PERFORMANCE ANALYSIS OF MINIMUM VARIANCE ASSET ALLOCATION WITH HIGH FREQUENCY DATA

Jian Yu[*], *Francisco Rubio*[†], *Matthew R. McKay*[*]

[*]Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong
[†]Genetic Finance, Ltd., Hong Kong

ABSTRACT

We investigate the asset allocation optimization under the time-varying high frequency global minimum variance portfolio framework. The overall performance strongly relies on the estimate of the portfolio covariance matrix. However, for such applications, the sample size is often of similar order to the number of assets and in this case, the performance of the conventional covariance estimators are not very satisfactory. Additionally, the time variation effects will further amplify the estimation error and thus lead to inaccurate and high-risk investment decisions. In this paper, we propose to use the recently developed time variation adjusted realized covariance (TVARCV) estimator in a shrinkage structure, in order to address the above-mentioned problems. For this shrinkage TVARCV estimator, we provide a deterministic characterization of the portfolio realized risk in terms of the shrinkage parameter and the covariance matrix. At last, aiming to minimize the portfolio risk with respect to the shrinkage parameter, we also provide a consistent estimator for the realized variance, which depends only on the observable returns. Numerical results show that the proposed estimator is robust to time variation and has smaller portfolio risk.

1. INTRODUCTION

Portfolio optimization is well known and has been studied for many years since Markowitz's mean-variance optimization framework was proposed [1]. This framework for solving the asset allocation problem relies on an accurate estimate of the expectation and covariance matrix of asset returns. In general, estimates of the latter are more stable, and thus many works focus on the performance of the covariance estimation in the so-called global minimum variance portfolio (GMVP) framework [2]. Moreover, covariance estimation is essential to many other applications, such as statistical signal processing, financial engineering and related fields [3–5]. Note that an accurate covariance estimation usually requires a large set of data samples, which may range over years. Moreover, a constant covariance is assumed in many traditional estimators, such as the sample covariance matrix (SCM) estimator. However, the observation data often exhibits strong non-stationary effects due to seasonal time variation, which are well known to amplify the estimation errors and hence lead to high risk and inaccurate investment decisions [6].

To reduce the effects of the seasonal time variation, the usage of intraday *high frequency* data is often considered (see e.g. [7, 8]). In such case, the problem is to estimate the integrated covariance (ICV), which is the intraday instantaneous covariance integrated over a day. Effectively, this model implies that the covariance between the different returns is relatively constant over time, but the individual variances of the returns change in some unknown way. Therefore, we can consider a short enough history to accommodate changes in the covariance. The standard estimator for ICV is realized covariance (RCV) estimator, see [9] and the references therein for a review. The RCV estimator performs well if a very large number of samples are available [10]. However, similar to SCM estimator, it performs poorly when the sample size is of similar order to the number of assets. Therefore, developing improved estimator for such cases is an important problem. For the traditional daily (i.e. none high frequency) framework, improved estimators have been derived by results from random matrix theory in a double limit regime, where the sample size goes to infinity at the same rate as the number of assets, see [11–13].

Another important characteristic of the high frequency data is the presence of "intraday" time variation [10, 14–16]. Earlier techniques, such as [10, 15, 16], used either autoregressive conditional heteroskedasticity (ARCH) or generalized ARCH model. However the model restrains to a single asset or a small number of assets because of its complexity. The recent state of art estimator, time variation adjusted realized covariance (TVARCV) estimator in [14] reduces the intraday time variation. However, it still performs poorly in the condition when the sample size is of similar order to the number of assets.

We focus on the analysis and estimation of the realized variance (portfolio risk) in a time-varying high frequency GMVP framework. To improve the performance of RCV estimator and solve the problem of intraday time variation, we propose to apply TVARCV estimator instead of RCV estimator in a shrinkage structure and get our new James-Stein shrinkage TVARCV estimator, which is suitable for large number of assets. Furthermore, for this shrinkage TVARCV estimator, we provide a deterministic characterization of the random portfolio realized risk in terms of only the shrinkage parameter and the covariance matrix in a double limit regime. Moreover, to minimize the portfolio risk, we provide a consistent estimator for the realized variance with regard to the shrinkage parameter which only depends on the observable returns. Our estimator is robust to time variation and outperforms the other benchmark estimators in minimizing the portfolio risk.

2. MODEL AND PROBLEM FORMULATION

It is widely recognized that the variances of intraday returns show strong time variation [10, 15, 16]. In our work, we consider discretely observed time-varying vector process of intraday returns with N independent samples per day: $\mathbf{y}_k = w_k^{1/2}\mathbf{u}_k$, $k = 1, 2, \ldots, N$, where w_k is the unknown time-varying real coefficient, and $\mathbf{u}_k = \mathbf{\Sigma}^{1/2}\mathbf{z}_k \in \mathbb{R}^M$ is independent and identical distributed (i.i.d.) zero-mean random vector with positive definite covariance $\mathbf{\Sigma}$, but not nec-

The work of J. Yu and M. R. McKay was supported by the research grant, SBI11EG15.

essarily Gaussian distributed. $\mathbf{\Sigma}$ is presumed to remain fixed during a day with $\mathrm{tr}(\mathbf{\Sigma}) = M$. $\mathbf{\Theta}_k = w_k\mathbf{\Sigma}$ is the intraday instantaneous covariance. As $\mathbf{\Theta}_k$ varies over time, it cannot be directly used in applications such as asset allocation. Therefore, the ICV matrix, defined as $\mathbf{\Sigma}_{\mathrm{ICV}} = \sum_{k=1}^{N}\mathbf{\Theta}_k = \sum_{k=1}^{N} w_k\mathbf{\Sigma} = \mathrm{tr}(\mathbf{W}_N)\mathbf{\Sigma}$ is typically considered instead of the traditional covariance matrix, where $\mathbf{W}_N$ is a diagonal matrix with w_k on its diagonal. The trace of $\mathbf{W}_N$ is normalized to be 1 and the ICV over a day becomes equivalent to the covariance $\mathbf{\Sigma}$ [10, 15, 16]. Thus, we simply use $\mathbf{\Sigma}$ instead of $\mathbf{\Sigma}_{\mathrm{ICV}}$ to represent ICV in the time-varying GMVP framework.

We concentrate on the asset allocation problem in a time-varying high frequency GMVP framework [17] which can be mathematically formulated as the following quadratic optimization problem with linear constraints:

$$\begin{aligned} \min_{\mathbf{v}} \quad & \sigma_P^2(\mathbf{v}) = \mathbf{v}^T\mathbf{\Sigma}\mathbf{v} \\ \text{s.t.} \quad & \mathbf{v}^T\mathbf{1}_M = 1\,, \end{aligned} \tag{1}$$

where $\mathbf{v}$ denotes the vector of asset holdings in units of currency, $\mathbf{1}_M \in \mathbb{R}^M$ with all entries equal to 1. The solution of (1) is

$$\mathbf{v}_{\mathrm{GMVP}} = \frac{\mathbf{\Sigma}^{-1}\mathbf{1}_M}{\mathbf{1}_M^T\mathbf{\Sigma}^{-1}\mathbf{1}_M}\,, \tag{2}$$

giving the following minimum portfolio variance $\sigma_P^2(\mathbf{v}_{\mathrm{GMVP}}) = \frac{1}{\mathbf{1}_M^T\mathbf{\Sigma}^{-1}\mathbf{1}_M}$. This represents the minimum possible portfolio risk that may be achieved, provided that we know $\mathbf{\Sigma}$ exactly.

In practice we do not know $\mathbf{\Sigma}$, thus we form some estimate $\hat{\mathbf{\Sigma}}$. We denote by $\hat{\mathbf{v}}_{\mathrm{GMVP}}$ the sample portfolio chosen based on $\hat{\mathbf{\Sigma}}$. In practice, the quality of a portfolio rule $\hat{\mathbf{v}}_{\mathrm{GMVP}}$ based on a forecast or in-sample prediction of $\mathbf{\Sigma}$ can be measured by the out-of-sample or realized variance, which is a measure of the portfolio risk:

$$\sigma_P^2(\hat{\mathbf{v}}_{\mathrm{GMVP}}) = \hat{\mathbf{v}}_{\mathrm{GMVP}}^T\mathbf{\Sigma}\hat{\mathbf{v}}_{\mathrm{GMVP}} = \frac{\mathbf{1}_M^T\hat{\mathbf{\Sigma}}^{-1}\mathbf{\Sigma}\hat{\mathbf{\Sigma}}^{-1}\mathbf{1}_M}{\left(\mathbf{1}_M^T\hat{\mathbf{\Sigma}}^{-1}\mathbf{1}_M\right)^2}\,. \tag{3}$$

3. RCV PERFORMANCE AND TIME VARIATION EFFECT TEST

The most widely used estimator for ICV is the RCV estimator: $\hat{\mathbf{\Sigma}}^{\mathrm{RCV}} = \mathbf{\Sigma}^{1/2}\mathbf{Z}\mathbf{W}_N\mathbf{Z}^T\mathbf{\Sigma}^{1/2}$, $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_N]^T$. This can be seen as a weighted sample covariance with unknown time variation weighting $\mathbf{W}_N$. In this section, we will show two main issues when the RCV estimator is used on time-varying samples in a more practical condition, when the sample size N is of similar order to the number of assets M. The test results first show that lack of samples may lead to high bias or variance, and hence high estimation error. Moreover, the time variation effect will further amplify the estimation error.

We compare the accurate approximation of three RCV curves to theoretical minimum variance (MV) bound $\sigma_P^2(\mathbf{v}_{\mathrm{GMVP}})$ in a time-varying GMVP framework based on $\hat{\mathbf{\Sigma}}^{\mathrm{RCV}}$. The test is to show the poor performance of RCV estimator and the error amplification caused by time variation effect. The samples $\mathbf{y}_k$ are assumed to be the product of $w_k^{1/2}$ and a Gaussian random vector $\mathbf{u}_k$ with positive definite covariance $\mathbf{\Sigma}$. $[\mathbf{\Sigma}]_{i,j} = 0.9^{|i-j|}$, $1 \le i, j \le M$. M varies from 60 to 320 with $N = 2M$. All the simulations are repeated for 50 trials and the average empirical performance is reported. The RCV estimators are used on three series of return samples with different time variations but sharing the same $\mathbf{\Sigma}$. The samples of $\mathrm{RCV}_{\mathrm{none}}$ are assumed to be i.i.d. with no time variation, i.e. $\mathbf{y}_k = \frac{1}{\sqrt{N}}\mathbf{u}_k$. We can find in figure Fig. 1 that the curve of $\mathrm{RCV}_{\mathrm{none}}$, when used on i.i.d. samples, significantly deviates from the MV bound when N and M are of similar order.

Moreover, to describe the time variation, the entries of $\mathbf{\Theta}_k$ are assumed to approximate U-shape distribution during a day like in [14]. The diagonal entries of $\mathbf{W}_N$ are assumed to be piecewise constants:

$$w_k = \begin{cases} \frac{a}{N}, & k = 1, 2, \ldots, \frac{N}{4}, \\ \frac{b}{N}, & k = \frac{N}{4}+1, \ldots, \frac{3N}{4}, \\ \frac{a}{N}, & k = \frac{3N}{4}+1, \ldots, N, \end{cases} \tag{4}$$

where a and b are two constants with different values for each of the three curves, $a+b = 2$. For all the three curves, we have $\mathbf{\Sigma}_{\mathrm{ICV}} = \mathbf{\Sigma}$. For $\mathrm{RCV}_{\mathrm{small}}$ curve, $a = \frac{6}{4}$, $b = \frac{2}{4}$ in (4), and for $\mathrm{RCV}_{\mathrm{big}}$ curve, $a = \frac{7}{4}$, $b = \frac{1}{4}$. The samples of $\mathrm{RCV}_{\mathrm{big}}$ are more volatile (i.e., have more time variation) than those of $\mathrm{RCV}_{\mathrm{small}}$. In Fig. 1, we can see that the curve of $\mathrm{RCV}_{\mathrm{big}}$ deviates further from the MV bound than $\mathrm{RCV}_{\mathrm{small}}$ does. $\mathrm{RCV}_{\mathrm{none}}$ is closest to the MV bound. The figure shows when the return process is more volatile, the curve will be further away from the MV bound. The time variation is shown to further amplify the estimation errors.

When we do not have sufficiently large number of samples, the traditional estimators such as SCM and RCV, even used on i.i.d. returns, are likely to underestimate the portfolio risk, which lead to high risk and overly optimistic investment decisions. Time variation effects further amplify the estimation errors. These problems motivate us to provide an estimator which can improve the accuracy of RCV estimator and also solve the problem of time variation effect in the time-varying high frequency GMVP problem.

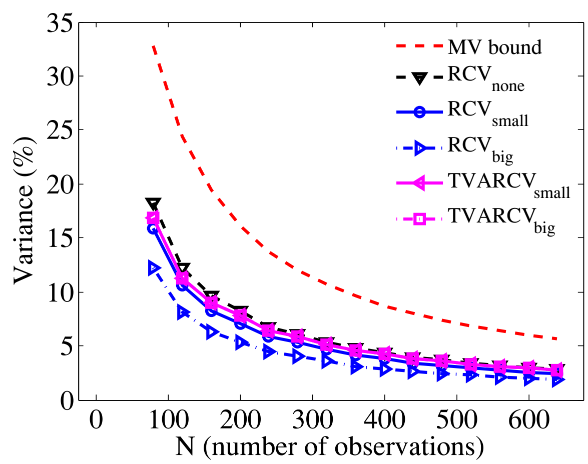

Fig. 1. Accurate approximation test in a time-varying GMVP framework using RCV, TVARCV

The TVARCV estimator in [14] is shown to be able to reduce time variation and is given as follows:

$$\hat{\mathbf{\Sigma}}^{\mathrm{TVARCV}} = \alpha_M \sum_{k=1}^{N} \frac{\mathbf{y}_k\mathbf{y}_k^T}{\|\mathbf{y}_k\|_2^2}\,, \tag{5}$$

where $\alpha_M = \mathrm{tr}(\hat{\mathbf{\Sigma}}^{\mathrm{RCV}})/N$.

The TVARCV estimator seeks to normalize the time variation by normalizing w_k of each sample. In Fig. 1, the samples of $\mathrm{TVARCV}_{\mathrm{small}}$ and $\mathrm{TVARCV}_{\mathrm{big}}$ curves use the same $\mathbf{W}_N$ as $\mathrm{RCV}_{\mathrm{small}}$ and $\mathrm{RCV}_{\mathrm{big}}$ respectively. What is interesting is that the curves of $\mathrm{TVARCV}_{\mathrm{small}}$ and $\mathrm{TVARCV}_{\mathrm{big}}$, although calculated based on samples with different time variation, are nearly the same. Moreover, we can find that when N is big enough (around 350), the TVARCV curves are nearly the same as the curve of $\mathrm{RCV}_{\mathrm{none}}$.

The TVARCV estimator can be used to reduce the time variation effect. However, lack of samples may lead to high bias or variance, therefore the curve of TVARCV estimator still significantly deviates from the MV bound when N and M are of similar order. In the next section, based on the TVARCV estimator, we will provide an estimator which will further improve the performance and significantly reduce the gap relative to the MV bound.

4. ASYMPTOTIC EQUIVALENCE AND CONSISTENT ESTIMATORS FOR TIME-VARYING GMVP FRAMEWORK

4.1. James-Stein shrinkage TVARCV estimator

We now provide an estimator which will further improve the performance of TVARCV estimator. The James-Stein shrinkage structure is widely used in many previous works to further improve the bias-variance tradeoff of the model [11, 12, 18]. It is often a combination of the SCM and the identity matrix. In the high frequency GMVP framework, [17] seeks to use a linear shrinkage of the RCV and the identity matrix to estimate $\mathbf{\Sigma}$. In this work, the need for addressing the time variation motivates the use of TVARCV instead of RCV in a shrinkage structure. Our proposed James-Stein shrinkage TVARCV estimator has the following structure:

$$\hat{\mathbf{\Sigma}}_{\mathrm{PSHR}} = (1-\rho_M)\alpha_M \sum_{k=1}^{N} \frac{\mathbf{y}_k \mathbf{y}_k^T}{\|\mathbf{y}_k\|_2^2} + \rho_M \mathbf{I}_M \,, \tag{6}$$

where ρ_M is the shrinkage parameter with $0 < \rho_M < 1$, $\mathbf{I}_M$ is the identity matrix.

The achieved realized variance based on $\hat{\mathbf{\Sigma}}_{\mathrm{PSHR}}$ in the time-varying GMVP framework:

$$\sigma_P^2(\hat{\mathbf{v}}_{\mathrm{GMVP}}) = \hat{\mathbf{v}}_{\mathrm{GMVP}}^T \mathbf{\Sigma} \hat{\mathbf{v}}_{\mathrm{GMVP}} = \frac{\mathbf{1}_M^T \hat{\mathbf{\Sigma}}_{\mathrm{PSHR}}^{-1} \mathbf{\Sigma} \hat{\mathbf{\Sigma}}_{\mathrm{PSHR}}^{-1} \mathbf{1}_M}{\left(\mathbf{1}_M^T \hat{\mathbf{\Sigma}}_{\mathrm{PSHR}}^{-1} \mathbf{1}_M\right)^2} \,, \tag{7}$$

which is random. The main challenge is to optimize the parameter ρ_M in (6), in order to minimize the realized variance in (7). To this end, we first provide a deterministic characterization of $\sigma_P^2(\hat{\mathbf{v}}_{\mathrm{GMVP}})$ in terms of $\mathbf{\Sigma}$ and ρ_M in a double limit regime. This gives a deterministic approximation for the true portfolio risk, and will be used subsequently to specify a consistent estimator for the optimal shrinkage parameter ρ_M.

4.2. Asymptotic equivalence of the time-varying GMVP realized variance

For our asymptotic analysis, we assume:

(**A1**) The double limit regime where $N, M \to \infty$ such that $0 < \liminf \frac{M}{N} \leq \limsup \frac{M}{N} < \infty$.

(**A2**) $\mathbf{\Sigma}$ has spectral norm bounded uniformly in M and N.

We now define

$$\xi_M := \frac{1}{N}\mathrm{tr}\left[\left(\mathbf{\Sigma}\left(\kappa_M \mathbf{\Sigma} + \rho_M \mathbf{I}_M\right)^{-1}\right)^2\right] \tag{8}$$
$$\tilde{\xi}_M := \kappa_M^2,$$

with κ_M the solution to [12]:

$$\kappa_M = \frac{1-\rho_M}{1+(1-\rho_M)\frac{1}{N}\mathrm{tr}\left[\mathbf{\Sigma}(\kappa_M\mathbf{\Sigma}+\rho_M\mathbf{I}_M)^{-1}\right]} \,. \tag{9}$$

Previous works [12, 13] have characterized the deterministic equivalence of the realized variance for i.i.d. samples in the traditional daily framework. However, these works depend on i.i.d. daily samples. In the time-varying high frequency framework, the asymptotic equivalence result does not only depend on $\mathbf{\Sigma}$, but also on how the time-variation process evolves [14]. As a result, when the time variation coefficients w_k are not constant and are unknown, results of [12, 13] do not directly apply. Therefore, we aim to derive an asymptotic equivalence result for the realized variance in the time-varying high frequency GMVP framework. To this end, the following technical lemma is important:

Lemma 1. *Under assumptions (A1) and (A2), the following asymptotic equivalence holds true:*

$$\mathrm{tr}\left[\mathbf{\Gamma}_M\left(\alpha_M\sum_{k=1}^{N}\frac{\mathbf{y}_k\mathbf{y}_k^T}{\|\mathbf{y}_k\|_2^2} - z\mathbf{I}_M\right)^{-1}\right] \asymp \mathrm{tr}\left[\mathbf{\Gamma}_M\left(\frac{\kappa_M}{1-\rho_M}\mathbf{\Sigma} - z\mathbf{I}_M\right)^{-1}\right] \,, \tag{10}$$

where $\asymp$ denotes asymptotic equivalence, $z = -\frac{\rho_M}{1-\rho_M}$, $\mathbf{\Gamma}_M = \mathbf{h}_M\mathbf{h}_M^H$, with $\mathbf{h}_M \in \mathbb{C}^M$ being an arbitrary nonrandom unit norm vector, and κ_M is given in (9).

The proof of *Lemma 1* involves the limiting empirical spectral distribution equivalence in [14] and the results in [12]. It is omitted due to space limitation.

Theorem 1. *(Time Variation Adjusted Asymptotic Deterministic Equivalence) Under assumptions (A1) and (A2), the following asymptotic equivalence holds true:*

$$\sigma_P^2(\hat{\mathbf{v}}_{\mathrm{GMVP}}) \asymp \frac{1}{1-\xi_M\tilde{\xi}_M} \cdot \frac{\mathbf{1}_M^T\left(\kappa_M\mathbf{\Sigma}+\rho_M\mathbf{I}_M\right)^{-1}\mathbf{\Sigma}\left(\kappa_M\mathbf{\Sigma}+\rho_M\mathbf{I}_M\right)^{-1}\mathbf{1}_M}{\left(\mathbf{1}_M^T\left(\kappa_M\mathbf{\Sigma}+\rho_M\mathbf{I}_M\right)^{-1}\mathbf{1}_M\right)^2} \,, \tag{11}$$

where ξ_M and $\tilde{\xi}_M$ are defined in (8), κ_M is calculated in (9).

Theorem 1 enables us to characterize $\sigma_P^2(\hat{\mathbf{v}}_{\mathrm{GMVP}})$ in terms of the only non-random variable $\mathbf{\Sigma}$ and ρ_M. Moreover, the result is robust to time variation and describes the true realized variance in the time varying framework. However, in reality, we do not know $\mathbf{\Sigma}$, we can not use this result to optimize ρ_M which minimizes the realized variance. Therefore, we aim to provide a consistent estimator for us to minimize the realized variance with regard to ρ_M.

4.3. Consistent estimators for minimizing the time-varying GMVP realized variance

Here we provide a time variation adjusted consistent estimator of the realized variance for us to optimize the shrinkage parameter ρ_M, which can be effectively used to minimize the realized variance in (7). To this end, we first denote the numerator of the right side of (7) as b_M, in which $\mathbf{\Sigma}$ in practice is unknown. In order to determine the optimal ρ_M, we need to obtain a consistent estimator for b_M, which is given in the following theorem.

Theorem 2. *(Time Variation Adjusted Generalized Consistent Estimator) Under assumptions (A1) and (A2), we have the following asymptotic equivalence*

$$b_M \asymp \frac{1}{(1-\zeta)^2} \mathbf{1}_M^T \hat{\mathbf{\Sigma}}_{\text{PSHR}}^{-1} \hat{\mathbf{\Sigma}}^{\text{TVARCV}} \hat{\mathbf{\Sigma}}_{\text{PSHR}}^{-1} \mathbf{1}_M \;, \tag{12}$$

where

$$\zeta = \frac{1}{N} \text{tr} \left[\hat{\mathbf{\Sigma}}^{\text{TVARCV}} \left(\hat{\mathbf{\Sigma}}^{\text{TVARCV}} + \frac{\rho_M}{1-\rho_M} \mathbf{I}_M \right)^{-1} \right].$$

Now substituting (12) into (7) yields

$$\sigma_P^2(\hat{\mathbf{v}}_{\text{GMVP}}) \asymp \frac{\mathbf{1}_M^T \hat{\mathbf{\Sigma}}_{\text{PSHR}}^{-1} \hat{\mathbf{\Sigma}}^{\text{TVARCV}} \hat{\mathbf{\Sigma}}_{\text{PSHR}}^{-1} \mathbf{1}_M}{(1-\zeta)^2 \left(\mathbf{1}_M^T \hat{\mathbf{\Sigma}}_{\text{PSHR}}^{-1} \mathbf{1}_M \right)^2} \,. \tag{13}$$

Therefore, the problem of obtaining the best asset allocation, as measured by the minimum realized variance, is reduced to optimizing (13) with regard to ρ_M. This can be done with a simple numerical search.

From (13), we see that the TVAGCE only depends on the *observable return samples* $\mathbf{y}_k$, and it is robust to intraday time variation. Moreover, the obtained estimator in (13) diminishes the time variation effect and further minimizes the realized variance. It is also important to note that compared with the generalized consistent estimator for the i.i.d. case in [12], we reduce the number of parameters from $N+1$ to 1, which significantly reduces the time of optimization.

5. NUMERICAL SIMULATIONS

In the simulation section, we demonstrated the advantages of our results in: (1) nicely approximating the MV bound in the in-sample (accurate approximation) test; (2) further minimizing the portfolio risk with respect to the shrinkage parameter ρ_M in the out-of-sample test in the time-varying high frequency GMVP framework. The samples are assumed to be the same as the previous test in Section 3.

In the first test, we compare the accurate approximation of five curves to the MV bound: (1) $\text{TVARCV}_{\text{big}}$ as the benchmark; (2) $\text{shrRCV}_{\text{small}}$, a shrinkage of the RCV estimator using the same data as $\text{RCV}_{\text{small}}$; (3) $\text{shrRCV}_{\text{big}}$, a shrinkage of the RCV estimator using the same data as RCV_{big}; (4) $\text{TVAGCE}_{\text{small}}$, the proposed method using the same data as $\text{RCV}_{\text{small}}$; (5) $\text{TVAGCE}_{\text{big}}$, the proposed method using the same data as RCV_{big}. We can find in Fig. 2 that the curves of $\text{shrRCV}_{\text{small}}$ and $\text{shrRCV}_{\text{big}}$ do not show much improvement. The curves of the proposed TVAGCE nicely approximate the MV bound. Moreover, the curves are above the MV bound, which solves the problem of risk underestimation which most traditional estimators have.

In the second test, we use the same five curves as in the first test and compare the abilities of each estimator in minimizing the realized portfolio risk. We consider a range of 30 days and assume the ICV matrix $\mathbf{\Sigma}$ does not change over the 30 days. The intraday time variation effects are assumed to be the same for each day. We use N intraday returns of the present day to predict the portfolio $\hat{\mathbf{v}}_{\text{GMVP}}$ of the next day. The portfolio risk is measured by the standard deviation of the out-of-sample returns. Figure 3 shows that our proposed TVAGCE is robust to time variation and has smaller portfolio risk than both the shrinkage RCV estimator and TVARCV estimator.

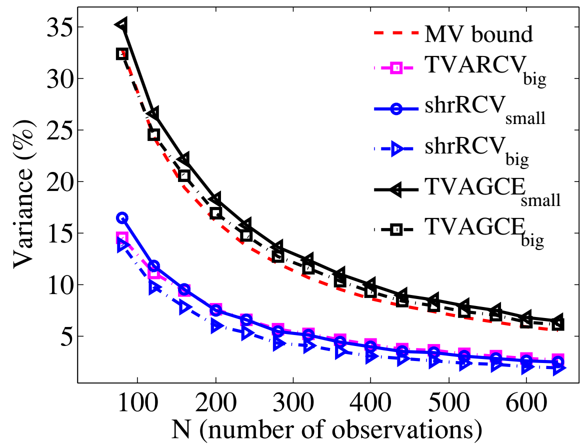

Fig. 2. Accurate approximation test of TVARCV estimator, shrinkage RCV estimator and the proposed TVAGCE in a time-varying GMVP framework.

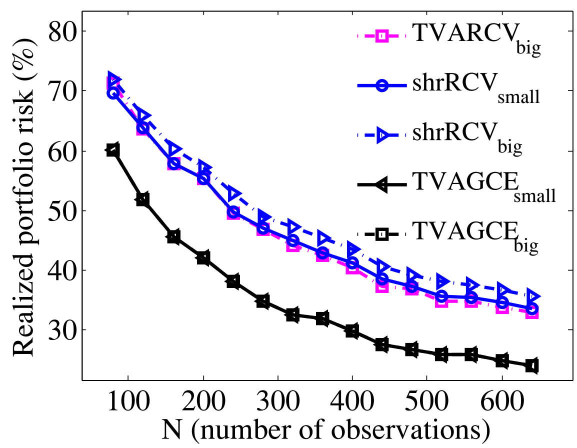

Fig. 3. Realized variance achieved over TVARCV estimator, shrinkage RCV estimator and the proposed TVAGCE in a time-varying GMVP framework.

6. REFERENCES

[1] H. Markowitz, "Portfolio selection," *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, Mar. 1952.

[2] R. Roll, "A mean/variance analysis of tracking error," *The Journal of Portfolio Management*, vol. 18, no. 4, pp. 13–22, Summer 1992.

[3] A. Wiesel and A. O. Hero, "Distributed covariance estimation in Gaussian graphical models," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 211–220, Jan. 2012.

[4] H. Li and H. V. Poor, "Large system spectral analysis of covariance matrix estimation," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 1395 –1422, Mar. 2009.

[5] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 62–75, May. 2010.

[6] D. Ruppert, *Statistics and data analysis for financial engineering*, Springer, 2010.

[7] T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys, "The distribution of realized exchange rate volatility," *Journal of the American Statistical Association*, vol. 96, pp. 42–55, Aug. 2001.

[8] T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys, "Modeling and forecasting realized volatility," *Econometrica*, vol. 71, no. 2, pp. 579–625, Mar. 2003.

[9] M. McAleer and M. C. Medeiros, "Realized volatility: A review," *Econometric Reviews*, vol. 27, no. 1-3, pp. 10–45, Mar. 2008.

[10] T. G. Andersen and T. Bollerslev, "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International Economic Review*, vol. 39, no. 4, pp. 885–905, Nov. 1998.

[11] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, no. 2, pp. 365–411, Feb. 2004.

[12] F. Rubio, X. Mestre, and D. P. Palomar, "Performance analysis and optimal selection of large minimum variance portfolios under estimation risk," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 4, pp. 337–350, Aug. 2012.

[13] M. Zhang, F. Rubio, and D. P. Palomar, "Improved calibration of high-dimensional precision matrices," *IEEE Transactions on Signal Processing*, vol. 61, no. 6, pp. 1509–1519, Mar. 2013.

[14] X. Zheng and Y. Li, "On the estimation of integrated covariance matrices of high dimensional diffusion processes," *The Annals of Statistics*, vol. 39, no. 6, pp. 3121–3151, 2012.

[15] J. Fleming, C. Kirby, and B. Ostdiek, "The economic value of volatility timing using realized volatility," *Journal of Financial Economics*, vol. 67, no. 3, pp. 473–509, Mar. 2003.

[16] V. Voev, *High Frequency Financial Econometrics*, pp. 293–312, Physica-Verlag HD, 2008.

[17] Q. Liu, "On portfolio optimization: How and when do we benefit from high-frequency data?," *Journal of Applied Econometrics*, vol. 24, no. 4, pp. 560–582, Mar. 2009.

[18] O. Ledoit and M. Wolf, "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection," *Journal of Empirical Finance*, vol. 10, no. 5, pp. 603–621, Dec. 2003.