On the robustness of the principal volatility components

Carlos Trucíos[1,2]*, Luiz K. Hotta[3] and Pedro Valls[1,2].

[1]Sao Paulo School of Economics - FGV. Brazil.

[2]Centre of Quantitative Studies in Economics and Finance. Brazil.

[3]Department of Statistics, University of Campinas. Brazil

Abstract

In this paper, we analyse the recently procedure of Hu and Tsay (2014) (Principal volatility component analysis. JBES, v32.2) and Li et al. (2016) (Modeling multivariate volatilities via latent common factors. JBES, v34.4) called principal volatility components. This procedure overcomes several difficulties in modelling and forecasting the conditional covariance matrix in large dimensions arising from the curse of dimensionality. We show that outliers have a devastating effect on the construction of the principal volatility components and on the forecast of the conditional covariance matrix and consequently in economic and financial applications based on this forecast. We propose a robust procedure and analyse its finite sample properties by means of Monte Carlo experiments and also illustrate it using empirical data. The robust procedure outperforms the classical method in simulated and empirical data.

Keywords: Conditional covariance matrix; Constant volatility; Curse of dimensionality; Jumps; Outliers; Principal components.

JEL Classification: C13; C51; C53; C55; G17.

AMS Classification: 62H99; 62G35; 62M20; 91B84; 91G70.

*Corresponding author: ctrucios@gmail.com

The authors acknowledge financial support from São Paulo Research Foundation (FAPESP) grants 2016/18599-4, 2013/00506-1 and 2013/22930-0 respectively. The first two authors also acknowledges support from Laboratory EPIFISMA while the third author is also grateful with the National Council for Scientific and Technological Development (CNPq) grant 309158/2016-8. Weiming Li is also gratefully acknowledged for their Matlab codes.

1 Introduction

Modelling and forecasting volatilities and co-volatilities play a crucial role in many economic and financial applications such as portfolio allocation, risk measures, option pricing, securities regulations and hedging strategies (Chiou and Tsay, 2008; Hammoudeh et al., 2010; Rombouts and Stentoft, 2011; Basher and Sadorsky, 2016; Wang and Liu, 2016).

Given the unfeasibility and inflexibility of most classical multivariate volatility models in large dimensions, researchers and practitioners have been looking for alternative tools to circumvent the curse of dimensionality when modelling and forecasting (co)volatilities in high-dimensional data. In this sense, some alternative approaches have been suggested in the last years. See, for instance, Lopes et al. (2012), Fan et al. (2012), Hafner and Reznikova (2012), Pakel et al. (2014), Gruber and West (2016), Kastner (2016), Li et al. (2016) and Engle et al. (2017) among others. Furthermore, based on the idea that comovements in the market can be driven by a few components, factor models appear in the economic and financial literature as an alternative way to achieve dimension reduction and to tackle the curse of dimensionality. See, for instance Fan et al. (2008), Pan et al. (2010), Matteson and Tsay (2011), García-Ferrer et al. (2012), Santos and Moura (2014), Matilainen et al. (2015) and Barigozzi and Hallin (2015) for some references.

In the spirit of dimensionality reduction, an innovative approach based on the classical principal component analysis (PCA), called principal volatility components (PVC), has been recently proposed by Hu and Tsay (2014a) and Li et al. (2016). This methodology produces two types of components. The first type corresponds to components with conditional covariance matrix evolving over time whilst the other type corresponds to components with constant conditional covariance matrix. This methodology is attractive because after obtaining the volatility components, the problem of modelling and forecasting the (co)volatilities of the entire system drop down into modelling and forecasting the (co)volatilities of the volatility components with heteroscedastic dynamic since the remaining components have constant volatility.

On the other hand, it is well know that outliers are not unusual in financial time series and several works show how outliers affect dramatically the forecast of (co)volatilities (Muler and Yohai, 2008; Boudt and Croux, 2010; Carnero et al., 2012; Boudt et al., 2013; Grané et al., 2014; Trucíos and Hotta, 2016; Trucíos et al., 2017b) and consequently financial applications (Vaz de Melo Mendes and Pereira Câmara Leal, 2005; Welsch and Zhou, 2007; Trucíos et al., 2017a). Furthermore, there are evidence showing that PCA is very sensitive to the presence of outliers (Croux and Haesbroeck, 2000; Hubert et al., 2005; Candès et al., 2011; Greco and Farcomeni, 2016). Thus, procedures based on similar methodology are also expected to be sensitive to outliers.

In this paper, we first analyse by means of Monte Carlo experiments the performance of the PVC in the presence of additive outliers showing that outliers have a devastating effect on this procedure even when moderate outliers are present. Then, we propose a robust principal volatility component (RPVC) procedure which shows to have good finite sample properties.

The rest of the paper is organized as follows. Section 2 presents the PVC of Hu and Tsay (2014a), the generalized version of Li et al. (2016) and our robust procedure. In Section 3 an extensive Monte Carlo experiment is carried out to evaluate the finite sample properties of the procedures in contaminated and uncontaminated series. Section 4 presents an empirical application of daily returns with 73 stocks of the Nasdaq-100 index and show that our robust procedure has better performance when applied to the selection of the minimum variance portfolio. Finally, Section 5 presents the main conclusions and future works.

2 Volatility components

Let $y_t = (y_{1t}, ..., y_{Nt})'$ a N-dimensional vector with $E(y_t|\mathcal{F}_{t-1}) = 0$ where $\mathcal{F}_{t-1}$ denotes the information available till time $t-1$ and let $M_{N\times N} = [A_{N\times r} \quad B_{N\times(N-r)}]$ an orthogonal matrix. Observe that if we denote $f_t = A'y_t$ and $\epsilon_t = BB'y_t$, we can rewrite y_t as

$$y_t = MM'y_t = (AA' + BB')y_t = Af_t + \epsilon_t. \tag{1}$$

Hu and Tsay (2014a) and Li et al. (2016) introduce methodologies on which, under mild conditions, it is possible to find B such that $Var(\epsilon_t|\mathcal{F}_{t-1}) = Var(\epsilon_t)$, i.e, the second term ϵ_t contains homocedastic components and all the conditional heteroscedastic components come from the first term. Although model (1) has the same form of the classical factor model, there are some differences between them. First, model (1) splits y_t into two terms, one explaining the conditional heteroscedastic dynamic (f_t) and the other one driven by components with constant volatility (ϵ_t). Additionally, f_t and ϵ_t are not necessary uncorrelated. Finally, none assumption is imposed directly on f_t and ϵ_t and all the features described previously are consequences of the eigenvalue-eigenvector decomposition described in Hu and Tsay (2014a) and Li et al. (2016) respectively. This model is also particularly useful because it reduces considerable the number of parameter to be estimated circumventing the curse of the dimensionality.

We briefly introduce the approaches of Hu and Tsay (2014a) and Li et al. (2016), denoted by PVC and GPVC respectively. These approaches allow to obtain components with the features described previously. Additionally, knowing the bad influence of outliers in classical methodologies and inspired on the comments of Franke (2014) and Hu and Tsay (2014b) about the robustness of the PVC procedure, we introduce a robust procedure which is less sensitive to additive outliers.

2.1 Principal volatility components (PVC)

Let us assume that the vector y_t defined previously is weakly stationary with finite four-order moment. Hu and Tsay (2014a) consider the eigenvalue-eigenvector decomposition of the cumulative generalized kurtosis matrix given by $\Gamma_\infty M = \Lambda M$ where Λ is a decreasing ordered diagonal matrix of eigenvalues, M is the associated normalized eigenvectors and Γ_∞ is the cumulative generalized kurtosis matrix defined as

$$\Gamma_\infty = \sum_{k=1}^{\infty} \sum_{i=1}^{N} \sum_{j=1}^{N} E^2 \left[\left(y_t y_t' - \Sigma \right) \left(x_{ij,t-k} - E\left(X_{ij} \right) \right) \right], \tag{2}$$

where Σ is the unconditional covariance matrix and $x_{ij,t-k}$ is a function of $y_{i,t-k} y_{j,t-k}$[1]. The kth volatility component is defined as $z_{kt} = m_k' y_t$ where m_k is the eigenvector associated with the kth eigenvalue and corresponds to the kth column of M. Hu and Tsay (2014a)

[1] In their simulations and empirical application Hu and Tsay (2014a) use the Huber's function defined as

$$r(x) = \begin{cases} x, & \text{if } |x| \leq c^2, \\ 2c\sqrt{x} - c^2, & \text{if } x > c^2, \\ c^2 - 2c\sqrt{|x|}, & \text{if } x < -c^2. \end{cases}$$

proves that if m_k is an eigenvector associated with a zero eigenvalue of Γ_∞, the linear combination $m_k' y_t$ has constant volatility (See Lemma 1 — Theorem 1 of Hu and Tsay (2014a)). Additionally, it can also be proved that under mild conditions (Theorem 1 of Hu and Tsay (2014a)) exist $N - r$ linearly independent combinations of y_t with constant volatility, where $r = rank(\Gamma_\infty)$.

In practice, (2) is estimated by

$$\hat{\Gamma}_r = \sum_{k=1}^{g} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(1 - \frac{k}{T}\right)^2 \left[\frac{1}{T} \sum_{t=k+1}^{T} \left[\left(y_t y_t' - \hat{\Sigma}\right)(x_{ij,t-k} - \bar{x}_{ij})\right]\right]^2, \tag{3}$$

where $\hat{\Sigma}$ is the sample covariance matrix, $\bar{x}_{ij}$ is the sample mean of $x_{ij,t}$, g is a positive integer that represents a lag order and T is the sample size. For more details see, Hu and Tsay (2014a) and Andreou and Ghysels (2014).

2.2 Generalized principal volatility components (GPVC)

The PVC of Hu and Tsay (2014a) assumes that the vector series has finite fourth-order moment. However, there is evidence showing that in many financial series this assumption does not hold (Zhu and Ling, 2011). To relax this assumption, Li et al. (2016), inspired by the paper of Pan et al. (2010), propose an alternative PVC procedure, denoted by GPVC, which requires only finite second-order moments.

In the GPVC, the cumulative generalized kurtosis matrix (2) is replaced by

$$G = \sum_{k=1}^{g} \sum_{t=1}^{T} \omega(y_t) E^2 \left[(y_t y_t' - \Sigma)\, I(\|y_{t-k}\| \leq \|y_t\|)\right], \tag{4}$$

where $\omega(\cdot)$ is a weight function and $\|\cdot\|$ is the L_1 norm. The matrix G is estimated in a natural way by

$$\hat{G} = \sum_{k=1}^{g} \sum_{\tau=1}^{T} \omega(y_\tau) \left[\frac{1}{T-k} \sum_{t=k+1}^{T} \left[\left(y_t y_t' - \hat{\Sigma}\right) I(\|y_{t-k}\| \leq \|y_\tau\|)\right]\right]^2. \tag{5}$$

Both procedures present a good performance with a slight better performance in favour of the GPVC procedure (Li et al., 2016). However, these procedures have two drawbacks. The first one, which is not discussed here, is related with the problem of dealing with $N/T \to 1$

or even $N > T$. The second one, which is the focus of this paper, is related with the presence of additive outliers that, as discussed previously, can have several implications in modelling and forecasting volatility (Boudt et al., 2013; Grané et al., 2014; Trucíos et al., 2017a,b). These outliers are not unusual and can be related with financial crashes, elections, wars, macroeconomic news and terrorist attack (Charles and Darné, 2014; Laurent et al., 2016).

Because both procedures are based on a methodology similar to the classical PCA, which is very sensitive to atypical observations (Croux and Haesbroeck, 2000; Hubert et al., 2005; Candès et al., 2011; Greco and Farcomeni, 2016) and in addition taking into account that both procedures focus on estimation and prediction of the conditional covariance matrix, which are badly affected by additive outliers (Carnero et al., 2012; Boudt et al., 2013; Trucíos et al., 2017b,a) it is important to know whether and how outliers affect the (G)PVC procedures and consequently their financial applications. In a second step, it is interesting to find an alternative or a robust procedure, which is pursued in the following section.

2.3 Robust principal volatility components (RPVC)

In order to obtain a procedure less sensitive to additive outliers, we robustify the estimator given in (5). The robust procedure is based on a robust estimator of the unconditional covariance matrix and a weighted estimator of $E\left[(y_t y_t' - \Sigma)\, I(\|y_{t-k}\| \leq \|y_t\|)\right]$. We replace the matrix (5) by a less sensitive matrix defined as

$$\hat{G}^R = \sum_{k=1}^{g} \sum_{\tau=1}^{T} \omega(y_\tau) \left[\sum_{t=k+1}^{T} \delta^*(d_t^2) \left\{ (y_t y_t' - \hat{\Sigma}^R) I(\|y_{t-k}\| \leq \|y_\tau\|) \right\} \right]^2, \qquad (6)$$

where d_t^2 is the robust square Mahalanobis distance given by $d_t^2 = (y_t - \hat{\mu}^R)' \hat{\Sigma}_t^{-1} (y_t - \hat{\mu}^R)$ with $\hat{\Sigma}_t = 0.01\rho(y_{t-1}' y_{t-1}) + 0.99\hat{\Sigma}_{t-1}$, $\hat{\Sigma}_1 = \hat{\Sigma}^R$ and $\hat{\mu}^R$ and $\hat{\Sigma}^R$ being a robust estimates of the unconditional mean and covariance matrix obtained using the minimum covariance determinant (MCD) estimator of Rousseeuw (1984) implemented with the algorithm of Hubert et al. (2012). Finally, $\rho(\cdot)$ and $\delta(\cdot)$ are given by

$$\rho(x_t) = \begin{cases} x_t, & \text{if } d_t^2 \leq c, \\ \hat{\Sigma}^R, & \text{if } d_t^2 > c, \end{cases} \qquad \delta(x) = \begin{cases} 1, & \text{if } x \leq c, \\ \frac{1}{x}, & \text{if } x > c, \end{cases}$$

and $\delta^*(\cdot) = \delta(\cdot)/||\delta(\cdot)||$, where $\|\cdot\|$ is the L_1 norm.

Observe that, to avoid that returns corresponding to periods with high volatility being considered as possible outliers we incorporate in the squared Mahalanobis distance a covariance matrix evolving over time which is obtained using a RiskMetrics 1994 Smoother with $\lambda = 0.99$. Similar approaches have been also used in Croux et al. (2010), Boudt and Croux (2010) and Boudt et al. (2013). Additionally, because the sample covariance matrix is sensitive to outliers (Hubert et al., 2012, 2015), we use the robust MCD estimator (Rousseeuw, 1984; Hubert et al., 2012). To maintain the robustness of d_t^2, we use $\hat{\Sigma}^R$ as the initial value in $\hat{\Sigma}_t$ and introduce the filter $\rho(\cdot)$ that mitigate the effect of outliers in the RiskMetrics Smoother. Finally, as a natural robust estimator of $E\left[(y_t y_t' - \Sigma)\, I(\|y_{t-k}\| \leq \|y_t\|)\right]$ we use a weighted estimator that penalize large values of d_t^2.[2]

3 Monte Carlo experiments

To evaluate the finite sample properties of the PVC, GPVC and RPVC, we carry out Monte Carlo experiments with small and large dimensions. Series of sample size 1000 and 1000 Monte Carlo replicates are considered. Different patterns of contamination, size of outliers and percentage of series contaminated are considered. We consider consecutive (C) and isolated (I) outliers in the middle and close to the end of the sample period. In cases contaminated by isolate outliers we put two outliers in the series at positions $t = 500$ and 999, in the same way cases with consecutive outliers are contaminated at positions $t = 500, 501$ and $998, 999$. Outliers of size 5 and 10 standard deviations of the univariate uncontaminated process are contemplated. Finally, we consider uncontaminated series (0% of series contaminated) and contamination of $p\%$ of the series, with $p\% = 25\%, 50\%$ and 100% of the series. The $p\%$ contaminated series were the first $p\%$ series which appear in the entire system.

[2] At the same time we are working in a robust PVC procedure, another robust procedure is being developed independently by Monte and Reisen (2016). The main differences between both approaches are that Monte and Reisen (2016) robustify the procedure of Hu and Tsay (2014a) while we robustify the procedure of Li et al. (2016). The procedure of Monte and Reisen (2016) replace the generalized covariance in (3) by a robust version based on Ma and Genton (2000) while we use the robust MCD estimator of Rousseeuw (1984). Additionally we mitigate the effect of outliers penalizing large values of the squared Mahalanobis distance taking into account high and low volatility period using a RiskMetrics Smoother that avoid returns corresponding to periods with high volatility being considered as outliers. Finally, the robust procedure proposed in this paper is fast and feasible in large dimensions. Because the GPVC has shown a slight better performance than the PVC and in addition because the robust procedure of Monte and Reisen (2016) is computationally more expensive than the other approaches we do not analyse this procedure here since Monte Carlo experiments even for small dimension ($N = 8$) is highly time consuming and in consequence infeasible in moderate/large dimensions.

Following Hu and Tsay (2014a), Andreou and Ghysels (2014) and Li et al. (2016), we use the factor model as DGP. In the simulation study we consider three analyses. First, we analyse if outliers affect the estimation of the number of volatility components with heteroskedastic dynamic. Second, considering that the number of components with heteroskedastic dynamic are known, we analyse the effect of the outliers in the estimation of the matrix A in (1). Finally, we are interested on the effects of outliers in the prediction of the conditional covariance matrix and its implications in economic and financial applications.

3.1 Number of volatility components

First, we are interested in knowing whether the selection of the number of volatility components with heteroskedastic dynamic is affected by outliers. The procedures proposed in the literature to estimate the number of components are not conclusive, for simplicity and illustrative purposes we use the ratio estimator criterion (Lam and Yao, 2012; Ahn and Horenstein, 2013), the BN criterion (Bai and Ng, 2002) and the Kaiser-Guttman criterion (Guttman, 1954).

We consider small and large dimensions ($N = 8$ and 60) and the factor model was generated with two and six common factors for $N = 8$ and 60 respectively. Each common factor follows a Gaussian GARCH(1,1) process with parameters $\omega = (1, 2)$, $\alpha = (0.07, 0.03)$ and $\beta = (0.83, 0.92)$ for $N = 8$ and $\omega = (1, 2, 1, 0.5, 2, 3)$, $\alpha = (0.07, 0.03, 0.05, 0.03, 0.02, 0.03)$ and $\beta = (0.83, 0.92, 0.90, 0.95, 0.78, 0.87)$ for $N = 60$, the factor load matrix A was randomly drawn as a matrix with orthogonal columns using the R package *pracma* of Borchers (2017), the idiosyncratic factors were simulated as $\epsilon_t = \frac{\bar{\epsilon}_t}{N}$ where $\bar{\epsilon}_t \sim Normal_N(0, I)$. The initial covariance matrix H_0 was simulated as a positive definite matrix using the procedure of Joe (2006) implemented in R package *clusterGeneration* of Qiu and Joe. (2015). In all cases we simulate 1500 observations and discard the first 500 to avoid the influence of the initial values.

Tables 1 and 2 report the average and standard deviation of the estimated number of volatility components using the ratio estimator (top panel), the BN (middle panel) and the Kaiser-Guttman (bottom panel) criteria for small and large dimensions respectively.

Additionally, the proportion of estimated components smaller, equal and larger than the true number of factors are also reported. Observe that, mainly for the large dimension case, when the non-robust procedures are implemented, the selected number of estimated volatility components obtained using any of the three criteria mentioned previously is highly affected by the presence of outliers. However, when the robust procedure is implemented, the presence of outliers does not affect significantly the selected number of volatility components. In large dimension, the ratio estimator and the BN criteria tends to overestimate the number of components when outliers are present in the series and the non-robust procedures are used, two exceptions are observed using the ratio estimator criterion in the PVC procedure with 100% of the series contaminated by consecutive outliers, in which cases the number of selected components is mostly underestimated when the size of outliers is $\omega = 10$ and almost 41.9% of cases is underestimated when $\omega = 5$.

In the presence of outliers, as expected, the Kaisser-Guttman criterion also missestimate the number of selected components. Furthermore, in small dimension the Kaisser-Guttman criterion also missestimate the number of selected components when no outliers are in the series and approximately 50% of cases underestimate the number of selected volatility components. However, in large dimensions when the series are uncontaminated, the number of estimated components obtaining is close to the true number of factors. Additional Monte Carlo experiments conclude that as the ratio *common factors/dimension* increase, the Kaiser-Guttman criterion estimate incorrectly the number of components (see supplementary material).

Note that, when the RPVC procedure is used, the ratio estimator criteria estimates the number of components pretty well in all cases. Furthermore, in uncontaminated cases and large dimension, all criteria estimate correctly the number of components.

3.2 Eigenvectors associated with non-zero eigenvalues

In this section, we analyse the effects of outliers on the estimation of the eigenvectors associated with the non-zero eigenvalues. Note that, these vectors are the columns of the matrix A in the model $y_t = Af_t + \epsilon_t$ and play an important role on the forecast of the conditional covariance matrix. To separate the source of error, we focus on the estimation of the

Table 1: Average and standard deviation of the estimated number of components for uncontaminated and contaminated series using the ratio estimator (top panel), the BN (middle panel) and the Kaiser-Guttman (bottom panel) criteria. For data coming from a factor model the proportion of estimated components lower (Comp = 1), equal (Comp =2) and upper (Comp >2) the number of true factors is also reported. Dimension $N = 8$, sample size $T = 1000$ and 1000 Monte Carlo replicates.

		0%	25%				50%				100%			
			$\omega = 5$		$\omega = 10$		$\omega = 5$		$\omega = 10$		$\omega = 5$		$\omega = 10$	
			I	C	I	C	I	C	I	C	I	C	I	C
Ratio estimator criterion														
	Mean	1.992	1.986	1.988	1.987	1.990	1.988	1.989	1.988	1.987	1.987	1.988	1.987	1.988
	SD	0.089	0.118	0.109	0.113	0.100	0.109	0.104	0.109	0.113	0.113	0.109	0.113	0.109
RPVC	Comp =1	0.008	0.014	0.012	0.013	0.010	0.012	0.011	0.012	0.013	0.013	0.012	0.013	0.012
	Comp =2	0.992	0.986	0.988	0.987	0.990	0.988	0.989	0.988	0.987	0.987	0.988	0.987	0.988
	Comp >2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Mean	1.998	2.383	2.607	3.106	3.366	2.868	3.100	3.249	3.798	2.281	2.624	2.032	3.108
	SD	0.045	0.830	1.017	0.737	0.864	1.031	1.465	0.952	1.385	0.957	1.370	0.830	1.176
GPVC	Comp =1	0.002	0.151	0.173	0.050	0.068	0.152	0.233	0.100	0.126	0.222	0.245	0.265	0.106
	Comp =2	0.998	0.390	0.271	0.074	0.053	0.150	0.115	0.059	0.060	0.409	0.302	0.502	0.193
	Comp >2	0.000	0.459	0.556	0.876	0.879	0.698	0.652	0.841	0.814	0.369	0.453	0.233	0.701
	Mean	1.972	2.214	2.694	3.037	2.779	2.517	2.944	3.286	2.437	2.090	1.880	2.143	1.987
	SD	0.165	0.794	1.012	0.785	1.268	1.065	1.523	0.908	1.626	0.878	0.975	0.960	0.886
PVC	Comp =1	0.028	0.189	0.168	0.069	0.291	0.242	0.290	0.085	0.434	0.258	0.409	0.284	0.303
	Comp =2	0.972	0.450	0.207	0.083	0.065	0.199	0.110	0.055	0.216	0.480	0.407	0.404	0.497
	Comp >2	0.000	0.361	0.625	0.848	0.644	0.559	0.600	0.860	0.350	0.262	0.184	0.312	0.200
BN criterion														
	Mean	2.066	2.400	2.502	2.588	2.620	2.449	2.517	2.598	2.599	2.328	2.416	2.532	2.579
	SD	0.293	0.726	0.783	0.818	0.833	0.743	0.799	0.814	0.825	0.650	0.732	0.791	0.815
RPVC	Comp =1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Comp =2	0.946	0.743	0.679	0.625	0.609	0.703	0.677	0.613	0.621	0.773	0.730	0.656	0.631
	Comp >2	0.054	0.257	0.321	0.375	0.391	0.297	0.323	0.387	0.379	0.227	0.270	0.344	0.369
	Mean	2.013	3.352	3.650	3.422	3.704	3.658	3.899	3.691	3.872	3.765	3.888	3.856	3.823
	SD	0.137	0.531	0.498	0.498	0.457	0.477	0.308	0.462	0.346	0.447	0.334	0.354	0.460
GPVC	Comp =1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001
	Comp =2	0.990	0.027	0.010	0.002	0.000	0.001	0.002	0.000	0.004	0.010	0.006	0.001	0.030
	Comp >2	0.010	0.973	0.990	0.998	1.000	0.999	0.998	1.000	0.996	0.990	0.994	0.999	0.969
	Mean	2.026	3.348	3.651	3.438	3.701	3.649	3.901	3.690	3.891	3.696	3.924	3.845	3.888
	SD	0.208	0.551	0.497	0.500	0.458	0.480	0.299	0.463	0.324	0.508	0.283	0.376	0.337
PVC	Comp =1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
	Comp =2	0.983	0.038	0.010	0.002	0.000	0.001	0.000	0.000	0.001	0.023	0.005	0.005	0.007
	Comp >2	0.017	0.962	0.990	0.998	1.000	0.999	1.000	1.000	0.998	0.977	0.995	0.995	0.993
Kaiser-Guttman criterion														
	Mean	1.562	1.549	1.522	1.552	1.521	1.550	1.520	1.548	1.518	1.544	1.517	1.546	1.518
	SD	0.496	0.498	0.500	0.498	0.500	0.498	0.500	0.498	0.500	0.498	0.500	0.498	0.500
RPVC	Comp =1	0.438	0.451	0.478	0.448	0.479	0.450	0.480	0.452	0.482	0.456	0.483	0.454	0.482
	Comp =2	0.562	0.549	0.522	0.552	0.521	0.550	0.520	0.548	0.518	0.544	0.517	0.546	0.518
	Comp >2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Mean	1.494	1.527	1.556	1.741	1.658	1.581	1.568	1.661	1.601	1.588	1.599	1.549	1.679
	SD	0.500	0.500	0.503	0.533	0.543	0.494	0.500	0.578	0.551	0.499	0.502	0.498	0.490
GPVC	Comp =1	0.506	0.473	0.447	0.305	0.377	0.419	0.434	0.394	0.431	0.415	0.407	0.451	0.332
	Comp =2	0.494	0.527	0.550	0.649	0.588	0.581	0.564	0.551	0.537	0.582	0.587	0.549	0.657
	Comp >2	0.000	0.000	0.003	0.046	0.035	0.000	0.002	0.055	0.032	0.003	0.006	0.000	0.011
	Mean	1.433	1.467	1.508	1.674	1.422	1.496	1.478	1.634	1.324	1.528	1.382	1.465	1.391
	SD	0.496	0.503	0.512	0.512	0.548	0.502	0.516	0.541	0.475	0.503	0.488	0.501	0.488
PVC	Comp =1	0.567	0.535	0.498	0.347	0.606	0.505	0.530	0.396	0.679	0.474	0.619	0.536	0.609
	Comp =2	0.433	0.463	0.496	0.632	0.366	0.494	0.462	0.574	0.318	0.524	0.380	0.463	0.391
	Comp >2	0.000	0.002	0.006	0.021	0.028	0.001	0.008	0.030	0.003	0.002	0.001	0.001	0.000

(a) 25%, 50% and 100% of series contaminated at time $t = 500$ and 999 (isolated outliers) or $t = 500, 501$ and $998, 999$ (consecutive outliers). (b) (c) Size of outliers $\omega = 5$ and 10 standard deviations of the univariate uncontaminated process. (d) the factor models are simulated with two factors.

Table 2: Average and standard deviation of the estimated number of components for uncontaminated and contaminated series using the ratio estimator (top panel), the BN (middle panel) and the Kaiser-Guttman (bottom panel) criteria. For data coming from a factor model the proportion of estimated components lower (Comp < 6), equal (Comp =6) and upper (Comp >6) the number of true factors is also reported. Dimension $N = 60$, sample size $T = 1000$ and 1000 Monte Carlo replicates.

		0%	25%				50%				100%			
			$\omega = 5$		$\omega = 10$		$\omega = 5$		$\omega = 10$		$\omega = 5$		$\omega = 10$	
			I	C	I	C	I	C	I	C	I	C	I	C
Ratio estimator criterion														
	Mean	6.000	6.000	6.000	6.000	6.000	6.000	6.000	6.000	6.000	6.000	6.000	6.000	6.000
	SD	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RPVC	Comp <6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Comp =6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Comp >6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Mean	6.000	7.980	9.926	7.990	9.960	7.999	9.997	7.999	9.997	8.000	10.000	7.766	9.604
	SD	0.000	0.140	0.298	0.100	0.196	0.032	0.055	0.032	0.055	0.000	0.000	1.162	1.490
GPVC	Comp <6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.039	0.066
	Comp =6	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Comp >6	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.961	0.934
	Mean	6.000	7.976	9.941	7.990	9.927	7.999	9.997	7.999	8.886	8.000	6.647	7.832	2.050
	SD	0.000	0.153	0.252	0.100	0.597	0.032	0.055	0.032	2.769	0.000	3.950	0.990	0.621
PVC	Comp <6	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.139	0.000	0.419	0.028	0.994
	Comp =6	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Comp >6	0.000	1.000	1.000	1.000	0.996	1.000	1.000	1.000	0.861	1.000	0.581	0.972	0.006
BN criterion														
	Mean	6.001	6.005	6.005	6.023	6.008	6.005	6.004	6.017	6.010	6.008	6.002	6.013	6.003
	SD	0.032	0.071	0.071	0.157	0.089	0.071	0.063	0.144	0.100	0.089	0.045	0.122	0.055
RPVC	Comp <6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Comp =6	0.999	0.995	0.995	0.978	0.992	0.995	0.996	0.985	0.990	0.992	0.998	0.988	0.997
	Comp >6	0.001	0.005	0.005	0.022	0.008	0.005	0.004	0.015	0.010	0.008	0.002	0.012	0.003
	Mean	6.000	7.986	8.000	7.991	7.999	7.999	7.999	7.999	8.000	8.000	8.000	8.000	8.000
	SD	0.000	0.118	0.000	0.094	0.032	0.032	0.032	0.032	0.000	0.000	0.000	0.000	0.000
GPVC	Comp <6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Comp =6	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Comp >6	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	Mean	6.000	7.987	8.000	7.990	8.000	7.999	8.000	7.999	7.999	8.000	8.000	8.000	8.000
	SD	0.000	0.113	0.000	0.100	0.000	0.032	0.000	0.032	0.032	0.000	0.000	0.000	0.000
PVC	Comp <6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Comp =6	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Comp >6	0.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Kaiser-Guttman criterion														
	Mean	5.999	5.999	6.000	5.999	6.000	6.000	6.000	6.000	6.000	6.000	5.998	6.000	6.000
	SD	0.032	0.032	0.000	0.032	0.000	0.000	0.000	0.000	0.000	0.000	0.045	0.000	0.000
RPVC	Comp <6	0.001	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000
	Comp =6	0.999	0.999	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000
	Comp >6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	Mean	6.000	7.218	7.868	7.341	7.453	7.574	8.008	4.901	4.654	5.963	5.701	2.183	3.995
	SD	0.000	0.534	0.630	0.660	0.854	0.532	0.737	0.765	0.677	0.508	0.636	0.389	0.071
GPVC	Comp <6	0.000	0.000	0.000	0.003	0.014	0.000	0.000	0.787	0.904	0.144	0.353	1.000	1.000
	Comp =6	1.000	0.057	0.007	0.096	0.115	0.019	0.017	0.209	0.091	0.747	0.572	0.000	0.000
	Comp >6	0.000	0.943	0.993	0.901	0.871	0.981	0.983	0.004	0.005	0.109	0.075	0.000	0.000
	Mean	6.000	7.088	7.968	7.617	3.315	7.599	5.089	5.588	2.117	6.188	2.294	3.288	2.000
	SD	0.000	0.568	0.501	0.615	0.783	0.502	0.832	0.995	0.322	0.539	0.458	0.860	0.000
PVC	Comp <6	0.000	0.000	0.000	0.000	0.989	0.000	0.690	0.429	1.000	0.062	1.000	0.996	1.000
	Comp =6	1.000	0.121	0.004	0.071	0.010	0.006	0.277	0.424	0.000	0.686	0.000	0.004	0.000
	Comp >6	0.000	0.879	0.996	0.929	0.001	0.994	0.033	0.147	0.000	0.252	0.000	0.000	0.000

(a) 25%, 50% and 100% of series contaminated at time $t = 500$ and 999 (isolated outliers) or $t = 500, 501$ and $998, 999$ (consecutive outliers). (b) 0% represents uncontaminated series. (c) Size of outliers $\omega = 5$ and 10 standard deviations of the univariate uncontaminated process. (d) the factor models are simulated with six factors.

eigenvectors and assume that we know the true number of components with hereroskedastic dynamic. We follow Li et al. (2016) and carry out a similar Monte Carlo experiments with 1000 replicates and consider small ($N = 8$) and large ($N = 100$) dimension cases. Following examples 1 and 4 of Li et al. (2016), the factor model is driven by just one common factor which follows a Gaussian GARCH(1,1) process with parameters $\omega = 1$, $\alpha = 0.07$ and $\beta = 0.83$. The idiosyncratic factors ϵ_t are simulated as $\epsilon_t = \frac{\bar{\epsilon}_t}{\sqrt{N}}$ where $\bar{\epsilon}_t \sim Normal_N(0, I)$ with N being the dimension of the system. The factor load matrix A is also normalized and each element is a random draw of $U(-1, 1)$. Given that the PVC and GPVC procedures have similar performance (Li et al., 2016) and considering the extreme computational cost of the Monte Carlo experiment using the PVC procedure when $N = 100$, for large dimensions we only consider the GPVC and RPVC procedures. To compare the estimation of the matrix A we use the two measures[3] defined in Li et al. (2016) and given by

$$d\left(\hat{\mathcal{M}}_1, \mathcal{M}_1\right) = \sqrt{1 - \frac{Tr\left(\hat{A}\hat{A}'AA'\right)}{r}}, \tag{7}$$

$$d\left(\hat{A}, A\right) = 1 - \frac{\left[\sum_t (y_t - \bar{y})'\hat{A}A'(y_t - \bar{y})\right]^2}{\left[\sum_t (y_t - \bar{y})'\hat{A}\hat{A}'(y_t - \bar{y})\right]\left[\sum_t (y_t - \bar{y})'AA'(y_t - \bar{y})\right]}, \tag{8}$$

where y_t is a vector of observed returns, A is the true load factor matrix, $\hat{A}$ is the estimated load factor matrix and r is the number of columns of $\hat{A}$. Figure 1 presents the box-plot of the results of the measure defined in (7) for small and large dimension cases. We can observed that, the effect of outliers in PVC and GPVC procedures is devastating even when just a few outliers are added in the series.

In general, the analyses of the results show that the RPVC is less sensitive to outliers and also stable despite the proportion of series contaminated, the size of outliers and if outliers are isolated or consecutive. Note also that, in the absence of outliers the performance of all procedures is almost similar with a slight better performance of the GPVC procedure. Results using (8) produces similar results and are available in the supplementary material.

[3]These measures and the measures used in the next section are implemented in the R package *StatPerMeco* of Trucios (2017).

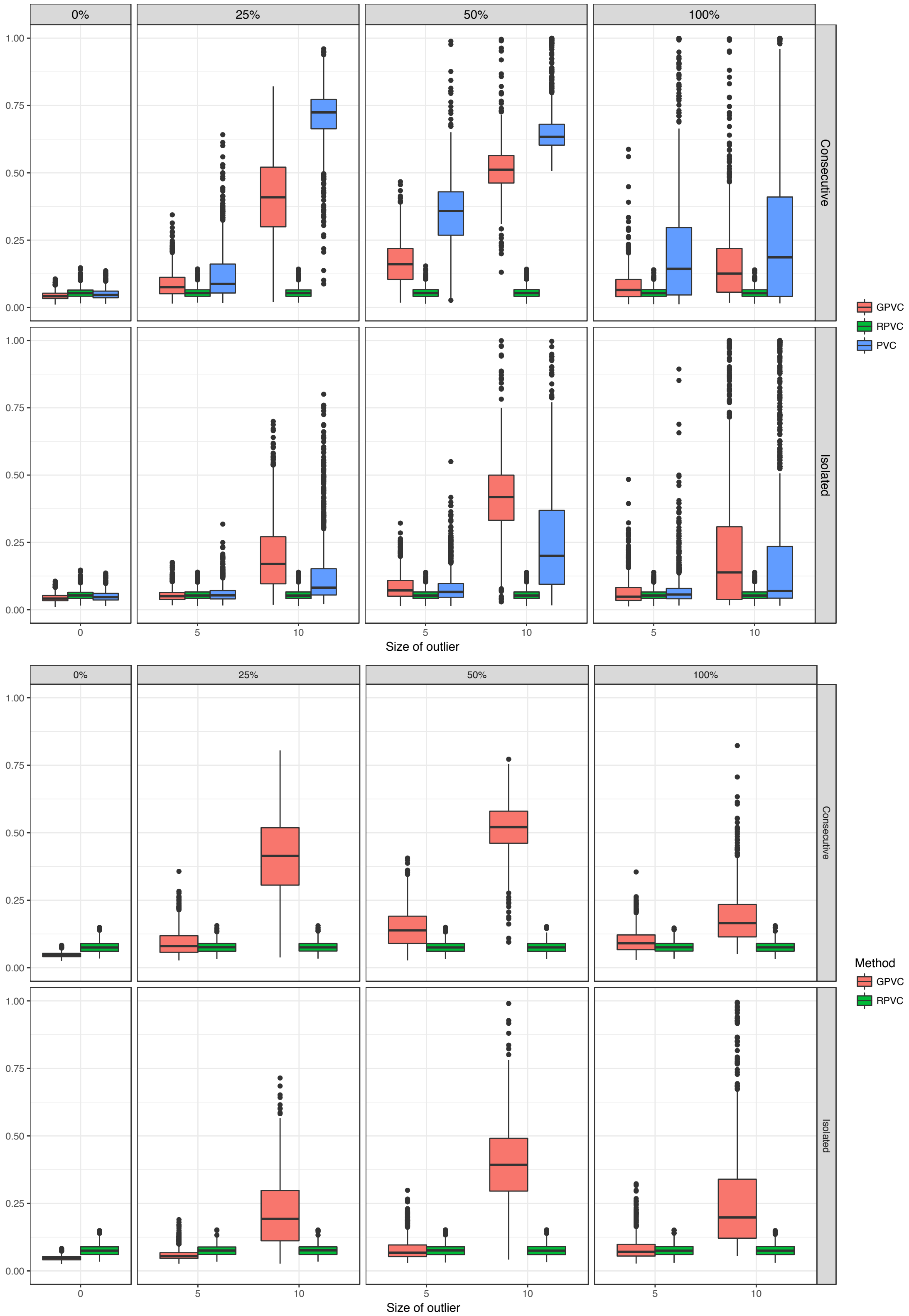

Figure 1: Boxplot of $d(\hat{\mathcal{M}}_1, \mathcal{M}_1)$ for uncontaminated (0%) and contaminated series with 25%, 50% and 100% of series contaminated. Dimension 8 (top panel) and 100 (bottom panel), $T = 1000$ and outliers of size $\omega = 0$, 5 and 10 standard deviations of the univariate uncontaminated process. 1000 replicates

3.3 Conditional covariance matrix

In this last section on Monte Carlo experiment, we use the same DGP used in the previous section to analyse the effects of outliers on the one-step-ahead forecast of the conditional covariance matrix. Because the PVC and the RPVC procedures have similar performance (Sections 3.1, 3.2 and Li et al. (2016)) and given the computational cost of the PVC procedure, hereafter we focus on the GPVC procedure and compare it with our robust proposal.

It can be shown (Section 2.2 of Li et al. (2016)) that the forecast h-steps-ahead of the conditional covariance matrix can be obtained through

$$\hat{\Sigma}_y(h) = \hat{A}\hat{\Sigma}_{\hat{f}}(h)\hat{A}' + \hat{A}\hat{A}'\hat{\Sigma}_y\hat{B}\hat{B}' + \hat{B}\hat{B}'\hat{\Sigma}_y, \tag{9}$$

where $\hat{\Sigma}_{\hat{f}}(h)$ is the conditional covariance matrix h-steps-ahead of the estimated components $\hat{f}$, $\hat{\Sigma}_y$ is the estimated unconditional covariance matrix of y and $\hat{A}$ and $\hat{B}$ are estimated eigenvectors. The one-step-ahead volatility of the volatility component is estimated using a Student-t quasi maximum likelihood (QML) GARCH(1,1) model for the GPVC and by the robust procedure of Boudt et al. (2013) with the filter used in Trucíos et al. (2017b) for the RPVC. It is important to mention that in real life data, a deep analyse should be made to choose the best model to be fitted.

Figures 2 and 3 report the MSE[4] and the MAE for contaminated and uncontaminated series in small and large dimension cases. The results show a devastating impact of outliers on the forecasting of the conditional covariance matrix when the non-robust procedure is used. Observe also that for uncontaminated series the both procedures have a similar performance. However, when outliers are present in the series the advantage of the robust procedure is clear, even in small dimensions. Also note that, for both criteria, when the dimension of the series is small and when only 25% of series are contaminated with outliers of size $\omega = 5$, the performance of both procedures are similar regardless if outliers are isolated or consecutive.

It is clear that, the non-robust volatility components procedure is very sensitive to outliers and can lead to improper estimation of the conditional covariance matrix, even

[4]Results for the MSE were cut-off in the value of 50 for small dimensions and in 3 for large dimensions to improve the visualization in the figure.

when the true number of volatility components with heteroskedastic dynamic is know. The consequences of using a non-robust procedure to forecast the conditional covariance matrix when outliers are present in the series can be disastrous leading for instance, as observed in our empirical application, to misspecified portfolio allocation and improper forecast of the portfolio volatility. These results are in concordance with the obtained by, for instance, Boudt et al. (2013) and Trucíos et al. (2017a) who showed that the effects of outliers affect the estimation of the conditional covariance matrix and consequently economic measures based on its estimation.

4 Empirical application

In this section, we implement the RPVC procedure to analyse the daily returns of stocks used in the construction of the Nasdaq-100 index traded from January 6, 2001 to May 12, 2017. Because not all stocks of the index were traded during the entire period, we ended up with $N = 73$ stocks. The daily prices are available at *finance.yahoo.com* and were downloaded on May 14 2017 using the R package *quantmod* of Ryan (2017). Returns are computed as usual by $r_{i,t} = 100 \times log\,(P_{i,t}/P_{i,t-1})$, where $P_{i,t}$ denotes the adjusted closing price of the ith stock at day t for $i = 1, ..., 73$.

With illustrative purposes we use the one-step-ahead forecast of the conditional covariance matrix to estimate the 1% and 5% Value-at-Risk (VaR) of the equal-weight portfolio as well as to construct the minimum variance portfolio (MVP) with short-sale constrains. The VaR is calculated assuming a Student-t distribution where the degrees of freedom is estimated by maximum likelihood using the portfolio innovations[5] and the MVP are rebalanced daily.

In the VaR and the MVP applications, we use a rolling windows approach with window size of 1000 days and all results are compared with the GPVC and OGARCH (Alexander and Chibumba, 1996) procedures. We use the OGARCH procedure because this method

[5] For $t = 1, ...1000$, the portfolio innovations were obtained through $e_{p,t} = r_{p,t}/\sqrt{\sigma^2_{p,t}}$ where $r_{p,t} = \omega \times (r_{1,t}, ...r_{73,t})'$ and $\sigma^2_{p,t} = \omega \hat{H}_t \omega'$ are the portfolio returns and variances at time t respectively with ω and $\hat{H}_t$ being a vector of equal-weights and the estimated conditional covariance matrix respectively. The VaR is obtained by $VaR = Z_\alpha \sigma_{p,T+1|T}$ where Z_α is the α quantile of the standardized Student-t distribution with ν degrees of freedom (ν is estimated using the portfolio innovations) and $\sigma^2_{p,T+1|T} = \omega \hat{H}_{T+1|T} \omega'$ is the one-step-ahead forecast conditional portfolio variance.

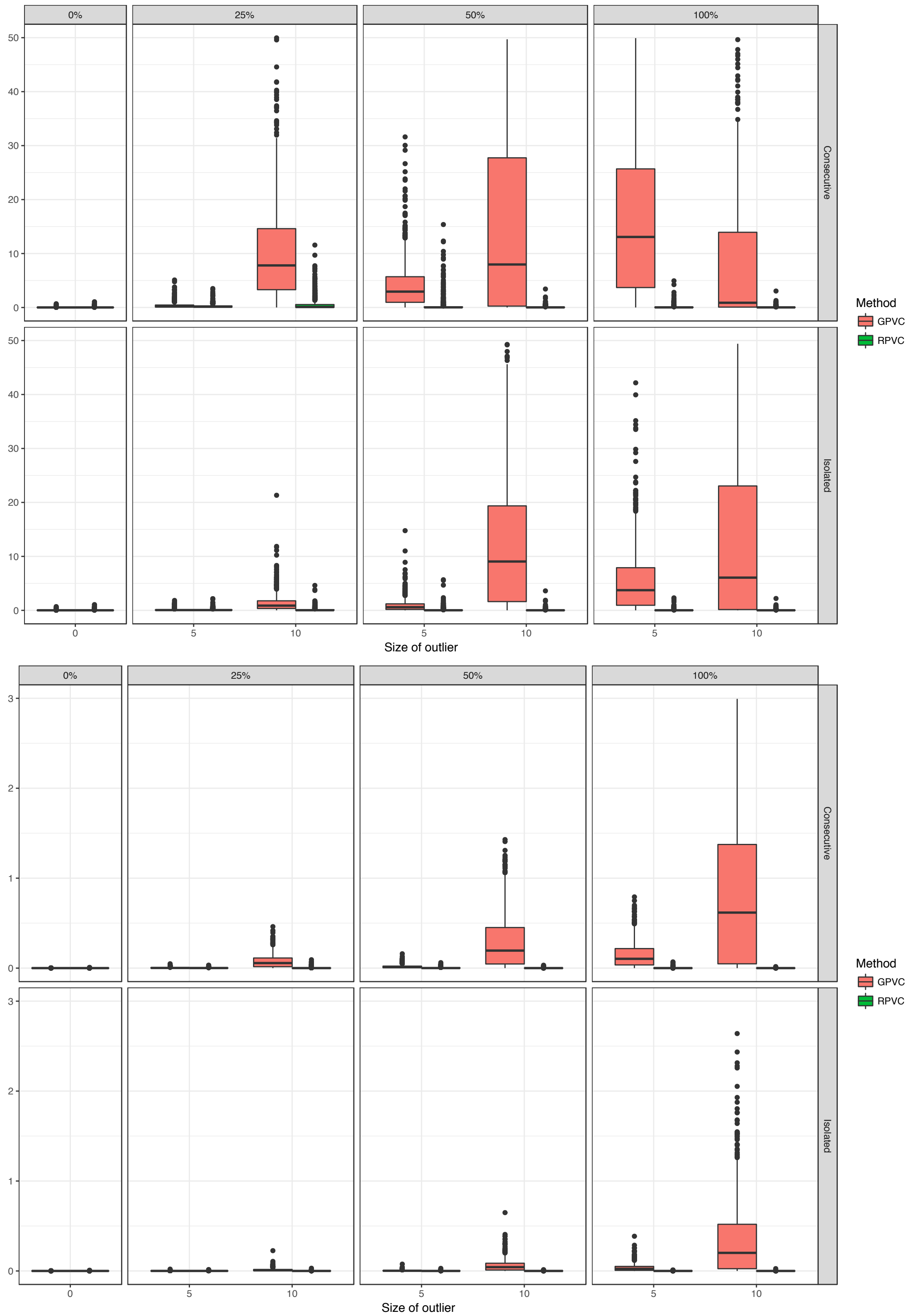

Figure 2: Boxplot of MSE for uncontaminated (0%) and contaminated series with 25%, 50% and 100% of series contaminated. Dimension 8 (top panel) and 100 (bottom panel), $T = 1000$ and outliers of size $\omega = 0$, 5 and 10 standard deviations of the univariate uncontaminated process. 1000 replicates.

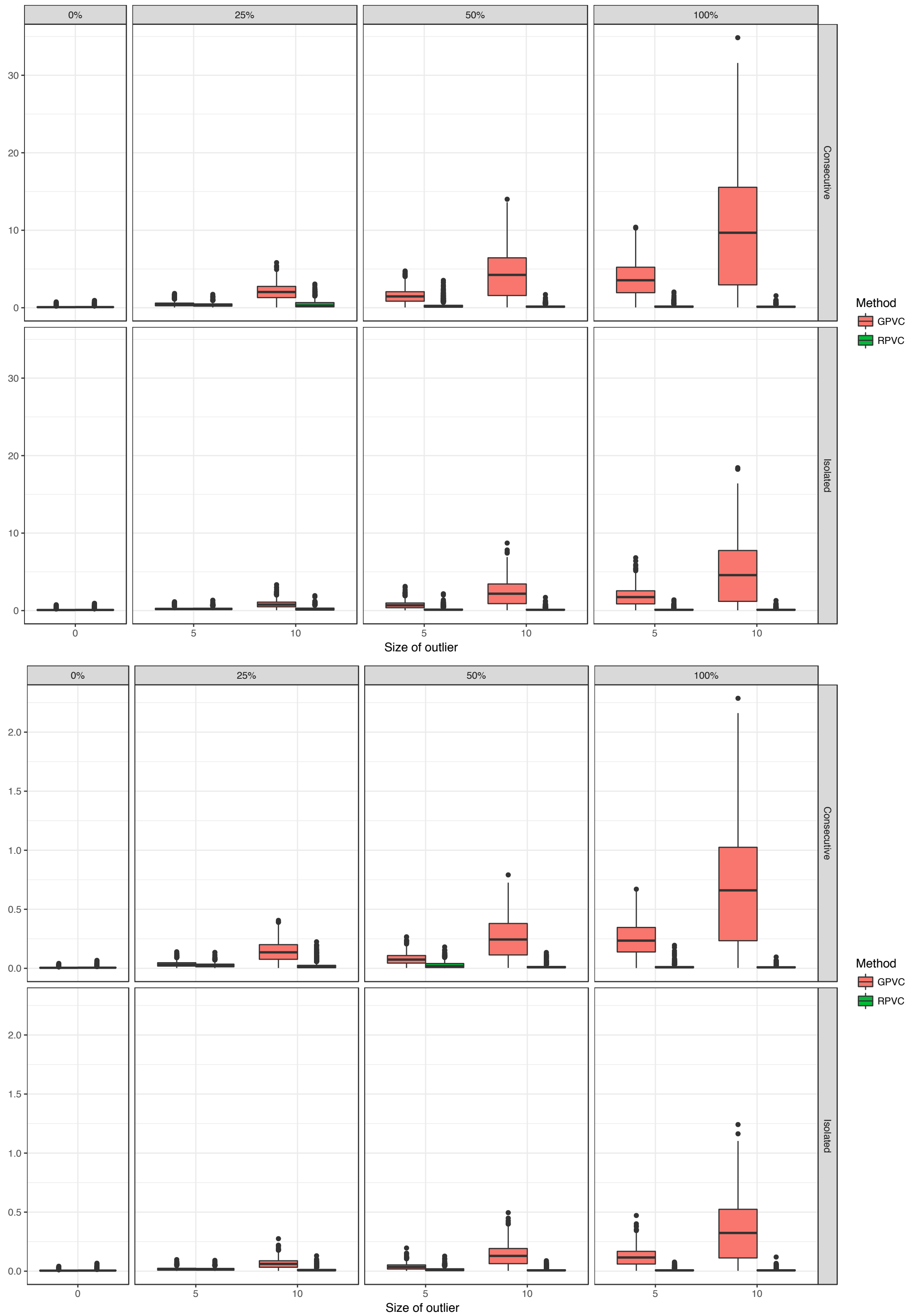

Figure 3: Boxplot of MAE for uncontaminated (0%) and contaminated series with 25%, 50% and 100% of series contaminated. Dimension 8 (top panel) and 100 (bottom panel), $T = 1000$ and outliers of size $\omega = 0$, 5 and 10 standard deviations of the univariate uncontaminated process. 1000 replicates.

has been used in several papers as a dimension reduction technique benchmark (Santos and Moura, 2014; Becker et al., 2015; Santos and Ferreira, 2017). The OGARCH model was estimated as in Becker et al. (2015), that means, considering the number of components equal to the number of series and each components being modelled as a GARCH(1,1) process. The GARCH model in all dimension reduction techniques was estimated assuming a Student-t distribution[6]. Additionally, we compare our robust procedure with the classical Risk Metrics (RM) methodology (Morgan, 1996), the new version, called, Risk Metrics 2006 (RM2006) methodology (Zumbach, 2007) and with the DCC model using composite likelihood and non-linear shrinkage as in Engle et al. (2017).

Table 3 reports the percentage of violation (returns smaller than the VaR) and the p-values of the back-testing tests of unconditional coverage (Kupiec, 1995), independence and conditional coverage (Christoffersen, 1998) for the 1% and 5% VaR of the equal-weight portfolio. The number of selected volatility components in the GPVC and RPVC procedures is estimated using the the ratio estimator (Lam and Yao, 2012; Ahn and Horenstein, 2013), the BN (Bai and Ng, 2002) and the Kaiser-Guttman (Guttman, 1954) criteria. The ratio estimator criterion suggests to use one component in both cases (GPVC and RPVC), the BN criterion suggests to use three components and the Kaiser-Guttman criterion suggest to use three components when the GPVC procedure is used and four components in the robust procedure. With illustrative and comparative purposes we fit the stocks of the Nasdaq 100 index to the models using one, two, three and four volatility components. The conditional variance of the volatility component is forecasted using the same strategy described in Subsection 3.3 and conditional covariance matrices in cases with more than one volatility component are forecasted using the DCC model. The DCC model was estimated by Aielli (2013) and Boudt et al. (2013) methods for the GPVC and RPVC procedures respectively.

The percentage of violations using the dimension reduction techniques (GPVC, RPVC, OGARCH) is close to the nominal one and in those cases the unconditional coverage (UC) and conditional coverage (CC) tests fail to reject the null hypotheses at 5% of significance level. Note that results for the 5% VaR using the GPVC procedure are the same regardless of the number of selected volatility components used. In fact, the values of the 5% VaR using the GPVC (not reporting here) are very similar and no differences are observed in the

[6]All codes in this paper were implemented in the R software (R Core Team, 2017)

proportion of violations. The RM2006 methodology slightly underestimate the percentage of violations but the UC and CC tests also fail to reject the null hypotheses. The RM methodology has a good performance for the 5% VaR, although, the percentage of violations is overestimated in the 1% VaR and the UC and CC tests reject the null hypotheses at 5% of significance level. The DCC procedure reports results in the 5% VaR similar than the obtained with the GPVC procedure.

Figure 4 shows the 1% VaR of the equal-weight portfolio using the GPVC and RPVC procedures with one volatility component. Note that after large returns, the VaR obtained using the GPVC (solid green line) procedure is unnecessarily larger than the obtained using the RPVC (dashed red line) procedure, implying in more capital requirements. Additional figures comparing the RPVC with the OGARCH, RM, RM2006 and DCC procedures are in the supplementary material.

Table 3: Percentage of violations (returns smaller than VaR) and p-values of the unconditional coverage (UC), independence (IND) and conditional coverage (CC) tests. VaR 1% (top panel) and VaR 5% (bottom panel) of the equal-weight portfolio

	Method	% violations	UC	IND	CC
1% VaR	GPVC 1VC	0.868	0.448	0.492	0.587
	RPVC 1VC	0.932	0.699	0.029	0.085
	GPVC 2VC	0.835	0.342	0.508	0.507
	RPVC 2VC	0.996	0.983	0.038	0.116
	GPVC 3VC	0.835	0.342	0.508	0.507
	RPVC 3VC	0.996	0.983	0.429	0.724
	GPVC 4VC	0.900	0.567	0.476	0.652
	RPVC 4VC	1.157	0.391	0.358	0.449
	OGARCH	0.867	0.447	0.492	0.586
	RM	1.478	0.012	0.185	0.017
	RM 2006	0.803	0.254	0.197	0.224
	DCC	0.964	0.839	0.034	0.101
5% VaR	GPVC 1VC	4.724	0.475	0.022	0.053
	RPVC 1VC	4.981	0.961	0.273	0.521
	GPVC 2VC	4.724	0.475	0.022	0.053
	RPVC 2VC	4.981	0.961	0.273	0.521
	GPVC 3VC	4.724	0.475	0.022	0.053
	RPVC 3VC	5.141	0.719	0.391	0.616
	GPVC 4VC	4.724	0.475	0.022	0.053
	RPVC 4VC	5.205	0.601	0.592	0.716
	OGARCH	4.916	0.830	0.141	0.314
	RM	4.948	0.895	0.132	0.303
	RM 2006	4.531	0.223	0.120	0.135
	DCC	4.724	0.475	0.417	0.531

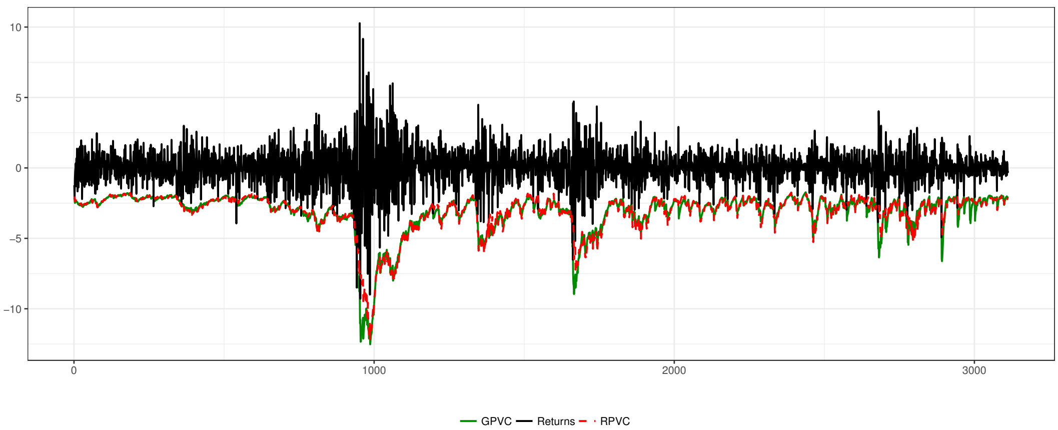

Figure 4: 1% VaR of the equal-weight portfolio using the GPVC (solid green line) and RPVC (dashed red line) procedures (considering one volatility component)

Following Engle et al. (2017), Gambacciani and Paolella (2017) and Trucíos et al. (2017a), Table 4 reports annualized performance measures. The results are presented from January 3, 2005 to May 12, 2017 (entire out-of-sample period) and also for August 1, 2007 to December 31, 2013 (high volatile out-of-sample period). For the sake of comparison, we also implement the equal-weighted portfolio. See, Fan et al. (2012), Engle et al. (2017) and Gambacciani and Paolella (2017) for some references where the naive equal-weighted portfolio has also used with comparison purpose.

Note that, as mentioned by Engle et al. (2017) and Ledoit and Wolf (2017), notwithstandingly high Sharpe/Sortino ratio as well as minimum observed standard deviation are all desirable properties, the MVP is calculated to achieve minimum variance and the observed annualized standard deviation should be the main focus in the comparison.

Results given in Table 4 show that in both periods the worst results are obtained when the equal-weighted portfolio is used, obtaining the highest standard deviation (SD) and also the smallest Sharpe and Sortino ratio. On the other hand, regardless of the number of selected volatility components used, the RPVC procedure always outperforms the GPVC in terms of SD and most of cases also in terms of Sharpe and Sortino ratio, the unique exception is observed in the period 2007-2013 when the Sharpe and Sortino ratios are smaller using the RPVC with three volatility components than using the GPVC procedure with the same number of components. The OGARCH procedure never outperforms the RPVC procedure in terms of SD, although, sometimes outperforms the GPVC. In terms of the Sharpe and Sortino ratios, in the high volatile period the OGARCH procedure is always outperformed by the RPVC and GPVC procedures. However, in the entire period the OGARCH method outperforms the GPVC in all cases but the RPVC procedure with one and two volatility components outperforms the OGARCH. Among the dimension reduction techniques used in this paper, the RPVC procedure presents the best performance. Note that among the dimension reduction techniques the smallest SD in both periods are obtained using the RPVC procedure with four volatility components and the highest Sharpe ratio is obtained using the RPVC procedure with two volatility components. The highest Sortino ratio is obtained using the RPVC with two volatility components in the entire period and four volatility components in the high volatile period.

Table 4 also reports results of the comparison between the RPVC procedure and the RM, RM2006 and DCC methodologies. In the entire period, the RPVC procedure presents a smaller SD than the obtained using the RM procedure and in the 2007-2013 period the RM methodology is outperformed by the RPVC procedure when using four volatility components. In all cases, the RM procedure presents largest Sharpe and Sortino ratios. The RM2006 and DCC procedures outperform all the other procedues in terms of SD. The smallest SD in the entire period is obtained using the DCC procedure while the smallest SD in the high volatile period is obtained using the RM2006 methodology . Similar results where dimension reduction techniques are outperformed by multivariate volatility models including simplest models such as EWMA, ORE (Foster and Nelson, 1996) and RiskMetrics in a MVP context can be found in, for instance, Han (2006), Santos and Moura (2014), Becker et al. (2015), Caldeira et al. (2017), Santos and Ferreira (2017) and Ledoit and Wolf (2017). Han (2006) and Han (2007) point out that not necessarily best statistical models have a better portfolio performance than simplest ones. Analysis in which scenarios simplest models could have a better performance than dimension reductions techniques is an interesting further research topic. Additionally, recent results of Barigozzi and Hallin (2017) shows that the introduction of a dynamic structure in volatility factor models improve the forecasting performance in comparison with static approaches.

Results obtained in this paper, probably, could be improved using a dynamic component approach (Barigozzi and Hallin, 2015; Peña and Yohai, 2016; Barigozzi and Hallin, 2017; Peña et al., 2017) and this topic is in our research agenda.

Table 4: Annualized performance measures for the selected MVP using equal-weight strategy, GPVC and RPVC procedures (with one to four volatility components)

	Jan 3, 2005 - May 12, 2017			Aug 1, 2007 - Dic 31, 2013		
Method	SD	ShR	SR	SD	ShR	SR
1/N	20.7347	0.6272	0.8796	24.8567	0.5064	0.7083
GPVC 1VC	16.8537	0.6773	0.9445	20.2005	0.6159	0.8535
RPVC 1VC	16.7884	0.7850	1.0955	19.9746	0.6481	0.8946
GPVC 2VC	16.8171	0.6888	0.9619	20.1418	0.6400	0.8885
RPVC 2VC	16.7383	0.8155	1.1409	19.9551	0.6683	0.9238
GPVC 3VC	16.8266	0.7105	0.9904	20.1810	0.6431	0.8901
RPVC 3VC	16.6621	0.7181	1.0022	19.9180	0.6071	0.8395
GPVC 4VC	16.7954	0.7131	0.9930	20.1550	0.6505	0.8989
RPVC 4VC	16.6261	0.7465	1.0481	19.8679	0.6642	0.9264
OGARCH	16.8062	0.7801	1.0935	20.1997	0.5559	0.7699
RM	16.8517	1.0880	1.5548	19.9113	0.9218	1.2948
RM 2006	16.3518	0.9929	1.3932	19.4390	0.7953	1.0948
DCC	16.3474	0.9307	1.3059	19.5645	0.7779	1.0735

(a) SD: Standard deviation of the out-of-sample portfolio returns multiplied by $\sqrt{252}$. (b) ShR: Annualized Sharpe ratio. (c) SR: Annualized Sortino ratio (Sortino and van der Meer, 1991).

Three important comments are necessary to point out before to finish our empirical application. First, although the criteria used in this paper to select the number of volatility components are not conclusive and each criteria suggest a different number, regardless of the number of selected volatility components, our robust procedure showed a better performance than the non-robust version. However, an optimal procedure to select the number of estimated volatility components its necessary. Second, the RPVC procedure showed to be better than the other dimension reduction techniques used in this paper as well as the equal-weight strategy. However, in the MVP context, RPVC is outperformed by the Risk Metrics 2006 procedure. This results could be explained by the static structure of the RPVC procedure. A new procedure, taking into account a dynamic approach in the volatility components could improve our results and it is a further research topic. Finally, the criteria used to asses the performance of forecast one-step-ahead conditional covariance matrix using RPVC procedure indicates that the RPVC procedure is superior to the GPVC, which is in concordance with the Monte Carlo experiments carried out in Section 3.

5 Conclusions and further research topics

In this paper, we focus on the principal volatility components procedure of Hu and Tsay (2014a) and Li et al. (2016). These procedures extract few components with time varying volatility and the remainder components with constant volatility tackling the problem of the curse of the dimensionality.

We analyse the problem of modelling and forecasting the conditional covariance matrix via principal volatility components in the presence of outliers and show that just a few outliers are sufficient to affect drastically the volatility components and the estimation of the conditional covariance matrix.

To estimate the number of selected volatility components we used the estimator criterion (Lam and Yao, 2012; Ahn and Horenstein, 2013), the BN criterion (Bai and Ng, 2002) and the Kaiser-Guttman criterion (Guttman, 1954). The use of the ration estimator and BN criteria estimated the number of components close (or equal) to the true number of factors. However, the Kaiser-Guttman criterion reports problems to estimate correctly the number of component when the ratio factors/dimension increase, for that reason we do not recommend its use in such cases.

A new and robust procedure with good finite sample properties based on a robust estimator of the unconditional covariance matrix, a weighted estimator and robust filters was introduced.

The principal volatility components approach brings an innovative contribution in the field of modelling and forecasting multivariate volatility, managing portfolios and quantifying risk. However, it is necessary caution when the data is contaminated by outliers because disastrous results can be obtained when using the non-robust procedures of Hu and Tsay (2014a) or Li et al. (2016). This paper contributes to the literature in two ways: it call the attention to the risk of using these procedures in the presence of outliers and introduces an approach robust to outliers and with a similar performance in uncontaminated series.

In our empirical application, the one-step-ahead forecast of conditional covariance matrix was used to estimate the VaR and also to construct the MVP. In both applications the RPVC procedure had a better performance than the GPVC. This results are in concordance with our Monte Carlo experiments and show the superiority of the RPVC procedure against the

GPVC one.

The problem of dealing with $N/T \to 1$ or even $N > T$ has not being addressed here. This topic as well as PVC in switching regime are in our research agenda.

The aim of this paper is not to compare the predictability of the volatility using different approaches but to robustify the principal volatility components method. An extensive comparison using other recently procedures such as Matteson and Tsay (2011), Pakel et al. (2014) Matilainen et al. (2015), Barigozzi and Hallin (2015), Peña and Yohai (2016), Barigozzi and Hallin (2017), Engle et al. (2017) among other in different scenarios is an interesting further research topic.

Finally, some papers such as Han (2006), Santos and Moura (2014), Becker et al. (2015), Caldeira et al. (2017), Santos and Ferreira (2017) and Ledoit and Wolf (2017) have reported in their empirical application that in a MVP context, dimension reduction techniques are sometimes outperformed by classical multivariate volatility models including simplest models such as EWMA, ORE (Foster and Nelson, 1996) and RiskMetrics. An interesting research topic is to evaluate some recently dimension reduction techniques and compare it in a MVP context with other multivariate volatility models as well as analyse the reasons why a better performance of the simplest models is observed in the papers previously mentioned.

References

Ahn, S. C. and Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3):1203–1227.

Aielli, G. P. (2013). Dynamic conditional correlation: on properties and estimation. *Journal of Business & Economic Statistics*, 31(3):282–299.

Alexander, C. and Chibumba, A. (1996). Multivariate orthogonal factor GARCH. *Working paper. Sussex: University of Sussex.*

Andreou, E. and Ghysels, E. (2014). Comment: Principal volatility component analysis. *Journal of Business & Economic Statistics*, 32(2):168–171.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Barigozzi, M. and Hallin, M. (2015). Generalized dynamic factor models and volatilities: recovering the market volatility shocks. *The Econometrics Journal*, 19(1):C33–C60.

Barigozzi, M. and Hallin, M. (2017). Generalized dynamic factor models and volatilities: estimation and forecasting. *Journal of Econometrics*, 201(21):307–321.

Basher, S. A. and Sadorsky, P. (2016). Hedging emerging market stock prices with oil, gold, VIX, and bonds: A comparison between DCC, ADCC and GO-GARCH. *Energy Economics*, 54(1):235–247.

Becker, R., Clements, A. E., Doolan, M., and Hurn, A. (2015). Selecting volatility forecasting models for portfolio allocation purposes. *International Journal of Forecasting*, 31(3):849–861.

Borchers, H. W. (2017). *pracma: Practical Numerical Math Functions.* R package version 2.0.4.

Boudt, K. and Croux, C. (2010). Robust M-estimation of multivariate GARCH models. *Computational Statistics & Data Analysis*, 54(11):2459–2469.

Boudt, K., Danielsson, J., and Laurent, S. (2013). Robust forecasting of dynamic conditional correlation GARCH models. *International Journal of Forecasting*, 29(2):244–257.

Caldeira, J. F., Moura, G. V., Nogales, F. J., and Santos, A. A. P. (2017). Combining multivariate volatility forecasts: an economic-based approach. *Journal of Financial Econometrics*, 15(2):247–285.

Candès, E. J., Li, X., Ma, Y., and Wright, J. (2011). Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):11:1–11:37.

Carnero, M. A., Peña, D., and Ruiz, E. (2012). Estimating GARCH volatility in the presence of outliers. *Economics Letters*, 114(1):86–90.

Charles, A. and Darné, O. (2014). Large shocks in the volatility of the Dow Jones Industrial Average index: 1928–2013. *Journal of Banking & Finance*, 43:188–199.

Chiou, S. C. and Tsay, R. S. (2008). A copula-based approach to option pricing and risk assessment. *Journal of Data Science*, 6(3):273–301.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, 29(4):841–862.

Croux, C., Gelper, S., and Mahieu, K. (2010). Robust exponential smoothing of multivariate time series. *Computational Statistics & Data Analysis*, 54(12):2999–3006.

Croux, C. and Haesbroeck, G. (2000). Principal component analysis based on robust estimators of the covariance or correlation matrix: influence functions and efficiencies. *Biometrika*, 87(3):603–618.

Engle, R. F., Ledoit, O., and Wolf, M. (2017). Large dynamic covariance matrices. *Journal of Business & Economic Statistics (Forthcoming)*.

Fan, J., Li, Y., and Yu, K. (2012). Vast volatility matrix estimation using high-frequency data for portfolio selection. *Journal of the American Statistical Association*, 107(497):412–428.

Fan, J., Wang, M., and Yao, Q. (2008). Modelling multivariate volatilities via conditionally uncorrelated components. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(4):679–702.

Foster, D. P. and Nelson, D. B. (1996). Continuous record asymptotics for rolling sample variance estimators. *Econometrica*, 64(1):139.

Franke, J. (2014). Comment: Principal volatility component analysis. *Journal of Business & Economic Statistics*, 32(2):171–172.

Gambacciani, M. and Paolella, M. S. (2017). Robust normal mixtures for financial portfolio allocation. *Econometrics and Statistics*, 3:91–111.

García-Ferrer, A., González-Prieto, E., and Peña, D. (2012). A conditionally heteroskedastic independent factor model with an application to financial stock returns. *International Journal of Forecasting*, 28(1):70–93.

Grané, A., Veiga, H., and Martín-Barragán, B. (2014). Additive level outliers in multivariate GARCH models. In Melas, V., Mignani, S., Monari, P., and Salmaso, L., editors, *Topics in Statistical Simulation*, volume 114, pages 247–255. Springer, New York.

Greco, L. and Farcomeni, A. (2016). A plug-in approach to sparse and robust principal component analysis. *Test*, 25(3):449–481.

Gruber, L. F. and West, M. (2016). Bayesian forecasting and scalable multivariate volatility analysis using simultaneous graphical dynamic models. *arXiv preprint arXiv:1606.08291*.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika*, 19(2):149–161.

Hafner, C. M. and Reznikova, O. (2012). On the estimation of dynamic conditional correlation models. *Computational Statistics & Data Analysis*, 56(11):3533–3545.

Hammoudeh, S. M., Yuan, Y., McAleer, M., and Thompson, M. A. (2010). Precious metals–exchange rate volatility transmissions and hedging strategies. *International Review of Economics & Finance*, 19(4):633–647.

Han, Y. (2006). Asset allocation with a high dimensional latent factor stochastic volatility model. *The Review of Financial Studies*, 19(1):237–271.

Han, Y. (2007). Return predictability, economic profits, and model mis-specification: How important are the better specified models? *Working Paper available at SSRN: https://ssrn.com/abstract=967564*.

Hu, Y.-P. and Tsay, R. S. (2014a). Principal volatility component analysis. *Journal of Business & Economic Statistics*, 32(2):153–164.

Hu, Y.-P. and Tsay, R. S. (2014b). Rejoinder: Principal volatility component analysis. *Journal of Business & Economic Statistics*, 32(2):176–177.

Hubert, M., Rousseeuw, P., Vanpaemel, D., and Verdonck, T. (2015). The detS and detMM estimators for multivariate location and scatter. *Computational Statistics & Data Analysis*, 81(1):64–75.

Hubert, M., Rousseeuw, P. J., and Vanden Branden, K. (2005). ROBPCA: a new approach to robust principal component analysis. *Technometrics*, 47(1):64–79.

Hubert, M., Rousseeuw, P. J., and Verdonck, T. (2012). A deterministic algorithm for robust location and scatter. *Journal of Computational and Graphical Statistics*, 21(3):618–637.

Joe, H. (2006). Generating random correlation matrices based on partial correlations. *Journal of Multivariate Analysis*, 97(10):2177–2189.

Kastner, G. (2016). Sparse Bayesian time-varying covariance estimation in many dimensions. *arXiv preprint arXiv:1608.08468*.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2):73–84.

Lam, C. and Yao, Q. (2012). Factor modeling for high-dimensional time series: inference for the number of factors. *The Annals of Statistics*, 40(2):694–726.

Laurent, S., Lecourt, C., and Palm, F. C. (2016). Testing for jumps in conditionally Gaussian

ARMA–GARCH models, a robust approach. *Computational Statistics & Data Analysis*, 100:383–400.

Ledoit, O. and Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. *The Review of Financial Studies*, 30(12):4349–4388.

Li, W., Gao, J., Li, K., and Yao, Q. (2016). Modeling multivariate volatilities via latent common factors. *Journal of Business & Economic Statistics*, 34(4):564–573.

Lopes, H. F., McCulloch, R. E., and Tsay, R. S. (2012). Cholesky stochastic volatility models for high-dimensional time series. *Discussion papers*.

Ma, Y. and Genton, M. G. (2000). Highly robust estimation of the autocovariance function. *Journal of Time Series Analysis*, 21(6):663–684.

Matilainen, M., Nordhausen, K., and Oja, H. (2015). New independent component analysis tools for time series. *Statistics & Probability Letters*, 105:80–87.

Matteson, D. S. and Tsay, R. S. (2011). Dynamic orthogonal components for multivariate time series. *Journal of the American Statistical Association*, 106(496):1450–1463.

Monte, E. Z. and Reisen, V. A. (2016). Robust principal volatility componnet analysis: an application to air pollution in the Greater Vitória Region, Espírito Santo, Brazil. In *Principal components analysis in multivariate time series with conditional heteroscedasticity and outliers: an application for air pollution, in the region of Greater Vitória, Espirito Santo, Brazil. (PhD Thesis)*.

Morgan, J. P. (1996). *RiskMetrics—Technical Document*. New York, NY.

Muler, N. and Yohai, V. J. (2008). Robust estimates for GARCH models. *Journal of Statistical Planning and Inference*, 138(10):2918–2940.

Pakel, C., Shephard, N., Sheppard, K., and Engle, R. (2014). Fitting vast dimensional time-varying covariance models. *Manuscript NYU and revision of SSRN, 1354497*.

Pan, J., Polonik, W., and Yao, Q. (2010). Estimating factor models for multivariate volatilities: an innovation expansion method. In *Proceedings of COMPSTAT 2010*, pages 305–314. Springer.

Peña, D., Smucler, E., and Yohai, V. J. (2017). Forecasting multiple time series with one-sided dynamic principal components. *arXiv preprint arXiv:1708.04705*.

Peña, D. and Yohai, V. J. (2016). Generalized dynamic principal components. *Journal of the American Statistical Association*, 111(515):1121–1131.

Qiu, W. and Joe., H. (2015). *clusterGeneration: Random Cluster Generation (with Specified Degree of Separation).* R package version 1.3.4.

R Core Team (2017). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Rombouts, J. V. and Stentoft, L. (2011). Multivariate option pricing with time varying volatility and correlations. *Journal of Banking & Finance*, 35(9):2267–2281.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association*, 79(388):871–880.

Ryan, J. A. (2017). *quantmod: Quantitative Financial Modelling Framework.* R package version 0.4-8.

Santos, A. A. P. and Ferreira, A. R. (2017). On the choice of covariance specifications for portfolio selection problems. *Brazilian Review of Econometrics*, 37(1):89–122.

Santos, A. A. P. and Moura, G. V. (2014). Dynamic factor multivariate GARCH model. *Computational Statistics & Data Analysis*, 76:606–617.

Sortino, F. A. and van der Meer, R. (1991). Downside risk. *Journal of Portfolio Management*, 17(4):27–31.

Trucios, C. (2017). *StatPerMeCo: Statistical Performance Measures to Evaluate Covariance Matrix Estimates.* R package version 0.1.0.

Trucíos, C. and Hotta, L. K. (2016). Bootstrap prediction in univariate volatility models with leverage effect. *Mathematics and Computers in Simulation*, 120:91–103.

Trucíos, C., Hotta, L. K., and Ruiz, E. (2017a). Robust bootstrap densities for dynamic conditional correlations: implications for portfolio selection and value-at-risk. *Working Paper available at SSRN: https://ssrn.com/abstract=2969908.*

Trucíos, C., Hotta, L. K., and Ruiz, E. (2017b). Robust bootstrap forecast densities for GARCH returns and volatilities. *Journal of Statistical Computation and Simulation*, 87(16):3152–3174.

Vaz de Melo Mendes, B. and Pereira Câmara Leal, R. (2005). Robust multivariate modeling in finance. *International Journal of Managerial Finance*, 1(2):95–106.

Wang, Y. and Liu, L. (2016). Crude oil and world stock markets: volatility spillovers, dynamic correlations, and hedging. *Empirical Economics*, 50(4):1481–1509.

Welsch, R. E. and Zhou, X. (2007). Application of robust statistics to asset allocation models. *REVSTAT–Statistical Journal*, 5(1):97–114.

Zhu, K. and Ling, S. (2011). Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA–GARCH/IGARCH models. *The Annals of Statistics*, 39(4):2131–2163.

Zumbach, G. O. (2007). A gentle introduction to the RM2006 methodology. *Working Paper available at SSRN: https://ssrn.com/abstract=1420183*.