases as the market constantly changes. In Section 2.1 and 2.2 we discuss two models that require less parameters to be estimated to determine the covariance matrix.

2.1 One-factor model

The simple one-factor model radically reduces the amount of parameters to estimate by assuming that the change of all assets is tied to one factor, e.g. a market index. It was first used to improve portfolio optimization by Sharpe (1963) and decreases the amount of parameters to estimate to $2n + 1$. A comprehensive description of the model is given by Jorion (2007), p. 192 ff. The one-factor model assumes that the stock returns can be described by

$$y_t = \alpha + \beta y_{m,t} + \eta_t \quad , \tag{1}$$

where the n-dimensional vector y_t contains the returns for all stocks $1 \ldots n$ at time t. The scalar $y_{m,t}$ describes the market return, e.g. is calculated from a stock market index. The vector $\beta = (\beta_1, \ldots, \beta_n)$ contains a constant for each stock which must be estimated, for example with a linear fit (ordinary least squares-estimator) separately for all stocks. The fixed intercepts' vector α can be neglected in the context of risk estimation as it contains no randomness. Assuming that the error terms in η_t are uncorrelated to each other the covariance matrix of y_t is given by

$$\text{Cov}(y_t) = \beta\beta'\sigma_m^2 + D_\eta \tag{2}$$

with a matrix $D_\eta = \text{diag}(\sigma_{\eta 1}^2, \ldots, \sigma_{\eta n}^2)$ that contains the variances of η_t on its diagonal and which can be estimated by standard ordinary least squares methods. The one-factor

model requires an estimation of n entries for the β vector and the n diagonal elements of D_η, plus one for the market volatility σ_m.

Given the parameter estimates, we can directly obtain a parametric estimate for $\text{Cov}(y_t)$.

2.2 Spatial dependence model

The spatial dependence model introduced by Arnold et al. (2013) is based on the assumption that a lot of the cross-sectional dependence between the stock returns can be captured by three different types of dependence: A general dependence, dependence within industrial branches and dependence based on geographic locations. For an overview of spatial dependence modeling see Anselin (1988), Cressie (1993) and LeSage and Pace (2009). Formally, we have the spatial autoregressive model

$$y_t = \rho_g W_g y_t + \rho_b W_b y_t + \rho_l W_l y_t + \varepsilon_t \quad , \quad t = 1, \ldots, T \tag{3}$$

where ρ_g is a scalar parameter measuring the general dependence, ρ_b is a scalar parameter measuring the dependence between industrial branches and ρ_l is a scalar parameter measuring the dependence based on geographic locations. W_g, W_b and W_l are spatial weighting matrices. The stochastic component in this model stems from the error vector ε_t. Given the basic model assumption it is plausible to assume that its covariance matrix has uncorrelated entries although heteroscedasticity is allowed.

The non-diagonal elements of the matrix W_g are set to the normalized market capitalization of the corresponding assets. For the matrices W_b and W_l the element in the i-th row and j-th column is non-zero if the i-th and j-th asset are in the same branch (W_b) or country (W_l). The non-zero elements are set to the normalized market capitalization of the asset in each row.

If the three parameters are known, the covariance matrix of the vector y_t is given by

$$\text{Cov}(y_t) = (I_n - \rho_g W_g - \rho_b W_b - \rho_l W_l)^{-1} \Sigma (I_n - \rho_g W_g' - \rho_b W_b' - \rho_l W_l')^{-1} \quad , \tag{4}$$

where Σ is the covariance matrix of the error term ε_t and I_n is an $n \times n$ identity matrix. The error terms are assumed to be uncorrelated, so all off-diagonal elements of Σ are zero. This leads to n additional parameters. The model uses $n + 3$ parameters which are best estimated by a two-step procedure that is based on the generalized methods of moments (GMM) approach, see Arnold et al. (2013) and in addition Lee and Liu (2009) and Lin and Lee (2010). Again, given the parameter estimates, we can directly obtain a parametric estimate for $\text{Cov}(y_t)$.

3 Refined methods of covariance estimation

We discuss four approaches to enhance the predictive capabilities of the methods discussed in Section 2. The GARCH residuals (Section 3.1) and local normalization (Section 3.2) reduce the empirically observed non-stationarity of return time series with regard

to the volatility. The power mapping method discussed in Section 3.3 is aimed at decreasing the noise in a correlation matrix. In Section 3.4 we explore additional methods to estimate the volatilities of the individual stocks.

3.1 GARCH residuals

The return time series has a fluctuating volatility which can lead to estimation errors in parameters derived from the time series. To improve the estimation it is desirable to remove these fluctuations from the return time series. This is possible by modeling the returns with a GARCH process introduced by Bollerslev (1986) as a generalization of the ARCH process invented by Engle (1982). We fit the GARCH(1,1)-model

$$X_t = \sigma_t \varepsilon_t \tag{5}$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{6}$$

to the historic data to estimate the parameters α_0, α_1 and β_1. Here, $(\varepsilon_t)_{t\in\mathbb{Z}}$ is a strong white noise process with $\text{var}\,\varepsilon_t = 1$ and $E[\varepsilon_t] = 0$. The conditional variances σ_t^2 can replicate the fluctuating volatilities in empirical time series. Hansen and Lunde (2005) have shown that in most cases a GARCH(1,1) is sufficient to capture the return time series. Then we use the GARCH residuals

$$\varepsilon_t = \frac{X_t}{\sigma_t} \tag{7}$$

to receive a return time series, where the volatility fluctuations are removed to the degree the return time series fits the GARCH process.

For the comparison in Section 5, we use a rolling window of $T = 100$ trading days to estimate model parameters. This window is too small for the GARCH fit to converge. Therefore we use a rolling window of $T_{\text{GARCH}} = 1000$ trading days to estimate the GARCH parameters. We emphasize that this larger window is only used to estimate the GARCH parameters.

3.2 Addressing local trends and changes in volatilty: local normalization

Estimating the GARCH parameters requires a rather large time window. Therefore we use a second method called local normalization introduced by Schäfer and Guhr (2010). It removes local trends and changes in volatility without altering the cross-correlations between time series. The local average of a function is defined as

$$\langle f_t \rangle_m = \frac{1}{m} \sum_{j=0}^{m-1} f_{t-j\Delta t} \quad , \tag{8}$$

where Δt is the return interval. Then the locally normalized returns are given by

$$\rho_{mt} = \frac{r_t - \langle r_t \rangle_m}{\sqrt{\langle r_t^2 \rangle_m - \langle r_t \rangle_m^2}} \quad , \tag{9}$$

where we first subtract the local mean value $\langle r_t \rangle_m$ from the return r_t and then divide by the local volatility. As shown by Schäfer and Guhr (2010) a value of $m = 13$ yields optimal results for daily stock returns.

3.3 Noise-reduction: power mapping

The correlation matrix of financial assets contains a significant amount of noise, which can be seen by comparing the eigenvalue density of a correlation matrix to a random matrix, see Laloux et al. (1999). The part of small eigenvalues, called the bulk part, exhibits the same shape for both matrices. They only differ for larger eigenvalues, which can be associated to industrial branches. The natural method to reduce the noise would be to increase the length of the time series to calculate the correlation matrix from. This is not a feasible way to predict the future correlation matrix, because the relationships between companies constantly change, as they start competing on new markets or discontinue their activities in one field. Several methods have been proposed in the past to reduce the noise, while keeping the times series short, e.g. see Gopikrishnan et al. (2001) and Giada and Marsili (2001).

Here we discuss the power mapping method introduced by Guhr and Kälber (2003) to reduce the noise in a correlation matrix. Every entry of the correlation matrix C is substituted by

$$C_{ij}^{(q)} = \text{sign}(C_{ij}) \, |C_{ij}|^q \tag{10}$$

yielding the noise reduced correlation matrix $C^{(q)}$. Notice that the diagonal elements are equal to one and thus not affected by power mapping. In general, the optimal value for the parameter q depends on the time horizon T on which the correlation matrix is calculated, $i.e.$, the degree of noise in the correlation matrix. However, power mapping is a very robust method which yields good results for a wide range of q values around the optimal one, as discussed in Schäfer et al. (2010). Here, we use $q = 1.5$.

3.4 Volatility forecast

The correlation matrix needed for the power mapping method can be calculated from the covariance matrix by dividing each element of the covariance matrix

$$C_{ij} = \frac{\text{Cov}(y_t)_{ij}}{\sigma_i \sigma_j} \tag{11}$$

by the respective volatilities σ_i and σ_j. In case of the sample covariance matrix we use the standard deviations of the returns calculated on the rolling window of $T = 100$ trading days. There is a plethora of possible methods or models to forecast volatility, see Poon and Granger (2003) for a review. Here, we use standard deviations from historic time series and the predicted volatilities from a GARCH fit. For the one-factor and the spatial dependence model we use the model specific volatilities from the diagonal of the covariance matrix in Equation (2) and (4), respectively. Then we can apply power

mapping to the correlation matrix. For the portfolio optimization in Section 4.2 we use the covariance matrix

$$\text{Cov}(y_t)_{ij} = C_{ij}\hat{\sigma}_i\hat{\sigma}_j \quad . \tag{12}$$

At this point we could use the same volatilities as in Equation (11) $\hat{\sigma}_{i,j} = \sigma_{i,j}$ or we can use other methods to estimate the volatility. We use two additional ways to calculate the volatilities. First, we calculate the standard deviation from the unaltered returns, i.e. with no further methods applied to them, in a rolling window of $T_{\text{vol}} = 14$ trading days. Here we assume that the volatility in the past three weeks is a better indicator for the future standard deviation compared to the longer horizon of $T = 100$ trading days.

Second, we can use the parameters from the GARCH fit described in Section 3.1 to predict the volatilities for the next trading day according to Equation (6).

4 Application to portfolio optimization

4.1 The data set

We use the adjusted daily closing prices for a collection of 49 stocks contained in the Euro Stoxx 50. It includes companies from various countries in the eurozone and spans across different branches. We had to remove GDF Suez because of incomplete data due to the merger. The data is taken from Thomson Reuters Datastream. A complete list of the stocks including their industrial branch and country as used in the spatial dependence model is given in Table 1. Nokia and CRH from Finland and Ireland are put together in the country group "others" because groups are not allowed to contain only one entry to avoid singularities. The observation period ranges from January 2001 to May 2012. We calculate the logarithmic returns from the adjusted prices. Table 1 gives an overview of the used stocks.

4.2 Portfolio optimization

We compare the effects of the methods discussed in Section 3 on the covariance estimation techniques of Section 2. For each covariance matrix, we perform a portfolio optimization to determine the minimum variance portfolio, see Markowitz (1952) and also Markowitz (1959) and Elton et al. (2006).

We estimate each covariance matrix $\text{Cov}(y_t) =: V$ on a rolling window of 100 trading days. The covariance matrix yields the portfolio weights

$$\frac{\hat{V}^{-1}\tau}{\tau'\hat{V}^{-1}\tau} \tag{13}$$

for the minimum variance portfolio, where τ is a vector containing only ones. The predicted portfolio variance is then

$$\hat{\sigma}^2_{\text{port}} := \left(\tau'\hat{V}^{-1}\tau\right)^{-1} \quad . \tag{14}$$

7

Table 1: The data set	
Automobile	BMW (Germany), Daimler (Germany), VW (Germany)
Basic industry	Arcelor Mittal (Benelux), CRH (Ireland), Saint-Gobain (France), Vinci (France)
Consumer electronics	Nokia (Finland), Philips (Benelux), SAP (Germany), Schneider (France), Siemens (Germany)
Consumer Retail	Anheuser Busch (Benelux), Carrefour (France), Danone (France), Inditex (Spain), L'Oreal (France), LVMH (France), Unilever (Benelux)
Energy	E.ON (Germany), ENEL (Italy), ENI (Italy), Iberdrola (Spain), RWE (Germany), Repsol (Spain), Total (France)
Finance	AXA (France), Allianz (Germany), BNP (France), Banco Bilbao (Spain), Banco Santander (Spain), Deutsche Bank (Germany), Deutsche Börse (Germany), Generali (Italy), ING (Benelux), Intesa (Italy), Münchener Rück (Germany), Société Générale (France), Unicredit (Italy), Unibail-rodamco (France)
Pharma and chemicals	Air Liquide (France), BASF (Germany), Bayer (Germany), Sanofi (France)
Telecom and media	Deutsche Telekom (Germany), France Telecom (France), Telecom Italia (Italy), Telefonica (Spain), Vivendi (France)

The estimated portfolio variance is then used to calculate a Gaussian value-at-risk (VaR) for a given α-quantile u_α

$$\widehat{\text{VaR}}_\alpha = u_\alpha \sqrt{\hat{\sigma}^2_{\text{port}}} \,. \tag{15}$$

In this setup it is possible to calculate the VaR on a daily basis and compare it to the realized portfolio returns.

5 Results

5.1 Spatial parameters

Figure 1 shows the influence of local normalization on the parameter estimation for the spatial dependence model. The three parameters of the spatial dependence model are calculated for a rolling time window of $T = 250$ days. The dashed lines show the parameters calculated from the original returns. The solid lines present the three parameters with local normalization applied to the returns. Especially during the financial crisis

of 2008 a strong jump is noticeable in the general and branch spatial parameters. This coincides with the peaking volatility during this turbulent time. The artifact which has the same width as the rolling window vanishes when applying the local normalization. We note that the use of GARCH residuals yields similar results.
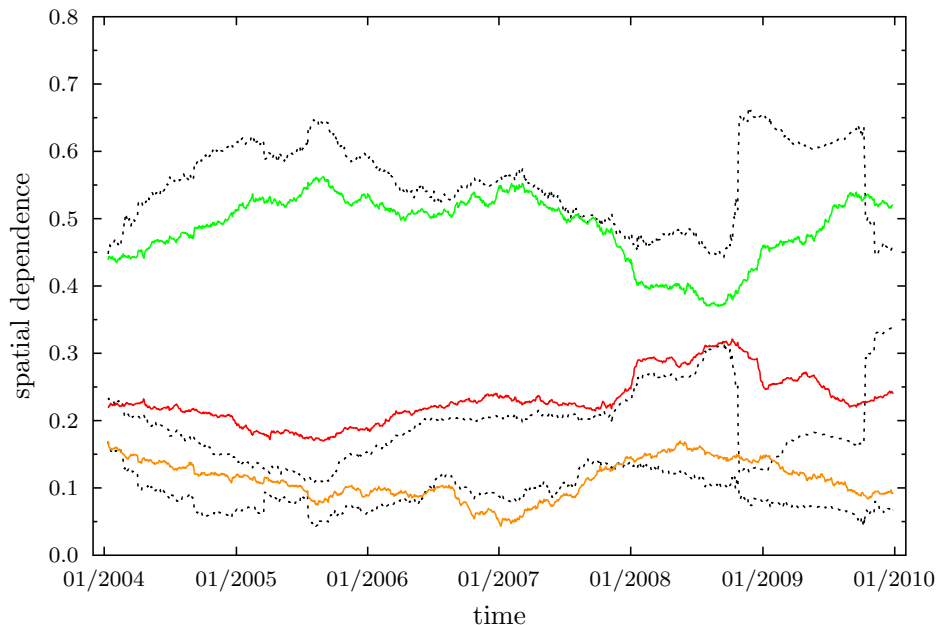

Figure 1: The parameters ρ_g, ρ_b and ρ_l are shown from top to bottom estimated at each trading day for an interval of 250 days. The solid lines are with local normalization applied to the returns, while the dotted lines are estimated from the unaltered returns.

5.2 Portfolio variances

First, we discuss the impact of each method presented in Section 3 on the realized portfolio variances. Then, we examine how combinations of the methods effect the risk assessment. The realized portfolio variances are given in Table 2. The table is structured as follows: The second column states which returns were used. We can use the original returns, the GARCH residuals or the locally normalized returns. The volatility forecast method is specified in the third column. The fourth column indicates whether or not power mapping was used to suppress the estimation noise. The last three columns show the results of the realized portfolio variances for the spatial dependence model, the one-factor model and the sample covariance matrix.

Using the GARCH residuals (row 2) or local normalization (row 3) yields similar effects for all models. The spatial dependence model benefits from both methods in a similar way, while the portfolio variances for the one-factor model get worse. There is

Table 2: Realized portfolio variances

	returns	volatility forecast	power mapping	realized portfolio variances		
				sdep	1-factor	sample
1	original	hist	no	0.000241	0.000093	0.000121
2	GARCH	hist	no	0.000171	0.000171	0.000115
3	normalized	hist	no	0.000172	0.000148	0.000127
4	original	hist	yes	0.000256	<u>0.000087</u>	<u>0.000086</u>
5	GARCH	hist	yes	0.000133	0.000207	<u>0.000086</u>
6	normalized	hist	yes	0.000135	0.000190	0.000093
7	original	GARCH	no	0.000100	<u>0.000087</u>	0.000126
8	GARCH	GARCH	no	<u>0.000087</u>	0.000142	0.000119
9	normalized	GARCH	no	<u>0.000087</u>	0.000121	0.000132
10	original	GARCH	yes	0.000097	<u>0.000086</u>	0.000098
11	GARCH	GARCH	yes	<u>0.000084</u>	0.000160	0.000095
12	normalized	GARCH	yes	<u>0.000084</u>	0.000143	0.000101
13	original	hist†	no	0.000101	<u>0.000086</u>	0.000116
14	GARCH	hist†	no	0.000091	0.000138	0.000119
15	normalized	hist†	no	<u>0.000088</u>	0.000119	0.000120
16	original	hist†	yes	0.000099	<u>0.000084</u>	0.000092
17	GARCH	hist†	yes	<u>0.000085</u>	0.000156	0.000093
18	normalized	hist†	yes	<u>0.000086</u>	0.000139	0.000097

no significant effect on the sample covariance matrix. Power mapping (row 4) slightly increases the portfolio variance in case of the spatial dependence model.

The one-factor model improves a little bit, but is very good from start. The strength of power mapping unveils when applied to the sample covariance matrix. Here power mapping greatly reduces the realized portfolio variance. The GARCH predicted volatilities (row 7) enormously decrease the portfolio variance for the spatial dependence model. They have a minor positive effect on the one-factor model, while a minuscule negative effect on the sample covariance matrix.

By combining different methods it is possible to further reduce the portfolio variance. The GARCH predicted volatilities with either GARCH residuals or local normalization (row 8,12 and 9,11) improve the spatial dependence model to be on par with the sample covariance matrix. The one-factor model does not benefit and the sample covariance matrix only gets better if power mapping is used (row 11,12). The effect of power mapping is marginal in case of the spatial dependence model.

Again, we notice that the one-factor model works best if used together with the original returns (row 10), while the spatial dependence model benefits from the GARCH volatility forecast. Local normalization or the GARCH residuals improve the spatial dependence model (row 5,6), but not to the same extent as a better volatility forecast (row 7-12 and 13-18).

If we shorten the time interval on which the volatility is estimated we can observe

comparable results to the GARCH predicted volatilities (row 13-18 and 7-12). In particular, local normalization in combination with the shorter historical volatilities achieve matching results compared to GARCH residuals with GARCH predicted volatilities and requires far shorter time series (row 12 and 18).

With regard to the realized portfolio variances we can conclude that the spatial dependence model works best in combination with methods that improve the volatility estimation, like GARCH predicted volatilities or local normalization with a shorter horizon for the volatility calculation. The one-factor works best without any refinements with the exception of minor improvements in combination with power mapping. For the sample covariance matrix only power mapping is required to achieve the best performance.

Table 3: Relative predicted portfolio variances in percent

	returns	volatility forecast	power mapping	$\frac{realized-predicted}{predicted}$ in % sdep	1-factor	sample
1	original	hist	no	$\underline{57}$	203	426
2	GARCH	hist	no	64	$\underline{45}$	362
3	normalized	hist	no	73	$\underline{53}$	397
4	original	hist	yes	$\underline{80}$	195	180
5	GARCH	hist	yes	$\underline{32}$	206	174
6	normalized	hist	yes	$\underline{39}$	234	184
7	original	GARCH	no	$\underline{137}$	256	581
8	GARCH	GARCH	no	$\underline{141}$	847	515
9	normalized	GARCH	no	$\underline{144}$	452	564
10	original	GARCH	yes	$\underline{108}$	260	283
11	GARCH	GARCH	yes	$\underline{119}$	1330	267
12	normalized	GARCH	yes	$\underline{121}$	755	276
13	original	hist†	no	323	401	903
14	GARCH	hist†	no	297	953	877
15	normalized	hist†	no	$\underline{233}$	480	727
16	original	hist†	yes	$\underline{186}$	319	353
17	normalized	hist†	yes	$\underline{187}$	765	359
18	GARCH	hist†	yes	$\underline{183}$	1345	350

We compare the predicted variances with the realized portfolio variances shown in Table 3. The prediction error of the sample covariance matrix is greatly reduced by the use of power mapping. Local normalization yields a small enhancement for the spatial dependence model, while there is no improvement for the one-factor model and the sample covariance matrix. In both cases, the one-factor model and the sample covariance matrix are not very well suited to predict the realized variance. The spatial dependence model is the best predictor for the realized portfolio variances.

5.3 VaR forecast

We calculate the VaR forecast according to Equation (15) on a daily basis for each α. The probability that the realized portfolio return is smaller than the VaR forecast is shown in Figure 2 for $\alpha \in (0, 0.5]$. The probability is calculated from all trading days in the observation period. For a perfect model the probability $P(y_{\text{port}} < \widehat{\text{VaR}}_\alpha)$ should be equal to α, which is indicated by a straight line.
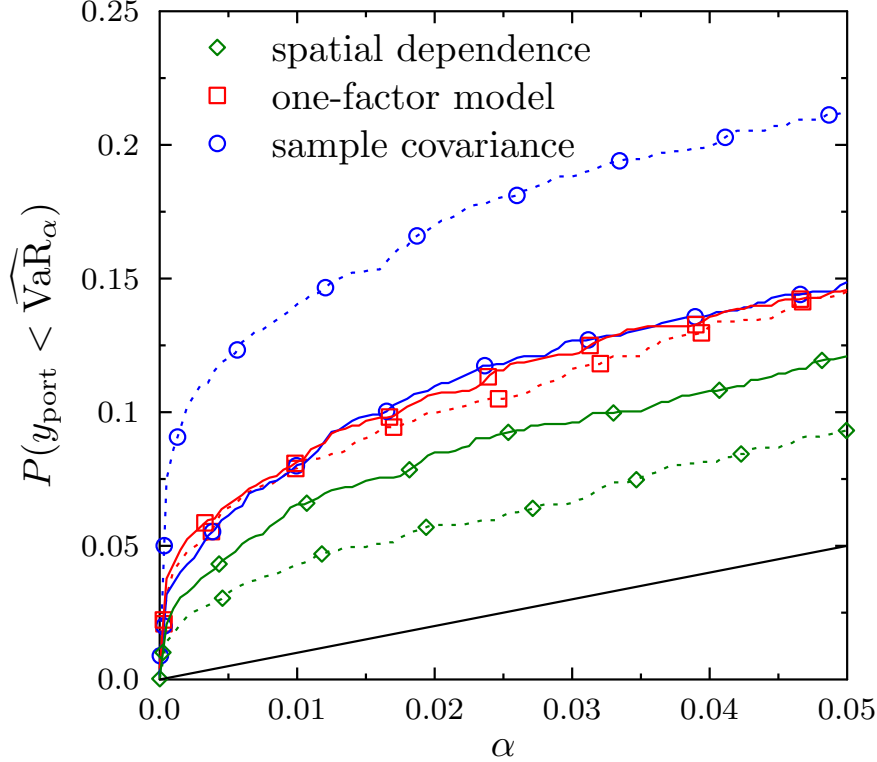

Figure 2: The probability of a portfolio return being smaller than the value at risk given a fixed quantile of α. The dashed lines show the results without any refinements for the spatial dependence model (diamond), the one-factor model (square) and the sample covariance matrix (circle). The solid lines show the effect of improved covariance estimation methods (see text for details).

We do not show the results for each case presented in Table 2; instead we limit ourselves to one case per model, where the realized portfolio variances are lowest and the prediction error is smallest for each model. For the spatial dependence model (diamond) we present the VaR forecast with GARCH residuals, GARCH-predicted volatilities and applied power mapping. The one-factor model (squares) uses a combination of original returns, historic volatilities and power mapping, while the sample covariance matrix (circle) uses GARCH residuals instead, historic volatilities and power mapping. The

dashed lines show the original models without any refinements while the solid lines show the refined cases.

Without any refinements the spatial dependence model produces the best VaR forecasts in the observed period, while the results for the one-factor model and the sample covariance matrix are rather poor in comparison. The large improvement with regard to the realized portfolio variances leads to a poorer VaR forecast. In contrast the sample covariance matrix not only gains better realized variances by applying power mapping and using GARCH residuals but also leads to a better VaR forecast. For the one-factor model there is no significant change to the predictive power of the VaR forecast.

Nonetheless, the sample covariance matrix and the one-factor model do not surpass the spatial dependence model in their risk estimation.

6 Conclusion

We compare three approaches for covariance estimation: The spatial dependence model, a one-factor model and the sample covariance matrix. As a benchmark for the quality of the covariance estimation, we use portfolio optimization. The realized portfolio variances and the relative prediction error are scrutinized. In addition to the original approaches, we investigate several refinement methods. An estimation error can arise from fluctuating volatilities; they can be removed from the return time series either by employing a GARCH fit and using the residuals, or by using a local normalization method. Volatilities of the individual return time series can be better predicted using a short-term historical estimate. This is due to the slowly decaying autocorrelation of empirical volatilities. Alternatively, we can use the volatility predictions of the GARCH fits. There is a large statistical estimation error, if the length of the time series is not much larger than the parameters to be estimated. This measurement noise can be reduced by noise reduction techniques such as power mapping.

The spatial dependence model captures the correlation between assets very well, a noise reduction of the covariance matrix is not necessary. However, it falls short on estimating the future volatilities of individual stocks. The realized portfolio variances can be immensely reduced by combining the spatial dependence model with better methods for volatility forecasting, and using locally normalized returns for the regression. The one-factor model produces quite good realized portfolio variances on its own. It works best with the original returns and should not be used with GARCH residuals or locally normalized returns. Slight improvements are possible with better volatility forecast methods or noise reduction. The sample covariance matrix suffers from noise due to the finite length of the time series. Noise reduction methods such as power mapping are sufficient to achieve results that are equally good compared to the other approaches.

Local normalization and the GARCH residuals effectively remove fluctuations in the volatility and reduce estimation artifacts for the spatial parameters. With the right choice of refinements all three approaches are capable of producing equally good realized portfolio variances, though the spatial dependence model provides the smallest prediction error for portfolio variances and VaR forecasts.

Funding

References

L. Anselin. *Spatial Econometrics: Methods and Models.* Studies in Operational Regional Science. Springer, 1988.

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

G. Bekaert and C. Harvey. Time-varying world market integration. *The Journal of Finance*, L(2):403–44, 1995.

T. Bollerslev. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

T. Bollerslev, R. Engle, and J. Wooldridge. A Capital Asset Pricing Model with Time-Varying Covariances. *The Journal of Political Economy*, 96(1):116–131, 1988.

J.-P. Bouchaud and M. Potters. *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management.* Cambridge University Press, 2 edition, 2009.

N. Cressie. *Statistics for spatial data.* Wiley series in probability and mathematical statistics: Applied probability and statistics. J. Wiley, 1993.

E. J. Elton, M. J. Gruber, S. J. Brown, and W. N. Goetzmann. *Modern Portfolio Theory and Investment Analysis.* Wiley, 7 edition, 2006.

R. Engle. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica: Journal of the Econometric Society*, 50 (4):987–1007, 1982.

R. Engle. Dynamic Conditional Correlation. *J. Bus. Econ. Stat.*, 20(3):339–350, July 2002.

L. Giada and M. Marsili. Data clustering and noise undressing of correlation matrices. *Physical Review E*, 63(6):061101, May 2001.

P. Gopikrishnan, B. Rosenow, V. Plerou, and H. Stanley. Quantifying and interpreting collective behavior in financial markets. *Physical Review E*, 64(3):035106, Aug. 2001.

T. Guhr and B. Kälber. A new method to estimate the noise in financial correlation matrices. *Journal of Physics A: Mathematical and General*, 36(12):3009–3032, Mar. 2003.

P. R. Hansen and A. Lunde. A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20(7):873–889, Dec. 2005.

P. Jorion. *Value at risk: the new benchmark for managing financial risk.* McGraw-Hill, New York, 3 edition, 2007.

L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters. Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*, 83(7):1467–1470, Aug. 1999.

L.-F. Lee and X. Liu. Efficient GMM Estimation of High Order Spatial Autoregressive Models With Autoregressive Disturbances. *Econometric Theory*, 26(01):187, Aug. 2009.

J. LeSage and R. Pace. *Introduction to Spatial Econometrics.* Statistics: a Series of Textbooks and Monographs. CRC PressINC, 2009.

X. Lin and L.-f. Lee. GMM estimation of spatial autoregressive models with unknown heteroskedasticity. *Journal of Econometrics*, 157(1):34–52, July 2010.

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

H. Markowitz. Portfolio Selection. *The Journal of Finance*, 7(1):77–91, 1952.

H. Markowitz. *Portfolio Selection: Efficient Diversification of Investment.* Yale University Press, 1959.

M. C. Münnix, T. Shimada, R. Schäfer, F. Leyvraz, T. H. Seligman, T. Guhr, and H. E. Stanley. Identifying states of a financial market. *Sci. Rep.*, 2, 09 2012.

S. Pafka and I. Kondor. Noisy covariance matrices and portfolio optimization. *Eur. Phys. J. B*, 27:277–280, 2002.

S. Pafka and I. Kondor. Noisy covariance matrices and portfolio optimization II. *Physica A*, 319:487–494, 2003.

E. Pantaleo, M. Tumminello, F. Lillo, and R. N. Mantegna. When do improved covariance matrix estimators enhance portfolio optimization? An empirical comparative study of nine estimators. *Quantitative Finance*, 11(7):1067–1080, 2011.

V. Plerou, P. Gopikrishnan, B. Rosenow, L. Amaral, and H. Stanley. Universal and nonuniversal properties of cross correlations in financial time series. *Phys. Rev. Lett.*, 83:1471–1474, 1999.

V. Plerou, P. Gopikrishnan, B. Rosenow, L. Amaral, T. Guhr, and H. Stanley. Random matrix approach to cross correlations in financial data. *Phys. Rev. E*, 65:066126, 2002.

S. Poon and C. Granger. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, XLI(June):478–539, 2003.

A. Santos, F. Nogales, and E. Ruiz. Comparing Univariate and Multivariate Models to Forecast Portfolio Value-at-Risk. *Journal of Financial Econometrics*, 11(2):400–441, 2013.

R. Schäfer and T. Guhr. Local normalization: Uncovering correlations in non-stationary financial time series. *Physica A*, 389(18):3856–3865, Sept. 2010.

R. Schäfer, N. F. Nilsson, and T. Guhr. Power mapping with dynamical adjustment for improved portfolio optimization. *Quant. Finance*, 10(1):107–119, Jan. 2010.

R. Schäfer, N. F. Nilsson, and T. Guhr. Power mapping with dynamical adjustment for improved portfolio optimization. *Quantitative Finance*, 10(1):107–119, 2010.

W. Sharpe. A simplified model for portfolio analysis. *Management science*, 9(2):277–293, 1963.

D. Wied. Cusum-type testing for changing parameters in a spatial autoregressive model for stock returns. *Journal of Time Series Analysis*, 34(1):221–229, 2013.