Time Varying Estimation and Inference with Application to Large Dimensional Covariance Estimation and Portfolio Management

G. Kapetanios (with L. Giraitis, Y. Dendramis et al)

14th January 2015

Introduction

- Modelling structural change is crucial for most econometric analyses and especially forecasting
- There are a variety of different approaches to such modelling. These include:
- Structural breaks: Change in parameters is rare and abrupt.
- Smooth deterministic change over time with no abrupt changes.
- Random Coefficient (RC) models: These are the major focus of the current presentation.

The usual approach

- RC models have become increasingly popular recently.
- They are estimated by being cast in state space form and then using filters, such as the Kalman filter, usually as part of a Bayesian estimation framework.
- Work has ranged across topics such as accounting for the Great Moderation, documenting changes in the effect of monetary policy shocks and documenting changes in the degree of exchange rate pass-through.
- A selection of papers that make use of such models include Cogley and Sargent (2001), Cogley and Sargent (2005), Cogley, Sargent, and Primiceri (2010), Benati (2010), Benati and Surico (2008), Mumtaz and Surico (2009), Pesaran, Pettenuzzo, and Timmermann (2006) and Koop and Potter (2007).
- Problems with these models: Unclear theoretical properties, heavy computational costs.

A new approach

- It is clear that easier, robust and less costly estimation methods with clear theoretical properties would be welcome.
- Kernel estimation of coefficient processes in models of smooth deterministic change is well established and fully analysed in the statistical literature. In the context of structural change this approach is simply a refinement of estimating models with rolling windows.
- Giraitis, Kapetanios and Yates (JoE, 2014) propose applying kernel estimation to RC models.
- They formalise a kernel estimator for the unobserved coefficient process and derive its theoretical properties.
- These estimator properties are very attractive: Consistency and asymptotic normality. Also great small sample properties.
- A number of follow-up papers take this approach to more complex and realistic settings.

Presentation Roadmap

- We will present the basic idea within the context of a simple AR model and discuss theoretical properties.
- Then, we move on to more realistic models that allow for time varying variances.
- Third, we discuss time varying Maximum Likelihood estimation and its theoretical properties coupled with an application to the modelling of the sterling money market.
- We consider the issue of selecting relevant tuning parameters using data
- The above allow us to develop time varying estimation for large dimensional covariance matrices and present a basic result that can be used for a variety of different covariance estimators
- Finally we will present some Monte Carlo evidence on these estimators which can feed directly into the construction of portfolios of assets.

Some preliminaries

- For simplicity of analysis we begin with a univariate dynamic model. We then extend in many directions.
- We consider the AR(1) model:

$$y_t = \rho_{t-1} y_{t-1} + u_t, \quad t = 1, 2, \cdots, n, \qquad u_t \sim IID(0, \sigma_u^2)$$

- AR is a workhorse class of models. Many possibilities depending on how $\rho_{n,t-1}$ is specified.
- The most closely related specification relates to Locally stationary models: Priestley (1965), Dahlhaus (1997)

$$\rho_{n,t-1} = \mu(t/n), \quad 1 \leq t \leq n \quad \text{deterministic, smooth}$$

Random ρ_{t-1}

- Random coefficient (RC) case: ρ_t random process and bounded between -1 and 1. Many ways to bound. This issue is not discussed in the macro/finance literature.
- We choose a straightforward standartization

$$\rho_t = \rho \frac{a_t}{\max_{0 \le k \le t} |a_k|}, \quad t = 1, 2, \cdots, n,$$

$\{a_t\}$ determines random drift. ρ restricts ρ_t away from -1 and 1. Both $\{a_t\}$, ρ unknown. Observe: $\rho_k \in [-\rho, \rho] \subset (-1, 1)$, for all $k = 1, \cdots, n$.
- a_t evolves as $I(d)$, $d > 1/2$ process: $\{v_t\}$ stationary

$$a_t = a_{t-1} + v_t, \quad t = 1, \cdots, n,$$

- Popular choice: v_t i.i.d., a_t driftless random walk,

Estimation and Inference

- We wish to estimate the coefficients $\rho_1, \cdots, \rho_n$
- We suggest as an estimator a weighted sample autocorrelation at lag 1 given by

$$\hat{\rho}_{n,t} := \frac{\sum_{k=1}^{n} K(\frac{t-k}{H}) y_k y_{k-1}}{\sum_{k=1}^{n} K(\frac{t-k}{H}) y_{k-1}^2}, \tag{1}$$

where $K(x) \geq 0$, $x \in \mathbb{R}$ is a continuous bounded kernel function.

- This is simply a generalisation of a rolling window estimator given by

$$\hat{\rho}_{n,t} := \frac{\sum_{k=t-H}^{t+H} y_k y_{k-1}}{\sum_{k=t-H}^{t+H} y_{k-1}^2},$$

which is a local sample correlation of y_t's at lag 1, based on $2H + 1$ observations $y_{t-H}, \cdots, y_{t+H}$.

- For the bandwidth we assume that $H \to \infty$ and $H = o(n)$.

Estimator properties

- This estimator is consistent and asymptotic normal as summarised below

Theorem 3

$$\hat{\rho}_{n,t} - \rho_t = \xi_{n,t} + O_P((H/n)^\gamma) = O_P(1/\sqrt{H}) + O_P((H/n)^\gamma),$$

$$\frac{T_{H,t}}{(1-\rho_t^2)^{1/2}}\xi_{n,t} \to_D N(0,1).$$

(ii) If $H = o(n^{\gamma/(0.5+\gamma)})$, then

$$\frac{T_{H,t}}{\sqrt{1-\hat{\rho}_{n,t}^2}}\Big(\hat{\rho}_{n,t} - \rho_t\Big) \to_D N(0,1).$$

where $\gamma = d - 1/2$ and

$$b_{tk} := K(\frac{t-k}{H}),\ T_{H,t} := \frac{\sum_{k=1}^n b_{tk}}{\left(\sum_{k=1}^n b_{tk}^2\right)^{1/2}},\ \xi_{n,t} := \frac{\sum_{k=1}^n b_{tk} u_k y_{k-1}}{\sum_{k=1}^n b_{tk} y_{k-1}^2}.$$

Some comments

- For $\gamma \geq 1/2$, we can take $H = o(n^{1/2})$ and then no knowledge of γ is needed. Using the above the estimation of standard errors is easily feasible.
- Estimator requires persistence of $\rho_{n,t}$
- $0 < \gamma < 1$ defines the magnitude of error term in normal approximation:
- Larger $\gamma \to$ stronger persistence $\to$ better approximation
- $\{v_j\}$ in $a_t = a_{t-1} + v_t$ can have short, long or negative memory.
- Bandwidth $H = o(n^{1/2})$ yields negligible error for short memory $\{v_j\}$ $(\gamma = 1/2)$ and long memory v_j $(1/2 < \gamma < 1)$.
- When $\gamma \to 0$, trending of a_t and the quality of approximation deteriorate.
- For stationary $\{a_t\}$, estimation is not consistent.

Further Extensions

- The above focuses on a simple univariate homoscedastic model.
- Good for proving the concept but not for actual empirical work
- Giratis, Kapetanios and Yates (2014) provide further extensions that enable analysing more realistic models
- The first extension is to heteroscedastic vector autoregressive models (VARs)
- We can show that kernel estimates are consistent under conditions that bound the norm of the coefficient matrices.
- The second crucial extension enables estimation of possible variation in the unconditional variances of the model shocks.

The setup

- We consider the m-dimensional VAR(1) model:

$$y_t = \mathbf{\Psi}_{t-1} y_{t-1} + u_t, \quad t = 1, 2, \cdots, n, \qquad u_t \sim IID(0, \Sigma_u)$$

- VAR is a workhorse class of models. Many possibilities depending on how $\mathbf{\Psi}_{t-1}$ is specified.
- The most closely related specification again relates to Locally stationary models: Priestley (1965), Dahlhaus (1997)

$$\mathbf{\Psi}_{n,t-1} = \mu(t/n), \quad 1 \leq t \leq n \quad \text{deterministic, smooth}$$

Random Ψ_{t-1}

- Random coefficient (RC) case: $\mathbf{\Psi}_t$ random process with eigenvalues bounded between -1 and 1. Many ways to bound. Let $\mathbf{\Psi}_{t-1} = [\psi_{t-1,ij}]$. Sufficient bounding can be implemented by defining

$$\tilde{\mathbf{\Psi}}_{t-1} = [\tilde{\psi}_{t-1,ij}], \qquad \tilde{\psi}_{t,ij} = \tilde{\psi}_{t-1,ij} + v_{\psi t,ij}, \quad t = 1, \cdots, n; \; i, j = 1,$$

  where $v_{\psi t,ij}$ is a zero mean i.i.d. sequence with finite variance. Then,

$$\psi_{t-1,ij} = \psi \frac{\tilde{\psi}_{t-1,ij}}{\max_{0 \le k \le t} \sum_{j=1}^{m} |\tilde{\psi}_{t-1,ij}|}, \quad t = 1, 2, \cdots, n,$$

  and $0 < \psi < 1$. This ensures that the maximum eigenvalue of $\mathbf{\Psi}_{n,t-1}$ is bounded above by one in absolute value.
- But this is just an example. What matters is that

$$||\mathbf{\Psi}_t - \mathbf{\Psi}_{t+h}|| = O_p(h/t)$$

  This kind of condition seems necessary for consistency.

Estimation

- We wish to estimate the coefficients $\Psi_1, \cdots, \Psi_n$
- We suggest as an estimator a weighted sample autocorrelation at lag 1 given by

$$\hat{\mathbf{\Psi}}_{n,t} := \left(\sum_{k=1}^{n} K(\frac{t-k}{H_\psi}) y_k y_{k-1}' \right) \left(\sum_{k=1}^{n} K(\frac{t-k}{H_\Psi}) y_{k-1} y_{k-1}' \right)^{-1}, \tag{2}$$

  where $K(x) \geq 0$, $x \in \mathbb{R}$ is a continuous bounded function.
- This is simply a generalisation of a rolling window estimator given by

$$\hat{\mathbf{\Psi}}_{n,t} := \left(\sum_{k=t-H_\psi}^{t+H_\psi} y_k y_{k-1}' \right) \left(\sum_{k=t-H_\psi}^{t+H_\psi} y_{k-1} y_{k-1}' \right)^{-1},$$

  which is a local sample correlation of y_t's at lag 1, based on $2H_\psi + 1$ observations $y_{t-H_\psi}, \cdots, y_{t+H_\psi}$.
- For the bandwidth we assume that $H_\psi \to \infty$ and $H_\psi = o(n)$

Estimation of time varying variance

- We wish to also allow a time varying variance for u_t and estimate it
- $$\mathbf{u}_t = \mathbf{H}_{t-1}\varepsilon_t, \qquad E[\mathbf{u}_t|\mathcal{F}_{t-1}] = \mathbf{0} \tag{3}$$

  with respect to some filtration $\mathcal{F}_t$, where $\mathbf{H}_t = \{h_{t,ij}\}$ is a $m \times m$ time varying random volatility process, and ε_t is a vector-valued standardized i.i.d. noise, $E\varepsilon_t = 0$, $E\varepsilon_t\varepsilon_t' = \mathbf{I}$. Denote by $\mathbf{\Sigma}_t = \mathbf{H}_{t-1}\mathbf{H}_{t-1}' = E[\mathbf{u}_t\mathbf{u}_t'|\mathcal{F}_{t-1}]$ the conditional variance-covariance matrix.
- We then obtain the residuals, $\hat{u}_t$ from the conditional mean model, and fit a time varying simple location model to every element of $\hat{u}_t\hat{u}_t'$. We denote the bandwidth parameter by H_h.

Estimator properties

Let $\kappa_{n,\psi} := (\bar{H}_\psi/n)^{1/2} + H_\psi^{-1/2}, \kappa_{n,h} := (\bar{H}_h/n)^{1/2} + H_h^{-1/2}$, $\mathrm{k}_{tj} := K\left((t-j)/H_\psi\right), K_t = \sum_{j=1}^{n} k_{tj}, K_{2,t} = \sum_{j=1}^{n} k_{tj}^2$. For $H_\psi = o(n/\log n),\ H_h = o(n/\log n)$,

$$\hat{\boldsymbol{\Psi}}_t - \boldsymbol{\Psi}_t = O_p(\kappa_{n,\psi}), \tag{4}$$

$$\boldsymbol{\Sigma}_{\hat{u}\hat{u},t} - \boldsymbol{\Sigma}_t = O_p\left(\kappa_{n,\psi}^2 + \kappa_{n,h}\right). \tag{5}$$

In addition, if $H_\psi \bar{H}_\psi = o(n)$, then for any real $m \times 1$- vector a such that $||a|| = 1$,

$$(K_t/K_{2,t})^{1/2}\mathbf{H}_{t-1}^{-1}(\widehat{\boldsymbol{\Psi}}_t - \boldsymbol{\Psi}_t)\Big(\sum_{j=1}^{n} k_{tj}\mathbf{y}_{j-1}\mathbf{y}_{j-1}'\Big)^{1/2} a \to_D \mathcal{N}(0, \mathbf{I}) \tag{6}$$

In addition, if $H_h\bar{H}_h = o(n)$ and $H_h^{1/2} << H_\psi << n/(H_h \log n)^{1/2}$, then

$$(L_t/L_{2,t}^{1/2})\mathbf{H}_{t-1}^{-1}(\boldsymbol{\Sigma}_{\hat{u}\hat{u},t} - \boldsymbol{\Sigma}_t)\mathbf{H}_{t-1}'^{\,-1} \to_D \mathbf{Z} \tag{7}$$

where the elements of $\mathbf{Z}$ are independent normal variables.

Time Varying Maximum Likelihood Estimation and an application to Bank Lending

- Giraitis, Kapetanios, Wetherilt and Zikes (2013) have recently extended the above method to handle time varying Maximum Likelihood estimation with random coefficients.
- They use a set of interdependent bivariate Tobit models to model interactions between bank lending of a set of banks over time in the sterling money market. Their model can handle parsimoniously a large set of variables while allowing for structural change.
- They consider the Tobit model

$$y_t = \begin{cases} \beta_{0,t}' x_t + u_t, & \text{if } \beta_{0,t}' x_t + u_t > 0 \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

where the latent variable $\beta_{0,t}' x_t + u_t$ is defined by a vector $\beta_{0,t}$, a $k \times 1$ vector of known random/deterministic regressors x_t and the noise $u_t = \sigma_{0,t}\varepsilon_t, \qquad \varepsilon_t \sim NID(0,1)$.

Time Varying Maximum Likelihood Estimation

- They assume that $\beta_{0,t}$ and $\sigma_{0,t}$ are bounded (truncated) random/deterministic processes independent of ε_t satisfying the following smoothness condition: for $1 \leq h \leq t$, as $h \to \infty$,

$$\sup_{j:|j-t|\leq h} ||\theta_{0,t} - \theta_{0,j}||^2 = O_p\left(h/t\right). \tag{9}$$

Time Varying Maximum Likelihood Estimation

- To accommodate estimation of time varying parameter $\theta_{0,t}$, they use the weighted likelihood function

$$L_{\theta,t,T} := \prod_j{}' (1 - F_{\theta,j})^{k_{tj}} \prod_j{}'' (g_{\theta,t}(y_j))^{k_{tj}},$$

with the weights $k_{tj} = \tilde{k}_{tj} / \left(\sum_{j=1}^{T} \tilde{k}_{tj} \right)$, $\tilde{k}_{tj} := K\big((t-j)/H\big)$, where $K(x) \geq 0$, $x \in \mathbb{R}$ is a continuous bounded function and the bandwidth parameter $H \to \infty$, $H = o(T/\log T)$.

- They define the MLE estimate $\hat{\theta}_t$ of $\theta_{0,t}$ as the maximiser of the weighted log-likelihood:

$$Q_{\theta,t,T} := \log L_{\theta,t,T} = \sum_j{}' k_{tj} \log(1 - F_{\theta,j}) + \sum_j{}'' k_{tj} \log g_{\theta,t}(y_j),$$

$$\hat{\theta}_t := \mathrm{argmax}_\theta Q_{\theta,t,T}. \tag{10}$$

Theorem

Let $y_1, ..., y_T$ be a sample of the Tobit model and $t = [\tau T]$ where $0 < \tau < 1$ is fixed. Denote $\kappa_{H,T} := (\bar{H}/T)^{1/2} + H^{-1/2}$. Then the MLE estimate $\hat{\theta}_t$ of the parameter $\theta_{0,t}$ has the following properties.

(i) *(Consistency). There exist an open neighborhood B_t of $\theta_{0,t}$ such that*

$$\hat{\theta}_t := \mathrm{argmin}_{\theta \in B_t} Q_{\theta,t,T} \to_P \theta_{0,t}, \text{ and } \hat{\theta}_t - \theta_{0,t} = O_p(\kappa_{H,T}).$$

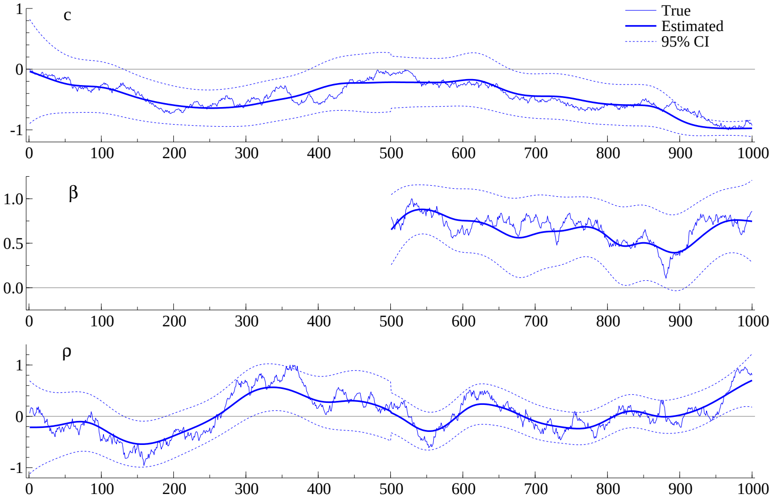
(ii) *(Asymptotic normality). In addition, if $\bar{H}H = o(T)$, then*

$$\Sigma_t^{-1/2}(\hat{\theta}_t - \theta_{0,t}) \to_D \mathcal{N}(0, I), \qquad \Sigma_t := K_{t,T}\Big(-\frac{\partial^2 Q_{\theta_{0,t},t}}{\partial\theta\partial\theta'}\Big)^{-1}$$

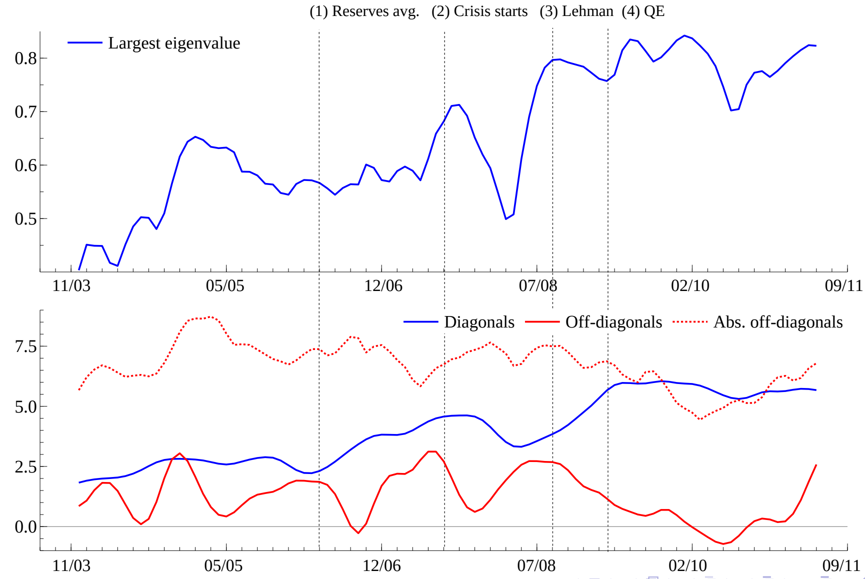
where $K_{t,T} := \sum_{j=1}^{T} k_{tj}^2 \sim const\ H^{-1}$, and $\big(-\frac{\partial^2 Q_{\theta_{0,t},t}}{\partial\theta\partial\theta'}\big)$ is a positive definite matrix.

Simulation

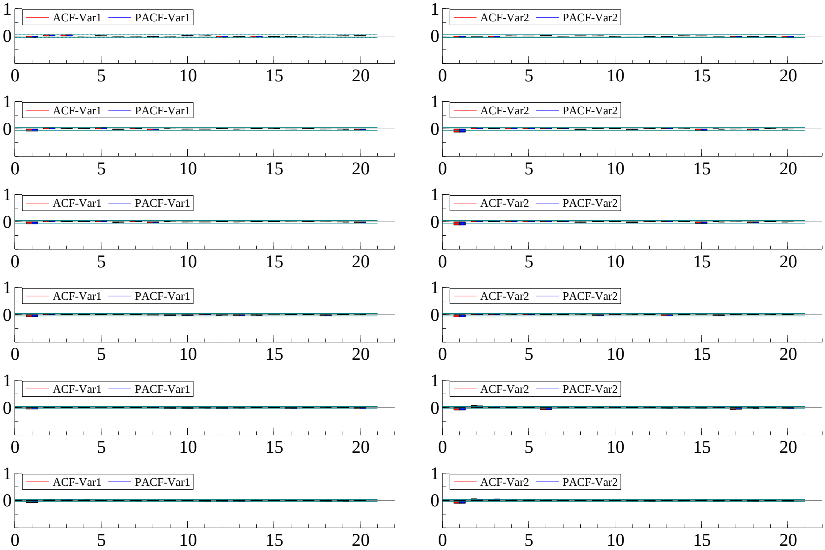
$$y_t = c_t + \rho_t y_{t-1} + \beta_t x_t + \epsilon_t$$



Summaries of VAR coefficient matrix



Diagnostics - generalized residuals



A diversion: How to get tuning parameters

- In all the above the bandwidth was taken as given and a choice based on random coefficients and MSE optimality was suggested.
- But one may wish to get a data dependent choice.
- Further, there may be other tuning parameters. For example, tuning parameters associated with the estimation of large covariance matrices.
- Giraitis, Kapetanios and Price (JoE, 2013) have addressed this in the context of forecasting under structural change.

A diversion: How to get tuning parameters

- They derive cross-validation approaches to determine the rate at which data should be downweighted when forecasting in the presence of structural change.
- They prove that one can retrieve the optimal bandwidth that minimises the MSE of forecasts under a variety of structural change settings.
- And show that such a strategy can be very effective for forecasting US spreads and price indices.
- We conjecture that such a strategy would work for determining more than one tuning parameters simultaneously.

- We can use the above methods to investigate the estimation of large dimensional covariance matrices in the presence of structural change.
- Let

$$\mathbf{y}_t = \mathbf{H}_{t-1}\boldsymbol{\varepsilon}_t, \qquad E[\mathbf{u}_t|\mathcal{F}_{t-1}] = \mathbf{0} \tag{11}$$

with respect to some filtration $\mathcal{F}_t$, where $\mathbf{H}_t = \{h_{t,ij}\}$ is an $p \times p$ time varying volatility process, and $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, ..., \varepsilon_{p,t})'$ is a vector-valued standardized α-mixing process, such that $E\boldsymbol{\varepsilon}_t = 0$, $E\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' = \mathbf{I}$.

- Denote by $\boldsymbol{\Sigma}_t = [\sigma_{ij,t}] = \mathbf{H}_{t-1}\mathbf{H}_{t-1}' = E[\mathbf{u}_t\mathbf{u}_t'|\mathcal{F}_{t-1}]$ the conditional variance-covariance matrix, where $\mathbf{H}_t$ is $\mathcal{F}_t$-measurable with respect to some filtration $\mathcal{F}_t$.

- We assume

$$h_{t,ij}^{4} \leq C, \quad \text{for} 1 \leq k \leq t/2 \tag{12}$$

Further, for $1 \leq k \leq t/2$ and some $\vartheta > 0$,

$$\max_{1 \leq i,j \leq N} \max_{1 \leq s \leq k} \left| \sigma_{ij,t} - \sigma_{ij,t+s} \right|^2 = O_{a.s.}(d(N)\,(k/t)^{\vartheta}). \tag{13}$$

- Assume further that

$$\sup_{i} \Pr\left[|\varepsilon_{i,t}| > a \right] \leq C_1 e^{-C_2 a^q},\ q > 1 \tag{14}$$

- Define

$$\hat{\boldsymbol{\Sigma}}_t = [\hat{\sigma}_{ij,t}] = L_t^{-1} \sum_{j=1}^{n} l_{tj} \mathbf{y}_j \mathbf{y}_j', \qquad l_{tj} := L(\frac{t-j}{H}), \quad L_t := \sum_{j=1}^{n} l_{tj}, \tag{15}$$

- Then we have the following theorem

Theorem

Under (12)-(14), and for any $0 < \delta < 1$, *we have that*

$$\max_{ij} |\hat{\sigma}_{ij,t} - \sigma_{ij,t}| = O_p\left(\frac{(\log p)^{(q+1)/q}}{H^{1-(1+\delta)/2}}\right) \tag{16}$$

- (16) is obtained under weaker than usual assumptions. We need to allow for heterogeneity over time and we also allow for mixing whereas usually an iid assumption is made. As a result of our weaker assumptions as well as time variation a slower rate that usual is obtained. See Theorems 3.3 and 3.4 of White and Wooldridge (1991).
- The full sample version of (16) forms the core of a variety of results on estimation of $\mathbf{\Sigma}$ in the case of full sample estimation. So (16) can be used to replicate all these results for the time varying case with minimal further modifications.

Time Varying Estimation of Large Dimensional Covariances

- Examples of full sample estimation include the work of Ledoit and Wolf on shrinkage estimators of Σ, and the work of Bickel and Levina and Cai and Liu on threshold estimators of Σ.
- It is worth mentioning that none of the above allow for non-iid data in their derivations. However, the mixing assumption we make is not novel for the full sample estimation since Fan, Liao and Micheva (2013) make it, although they base their work on an exponential inequality in a recent working paper and do not seem to be aware of either White and Wooldridge (1991) or the work on exponential inequalities cited therein.

Ledoit and Wolf (2003, 2004)

- Relies on an asymptotically optimal shrinkage approach

$$\widehat{\Sigma}_{LW} = \rho_1 \widehat{\Sigma}_{t\arg et} + \rho_2 \widehat{\Sigma} \tag{17}$$

$\widehat{\Sigma}_{t\arg et}$:variance target estimator, $\widehat{\Sigma}$:full sample covariance estimator, ρ_1, ρ_2, μ are positive constants. When the $\Sigma_{t\arg et}$ is the identity ($\Sigma_{t\arg et} = I$) the optimal weights are given by:$\widehat{\rho}_1 = m_T b_T^2 / d_T^2$, $\widehat{\rho}_2 = a_T^2 / d_T^2$, $m_T = N^{-1} tr\left(\widehat{\Sigma}\right)$, $d_T^2 = N^{-1} tr\left(\widehat{\Sigma}^2\right) - m_T^2$, $a_T^2 = d_T^2 - b_T^2$, $b_T^2 = \min\left(\overline{b}^2, d_T^2\right)$ and $\overline{b}_T^2 = \frac{1}{NT^2} \sum_{t=1}^{T} \left(\sum_{i=1}^{N} \left(x_{it} - \overline{x}_i\right)^2\right) - \frac{1}{NT} tr\left(\widehat{\Sigma}^2\right)$

- One could use several candidates for the $\widehat{\Sigma}_{t\arg et}$, such as the diagonal of $\widehat{\Sigma}$

$$\widehat{\Sigma}_{LW} = \rho_1 diag\left(\widehat{\Sigma}\right) + \rho_2 \widehat{\Sigma} \tag{18}$$

Cai and Liu (2013)

The adaptive thresholding estimator:

$$\widehat{\Sigma}_{CL}(\delta) = (\widetilde{\sigma}_{ij})_{N\times N} \text{ with } \widetilde{\sigma}_{ij} = s_{\lambda_{ij}}(\widehat{\sigma}_{ij}), \tag{19}$$

$\widehat{\sigma}_{ij}$ is the i, j element of the full sample covariance matrix estimate $\widehat{\Sigma}$

$s_{\lambda_{ij}}$ is a thresholding rule (hard, soft, adaptive lasso)

$$\lambda_{ij} := \lambda_{ij}(\delta) = \delta\sqrt{\frac{\widehat{\theta}_{ij}\log N}{T}}$$

and

$$\widehat{\theta}_{ij} = \frac{1}{T}\sum_{t=1}^{T}\left[(x_{it} - \overline{x}_i)(x_{jt} - \overline{x}_j) - \widehat{\sigma}_{ij}\right]^2 \tag{20}$$

δ is a regularization parameter that can be fixed at $\delta = 2$ or chosen through cross validation

A theorem on the time varying Cai-Liu estimator

Theorem

Let $\{\mathbf{\Sigma}_t\}_{t=1}^{\infty}$ *belong to* $\mathcal{U}(q, c_0(N), M)$ *and let our Assumptions hold. Let* $\lambda_{ij,t} = \kappa\hat{\theta}_{ij,t}^{1/2}\frac{(\log H^{\varkappa})^{1/2}(\log N)^{(q+1)/q}}{\bar{H}^{1-(1+\delta)/2}}$, *for any* $0 < \delta < 1$, *some finite* κ, *where* $\bar{H} = H(\log H^{\varkappa})^{1/2}$ *for some* $\varkappa > 1$. *Let* $\bar{H} = o\left(d(N)^{-\frac{1}{\vartheta+1}} T^{\frac{\vartheta}{\vartheta+1}}\right)$. *Then,*

$$\left\| T_{\lambda_{ij}}\left(\hat{\Sigma}_t\right) - \Sigma_t \right\| = O_p\left(c_0(N)\frac{(\log H^{\varkappa})^{1/2}(\log N)^{(q+1)/q}}{\bar{H}^{1-(1+\delta)/2}}\right), \text{ for all } t. \quad (21)$$

Further, let $\{\mathbf{\Sigma}_t\}_{t=1}^{\infty}$ *belong to* $\mathcal{U}(q, c_0(N), M, \epsilon)$, *for some* $\epsilon > 0$. *Then, for any* $0 < \delta < 1$,

$$\left\| T_{\lambda_{ij}}\left(\hat{\Sigma}_t\right)^{-1} - \Sigma_t^{-1} \right\| = O_p\left(c_0(N)\frac{(\log H^{\varkappa})^{1/2}(\log N)^{(q+1)/q}}{\bar{H}^{1-(1+\delta)/2}}\right), \text{ for all} \quad (22)$$

Relaxing our assumptions

The Theorem holds under restrictive conditions. We impose both mixing and exponentially declining tails. We can relax both at the cost of lower rates. The following Corollary gives this result.

Corollary

Let $\{\mathbf{\Sigma}_t\}_{t=1}^{\infty}$ *belong to* $\mathcal{U}\left(q, c_0(N), M\right)$ *and let our Assumptions hold. Let* ε_t *have finite eighth moments. Let* $\lambda_{ij,t} = \kappa \hat{\theta}_{ij,t}^{1/2} N H^{-\eta}$*, for any* $0 < \eta < 1/2$*, and some finite* κ*. Let* $\bar{H} = H\left(\log H^{\varkappa}\right)^{1/2}$ *for some* $\varkappa > 1$*, and* $\bar{H} = o\left(d(N)^{-\frac{1}{2\eta+\vartheta}} T^{\frac{\vartheta}{2\eta+\vartheta}} N^{\frac{2}{2\eta+\vartheta}}\right)$*. Then, for all* $\eta < 1/2$*,*

$$\left\| T_{\lambda_{ij}}\left(\hat{\Sigma}_t\right) - \Sigma_t \right\| = o_p\left(c_0(N) N H^{-\eta}\right), \text{ for all } t. \quad (23)$$

Evaluation Exercises

- We carry out a Monte Carlo and an empirical exercise
- The Monte Carlo considers the estimation of the time varying covariance matrix both within the sample and at the end of the sample.
- The empirical exercise focuses on the construction of minimum variance portfolios and considers their out-of-sample performance.

Minimum Variance portfolio

- The global minimum variance (GMV) portfolio is the portfolio which is designed to minimize investors exposure to risk

$$w_t := w\left(\widehat{\Sigma}_t\right) = \frac{\widehat{\Sigma}_t^{-1} 1_N}{1_N' \widehat{\Sigma}_t^{-1} 1_N}$$

- Cross Validation is designed to minimize the out of sample variance of the portfolio.

Cross Validation Objective Functions

To determine to bandwidth (h), shrinkage (ρ), and thresholding (δ) coefficients, we rely on cross validation.
Given a sample of size T, of N returns $\{R_t\}_{t=1}^{T}$, we select the parameter of interest ($\theta = (h, \rho, \delta)$) by numerically minimizing the objective function, in the sample. The objective function could be the portfolio variance

$$Q_{T,\theta}^{ov} := \frac{1}{T_n} \sum_{\tau=T_0}^{T} \left(w_{\theta,\tau|\tau-1} R_\tau - \frac{1}{T_n} \sum_{\tau=T_0}^{T} w_{\theta,\tau|\tau-1} R_\tau \right)^2 \quad (24)$$

where $w_{\theta,\tau|\tau-1}$ is the GMV portfolio computed using the data available up to time τ-1, $T_0 = o(T)$ and $T_n := T - T_0 + 1$. Or it could be the MSE of the covariance matrix estimator

$$Q_{T,\theta}^{oc} := \frac{1}{T_n} \sum_{\tau=T_0}^{T} ||\widehat{\Sigma}_{\theta,-\tau} - (R_\tau - \hat{\mu}_\tau)' (R_\tau - \hat{\mu}_\tau)|| \quad (25)$$

where $\widehat{\Sigma}_{\theta,-\tau}$ obtained by dropping data at time τ.

Time Varying Estimation of Large Dimensional Covariances: Monte Carlo

- To generate Σ_t we do the following
- First, we draw T $N \times 1$ vectors $\mathbf{b}_t = (b_{1t}, b_{2t}, ..., b_{Nt})'$ where for $i \leq N_b (< N)$, $N_b = \left[N^{d_b}\right]$,

$$b_{it} = \left(\frac{2.5\tilde{b}_{it}}{\max_{1 \leq j \leq t} \tilde{b}_{ij}} \right) + 2.5$$

where $\tilde{b}_{it} = \tilde{b}_{it-1} + \xi_{it}$ and $\xi_{it} \sim niid(0.1, 1)$. For $i > N_b$, we set $b_{it} = 0$.

Time Varying Estimation of Large Dimensional Covariances: Monte Carlo

- Second, let

$$h_{it} = \left(\frac{10\widetilde{h}_{it}}{\max_{1 \le j \le t} \left| \widetilde{h}_{it} \right|} \right) + 10, \; i = 1, .., N$$

  where $\widetilde{h}_{it} = \widetilde{h}_{it-1} + \eta_{it}$ and $\eta_{it} \sim niid(0.1, 1)$. This ensures that h_{it}, is bounded between 0 and 20.
- Third, sample an $N \times 1$ vector $d = (d_1, d_2, .., d_N)$, $d_i \tilde{} \chi^2_2$. Then, define T $N \times 1$ vectors $\mathbf{e}_t = (e_{1t}, e_{2t}, .., e_{Nt})$, $e_{it} = h_{it} d_i$. Then, compute the $N \times N$ matrix $D_t = diag\left(\mathbf{e}_t\right)$. Set

$$\tilde{\mathbf{\Sigma}}_t = D_t + \mathbf{b}_t' \mathbf{b}_t$$

- Finally, denote $\tilde{\mathbf{\Sigma}}_t = [\tilde{\sigma}_{ijt}]$, set $\mathbf{\Sigma}_t = [\frac{\tilde{\sigma}_{ijt}}{\sqrt{\tilde{\sigma}_{ii1}\tilde{\sigma}_{jj1}}}]$ and generate data as $\mathbf{y}_t = \mathbf{\Sigma}_t^{1/2} \boldsymbol{\epsilon}_t$ where $\boldsymbol{\epsilon}_t \sim N(0, I)$.

Time Varying Estimation of Large Dimensional Covariances: Monte Carlo

- We consider $d_b = 0.5, 0.75, 1$, $N = 10, 40, 100$ and $T = 100, 200, 400$.
- We use two cross-validation exercises.
- One focuses on in-sample estimation and uses a leave-one out approach in minimising $||\hat{\mathbf{\Sigma}}_t^{(-1)} - \mathbf{y}_t\mathbf{y}_t'||$, where $||.||$ denotes the Frobenius norm, to determine H and any other tuning parameters, using the whole sample. The notation $.^{(-1)}$ indicates the leave-one out estimator.
- The other focuses on pseudo out-of-sample estimation, more relevant for portfolio optimisation, and minimises $||\hat{\mathbf{\Sigma}}_t - \mathbf{y}_{t+1}\mathbf{y}_{t+1}'||$ over the last 20 periods of the sample.
- The above define estimation and forecasting strategies that are then evaluated using $||\hat{\mathbf{\Sigma}}_t - \mathbf{\Sigma}_t||$ and $||\hat{\mathbf{\Sigma}}_t - \mathbf{\Sigma}_{t+1}||$, respectively as the evaluation criterion.

A summary of the Monte Carlo results

	Whole sample											
N	10	10	10	50	50	50	10	10	10	50	50	50
T	200	200	200	200	200	200	400	400	400	400	400	400
d_b	1	0.75	0.5	1	0.75	0.5	1	0.75	0.5	1	0.75	0.5
TV methods	10	10	10	10	10	8	10	10	10	10	10	10
Whole Sample methods	0	0	0	0	0	2	0	0	0	0	0	0
TV CL	0	6	8	0	10	8	2	7	9	1	9	10
TV FAN	0	0	0	0	0	0	0	0	0	0	0	0
TV LW	5	3	2	5	0	0	5	2	1	5	1	0
TV sample covariance	5	1	0	5	0	0	3	1	0	4	0	0
TV method with h=.5	0	0	0	0	0	0	0	0	0	0	0	0
TV method with h=.6	1	0	0	0	0	0	0	0	1	0	0	1
TV method with h=.7	2	2	1	2	1	1	2	3	1	2	1	1
TV method with h=.8	4	4	4	4	3	2	6	3	3	4	3	2
TV method with h=.9	2	3	2	2	3	2	1	2	2	1	3	3
TV CV methods	1	1	3	2	3	3	1	2	3	3	3	3

A summary of the Monte Carlo results

	End of sample											
N	10	10	10	50	50	50	10	10	10	50	50	50
T	200	200	200	200	200	200	400	400	400	400	400	400
d_b	1	0.75	0.5	1	0.75	0.5	1	0.75	0.5	1	0.75	0.5
TV methods	10	10	10	10	10	6	10	10	10	10	10	7
Fixed methods	0	0	0	0	0	4	0	0	0	0	0	3
TV CL	2	5	8	1	9	6	3	5	8	3	9	7
TV FAN	0	0	0	0	0	0	0	0	0	0	0	0
TV LW	4	2	2	4	0	0	3	1	1	3	0	0
TV sample covariance	4	3	0	5	1	0	4	4	1	4	1	0
TV method with h=.5	0	0	0	0	0	0	0	0	0	0	0	0
TV method with h=.6	0	0	0	0	0	0	0	0	0	0	0	0
TV method with h=.7	2	3	1	2	1	1	4	3	1	3	1	1
TV method with h=.8	5	3	2	5	3	1	6	5	2	6	3	1
TV method with h=.9	3	4	4	3	3	1	0	2	4	1	3	2
TV CV methods	0	0	3	0	3	3	0	0	3	0	3	3

Insample Simulations

T=200 N=10 db=0.5		T=200 N=40 db=0.5		T=200 N=100 db=0.5	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{tr}$
TV-n-h=.7-CAI-h	0.48	TV-n-h=.8-CAI-h	0.345	TV-n-h=.8-CAI-h	0.184
TV-n-h=.8-CAI-h	0.486	TV-n-h=.7-CAI-h	0.363	TV-n-h=.9-CAI-h	0.219
TV-n-h=.9-CAI-h	0.62	TV-n-h=.9-CAI-h	0.412	TV-n-h=.7-CAI-h	0.225
TV-n-h=.8-LW	0.634	TV-n-h=.95-CAI-h	0.441	TV-n-h=.95-CAI-h	0.235
TV-n-h=.7-LW	0.638	CAI-h	0.495	CAI-h-oc	0.258
TV-n-h=.6-CAI-h	0.671	CAI-h-oc	0.499	CAI-h	0.268
TV-n-h=.95-CAI-h	0.681	CAI-al	0.537	CAI-al	0.301
TV-n-h=.8	0.738	TV-n-h=.6-CAI-h	0.611	TV-n-h=.9-LW	0.46
TV-n-oc	0.739	TV-n-h=.8-LW	0.647	TV-n-h=.95-LW	0.467
TV-e-h=.9	0.745	TV-n-h=.9-LW	0.666	TV-n-h=.6-CAI-h	0.477
CAI-h	0.794	TV-n-h=.95-LW	0.687	TV-n-h=.8-LW	0.493
CAI-h-oc	0.85	LW	0.738	LW	0.495
LW	0.927	LW-oc	0.739	LW-oc	0.495
LW-oc	0.927	TV-n-oc	0.908	TV-n-oc	0.983
sample-estimate	1	sample-estimate	1	sample-estimate	1

Insample Simulations

T=200 N=10 db=.75		T=200 N=40 db=0.75		T=200 N=100 db=0.75	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}$
TV-n-oc	0.394	TV-n-h=.7-CAI-h	0.584	TV-n-h=.7-CAI-h	0.584
TV-n-h=.5-LW	0.442	TV-n-h=.8-CAI-h	0.603	TV-n-h=.8-CAI-h	0.603
TV-n-h=.5-CAI-h	0.444	TV-n-h=.8-LW	0.695	TV-n-h=.9-CAI-h	0.72
TV-n-h=.5	0.456	TV-n-h=.7-LW	0.717	CAI-h-oc	0.72
TV-n-h=.6-CAI-h	0.469	TV-n-h=.6-CAI-h	0.727	TV-n-h=.6-CAI-h	0.732
TV-n-h=.6-LW	0.473	TV-n-h=.9-CAI-h	0.735	TV-n-h=.8-LW	0.737
TV-e-h=.6	0.477	TV-n-oc	0.753	TV-n-h=.95-CAI-h	0.771
TV-n-h=.6	0.485	TV-n-h=.8	0.758	TV-n-h=.9-LW	0.801
TV-e-h=.7	0.507	TV-e-h=.9	0.768	TV-n-h=.7-LW	0.812
TV-n-h=.7-CAI-h	0.531	TV-e-h=.95	0.787	TV-n-h=.95-LW	0.841
CAI-h	0.991	CAI-h-oc	0.839	TV-n-oc	0.848
CAI-h-oc	0.993	CAI-h	0.909	CAI-h	0.87
LW	0.996	LW-oc	0.959	LW	0.927
LW-oc	0.996	LW	0.96	LW-oc	0.927
sample-estimate	1	sample-estimate	1	sample-estimate	1

Insample Simulations

T=200 N=10 db=1		T=200 N=40 db=1		T=200 N=100 db=1	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$
TV-n-oc	0.446	TV-n-h=.6-LW	0.46	TV-n-h=.6	0.542
TV-n-h=.5	0.47	TV-n-oc	0.461	TV-e-h=.8	0.548
TV-n-h=.5-LW	0.476	TV-n-h=.6	0.464	TV-n-h=.7	0.553
TV-e-h=.6	0.48	TV-e-h=.8	0.479	TV-n-oc	0.555
TV-n-h=.6	0.505	TV-n-h=.7-LW	0.48	TV-n-h=.6-LW	0.562
TV-n-h=.6-LW	0.508	TV-n-h=.7	0.486	TV-n-h=.7-LW	0.574
TV-e-h=.7	0.51	TV-e-h=.7	0.497	TV-e-h=.7	0.576
TV-n-h=.5-CAI-h	0.514	TV-f-h=.7	0.509	TV-f-h=.7	0.583
TV-e-h=.5	0.524	TV-n-h=.5-LW	0.552	TV-f-h=.8	0.617
TV-f-h=.6	0.528	TV-n-h=.5	0.558	TV-e-h=.9	0.639
sample-estimate	1	LW-oc	0.995	sample-estimate	1
CAI-h-oc	1	LW	0.996	CAI-h-oc	1.001
LW	1.002	sample-estimate	1	LW	1.025
LW-oc	1.004	CAI-h-oc	1.001	LW-oc	1.03
CAI-h	1.025	CAI-h	1.203	CAI-h	1.55

Insample Simulations

T=400 N=10 db=0.5		T=400 N=40 db=0.5		T=400 N=100 db=0.5	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}$
TV-n-h=.7-CAI-h	0.378	TV-n-h=.8-CAI-h	0.27	TV-n-h=.8-CAI-h	0.17
TV-n-h=.8-CAI-h	0.394	TV-n-h=.7-CAI-h	0.344	TV-n-h=.9-CAI-h	0.236
TV-n-h=.8-LW	0.532	TV-n-h=.9-CAI-h	0.385	TV-n-h=.7-CAI-h	0.26
TV-n-h=.7-LW	0.535	TV-n-h=.95-CAI-h	0.445	TV-n-h=.95-CAI-h	0.271
TV-n-oc	0.567	CAI-h	0.544	CAI-h-oc	0.328
TV-e-h=.9	0.584	CAI-al	0.558	CAI-h	0.329
TV-n-h=.8	0.589	CAI-h-oc	0.559	CAI-al	0.346
TV-n-h=.9-CAI-h	0.594	TV-n-h=.8-LW	0.639	TV-n-h=.9-LW	0.569
TV-n-h=.6-CAI-h	0.604	TV-n-h=.9-LW	0.683	TV-n-h=.95-LW	0.592
TV-n-h=.7	0.62	TV-n-h=.95-LW	0.733	TV-n-h=.8-LW	0.602
CAI-h	0.839	TV-n-h=.6-CAI-h	0.753	LW	0.649
CAI-h-oc	0.89	LW	0.832	LW-oc	0.649
LW	0.967	LW-oc	0.832	TV-n-h=.6-CAI-h	0.706
LW-oc	0.967	TV-n-oc	0.848	TV-n-oc	0.95
sample-estimate	1	sample-estimate	1	sample-estimate	1

Insample Simulations

T=400 N=10 db=0.75	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$	T=400 N=40 db=0.75	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$	400.0000 N=100 db=0.75	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}$
TV-n-h=.7-CAI-h	0.355	TV-n-h=.7-CAI-h	0.373	TV-n-h=.7-CAI-h	0.399
TV-n-h=.7-LW	0.398	TV-n-h=.8-CAI-h	0.393	TV-n-h=.8-CAI-h	0.402
TV-n-oc	0.402	TV-n-h=.8-LW	0.526	TV-n-h=.9-CAI-h	0.58
TV-n-h=.7	0.406	TV-n-oc	0.544	TV-n-h=.6-CAI-h	0.628
TV-e-h=.8	0.423	TV-n-h=.7-LW	0.552	TV-n-h=.8-LW	0.636
TV-n-h=.8-CAI-h	0.426	TV-e-h=.9	0.553	TV-n-h=.95-CAI-h	0.665
TV-f-h=.8	0.446	TV-n-h=.6-CAI-h	0.559	TV-n-h=.8	0.681
TV-n-h=.6-CAI-h	0.451	TV-n-h=.8	0.559	TV-n-oc	0.682
TV-e-h=.9	0.456	TV-n-h=.7	0.593	TV-e-h=.9	0.691
TV-n-h=.8-LW	0.467	TV-f-h=.8	0.613	TV-e-h=.95	0.715
CAI-h	0.966	CAI-h-oc	0.875	CAI-h-oc	0.764
CAI-h-oc	0.975	CAI-h	0.886	CAI-h	0.81
sample-estimate	1	LW	0.983	LW	0.982
LW	1.001	LW-oc	0.983	LW-oc	0.984
LW-oc	1.002	sample-estimate	1	sample-estimate	1

Insample Simulations

T=400 N=10 db=1		T=400 N=40 db=1		T=400 N=100 db=1	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$
TV-n-h=.7-LW	0.293	TV-n-h=.7-LW	0.329	TV-n-h=.7-LW	0.345
TV-n-h=.7	0.297	TV-n-h=.7	0.332	TV-n-h=.7	0.346
TV-e-h=.8	0.299	TV-n-oc	0.335	TV-n-oc	0.348
TV-n-oc	0.304	TV-e-h=.8	0.337	TV-e-h=.8	0.35
TV-n-h=.6-LW	0.305	TV-n-h=.6-LW	0.356	TV-n-h=.6-LW	0.369
TV-n-h=.6	0.307	TV-n-h=.6	0.357	TV-n-h=.6	0.369
TV-f-h=.7	0.324	TV-f-h=.7	0.378	TV-f-h=.7	0.388
TV-n-h=.7-CAI-h	0.328	TV-f-h=.8	0.381	TV-e-h=.7	0.399
TV-e-h=.7	0.331	TV-e-h=.7	0.387	TV-f-h=.8	0.4
TV-f-h=.8	0.351	TV-e-h=.9	0.422	TV-e-h=.9	0.44
LW	1	LW	1	LW-oc	1
sample-estimate	1	sample-estimate	1	sample-estimate	1
LW-oc	1	LW-oc	1	CAI-h-oc	1
CAI-h-oc	1.001	CAI-h-oc	1	LW	1.001
CAI-h	1.018	CAI-h	1.069	CAI-h	1.127

Out of Sample Simulations

T=200 N=10 db=0.5		T=200 N=40 db=0.5		T=200 N=100 db=0.5	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$
TV-n-CAI-h-oc	0.59	TV-n-CAI-h-oc	0.46	TV-n-CAI-h-oc	0.31
TV-n-h=.8-LW	0.67	CAI-h	0.61	CAI-h	0.43
TV-e-h=.95	0.69	CAI-al	0.73	CAI-al	0.55
TV-n-h=.8	0.69	TV-n-h=.9-LW	0.82	TV-n-LW-oc	0.7
TV-f-h=.9	0.72	TV-n-LW-oc	0.82	TV-n-h=.9-LW	0.7
TV-e-h=.9	0.72	TV-n-LW(I)-oc	0.85	TV-n-h=.95-LW	0.71
TV-n-h=.9	0.74	TV-n-h=.95-LW	0.85	TV-n-LW(I)-oc	0.72
TV-f-h=.95	0.75	TV-n-h=.8-LW	0.87	LW	0.76
CAI-h	0.76	TV-n-h=.95	0.93	TV-n-h=.8-LW	0.89
TV-n-oc	0.76	TV-n-oc	0.93	LW-oc	0.9
TV-n-h=.7-LW	0.76	LW	0.93	TV-n-oc	1
TV-n-h=.9-LW	0.77	TV-n-h=.9	0.95	sample estimate	1
TV-n-h=.95	0.78	TV-f-h=.95	0.99	TV-n-h=.95	1.03
TV-n-LW(I)-oc	0.78	sample estimate	1	TV-n-h=.9	1.1
TV-n-LW-oc	0.79	LW-oc	1.06	TV-f-h=.95	1.17
LW	0.94	TV-e-h=.95	1.11	CAI-s	1.25
sample estimate	1	TV-f-h=.9	1.15	TV-e-h=.95	1.43
LW-oc	1.21	TV-n-h=.8	1.16	TV-f-h=.9	1.48

Out of Sample Simulations

T=200 N=10 db=0.75		T=200 N=40 db=0.75		T=200 N=100 db=0.75	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$
TV-e-h=.9	0.41	TV-n-CAI-h-oc	0.59	TV-n-CAI-h-oc	0.61
TV-n-h=.7	0.42	TV-n-h=.8	0.67	TV-n-h=.9	0.79
TV-n-h=.7-LW	0.43	TV-e-h=.95	0.68	TV-n-h=.95	0.8
TV-n-h=.8	0.43	TV-e-h=.9	0.69	TV-e-h=.95	0.81
TV-f-h=.8	0.44	TV-n-h=.8-LW	0.71	TV-n-h=.8	0.81
TV-e-h=.95	0.45	TV-f-h=.9	0.71	TV-n-oc	0.82
TV-e-h=.8	0.46	TV-n-oc	0.72	TV-f-h=.95	0.82
TV-n-h=.8-LW	0.46	TV-n-h=.9	0.72	TV-n-h=.8-LW	0.83
TV-n-oc	0.46	TV-f-h=.95	0.75	TV-f-h=.9	0.85
TV-n-CAI-h-oc	0.46	TV-n-h=.95	0.76	TV-e-h=.9	0.88
TV-f-h=.9	0.46	TV-n-h=.7-LW	0.77	TV-n-h=.9-LW	0.88
TV-n-LW-oc	0.5	TV-n-LW-oc	0.78	TV-n-LW-oc	0.89
TV-n-LW(I)-oc	0.51	TV-n-LW(I)-oc	0.78	TV-n-LW(I)-oc	0.89
TV-n-h=.9	0.52	TV-n-h=.7	0.79	TV-n-h=.95-LW	0.93
CAI-h	0.68	CAI-h	0.89	CAI-h	0.95
LW	0.73	LW	0.97	sample estimate	1
sample estimate	1	sample estimate	1	LW	1.02
LW-oc	1.18	LW-oc	1.66	LW-oc	1.62

Out of Sample Simulations

T=200 N=10 db=1		T=200 N=40 db=1		T=200 N=100 db=1	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$
TV-n-h=.7	0.47	TV-n-h=.7	0.45	TV-e-h=.8	0.54
TV-n-h=.7-LW	0.48	TV-n-h=.7-LW	0.45	TV-n-h=.6	0.54
TV-e-h=.8	0.49	TV-e-h=.8	0.45	TV-n-h=.7	0.54
TV-e-h=.9	0.51	TV-f-h=.8	0.48	TV-n-h=.6-LW	0.55
TV-f-h=.8	0.51	TV-n-h=.6	0.49	TV-n-h=.7-LW	0.56
TV-n-h=.6	0.53	TV-n-h=.6-LW	0.49	TV-f-h=.7	0.57
TV-n-h=.6-LW	0.53	TV-e-h=.9	0.5	TV-e-h=.7	0.58
TV-n-h=.8	0.54	TV-f-h=.7	0.52	TV-f-h=.8	0.58
TV-f-h=.7	0.56	TV-n-oc	0.53	TV-e-h=.9	0.6
TV-n-h=.8-LW	0.57	TV-n-h=.8	0.53	TV-n-oc	0.62
TV-n-oc	0.58	TV-n-CAI-h-oc	0.53	TV-n-CAI-h-oc	0.63
TV-n-CAI-h-oc	0.62	TV-n-h=.8-LW	0.54	TV-n-h=.8	0.64
TV-n-LW-oc	0.63	TV-n-LW-oc	0.57	TV-n-LW-oc	0.67
TV-n-LW(I)-oc	0.64	TV-n-LW(I)-oc	0.57	TV-n-LW(I)-oc	0.67
LW	0.96	LW	0.89	LW	0.98
CAI-h	1	sample estimate	1	sample estimate	1
sample estimate	1	CAI-h	1.04	CAI-h	1.34
LW-oc	1.89	LW-oc	1.77	LW-oc	1.79

Out of Sample Simulations

T=400 N=10 db=0.5		T=400 N=40 db=0.5		T=400 N=100 db=0.5	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$
TV-n-CAI-h-oc	0.41	TV-n-CAI-h-oc	0.35	TV-n-CAI-h-oc	0.29
TV-e-h=.9	0.52	CAI-h	0.69	CAI-h	0.53
TV-n-h=.8	0.52	CAI-al	0.76	CAI-al	0.61
TV-n-h=.8-LW	0.53	TV-n-h=.8-LW	0.81	TV-n-h=.9-LW	0.86
TV-e-h=.95	0.55	TV-n-h=.9	0.82	TV-n-h=.95-LW	0.89
TV-n-h=.7-LW	0.57	TV-n-h=.9-LW	0.83	TV-n-LW(I)-oc	0.89
TV-f-h=.9	0.57	TV-n-oc	0.84	TV-n-LW-oc	0.89
TV-n-oc	0.6	TV-n-h=.95	0.84	TV-n-oc	0.96
TV-f-h=.8	0.6	TV-n-LW(I)-oc	0.84	TV-n-h=.95	0.97
TV-n-h=.7	0.6	TV-f-h=.95	0.84	LW	0.99
TV-n-LW(I)-oc	0.61	TV-n-LW-oc	0.84	sample estimate	1
TV-n-LW-oc	0.62	TV-e-h=.95	0.85	TV-n-h=.9	1.01
TV-e-h=.8	0.66	TV-n-h=.95-LW	0.9	TV-n-h=.8-LW	1.04
TV-n-h=.9	0.68	TV-f-h=.9	0.9	TV-f-h=.95	1.06
CAI-h	0.85	TV-n-h=.8	0.92	LW-oc	1.21
LW	0.99	sample estimate	1	TV-e-h=.95	1.23
sample estimate	1	LW	1.03	TV-f-h=.9	1.31
LW-oc	1.29	LW-oc	1.36	TV-n-h=.8	1.42

Out of Sample Simulations

T=400 N=10 db=0.75		T=400 N=40 db=0.75		T=400 N=100 db=0.75	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$
TV-n-h=.7	0.36	TV-n-CAI-h-oc	0.45	TV-n-CAI-h-oc	0.5
TV-n-h=.7-LW	0.36	TV-e-h=.9	0.52	TV-n-h=.8	0.66
TV-f-h=.8	0.37	TV-n-h=.8	0.53	TV-e-h=.95	0.67
TV-e-h=.9	0.37	TV-n-h=.8-LW	0.56	TV-e-h=.9	0.68
TV-e-h=.8	0.37	TV-e-h=.95	0.56	TV-n-h=.8-LW	0.69
TV-n-h=.8	0.4	TV-f-h=.9	0.58	TV-f-h=.9	0.71
TV-n-CAI-h-oc	0.41	TV-f-h=.8	0.59	TV-n-h=.9	0.75
TV-n-h=.8-LW	0.42	TV-n-h=.7-LW	0.59	TV-n-oc	0.75
TV-n-oc	0.43	TV-n-h=.7	0.59	TV-f-h=.95	0.77
TV-e-h=.95	0.45	TV-n-oc	0.61	TV-n-LW(I)-oc	0.8
TV-f-h=.9	0.46	TV-n-LW-oc	0.64	TV-n-LW-oc	0.8
TV-n-LW-oc	0.47	TV-n-LW(I)-oc	0.65	TV-f-h=.8	0.81
TV-n-LW(I)-oc	0.48	TV-e-h=.8	0.65	TV-n-h=.95	0.81
TV-n-h=.6-LW	0.48	TV-n-h=.9	0.67	TV-n-h=.7-LW	0.82
CAI-h	0.81	CAI-h	0.86	CAI-h	0.9
LW	0.87	LW	0.98	sample estimate	1
sample estimate	1	sample estimate	1	LW	1.07
LW-oc	1.78	LW-oc	1.87	LW-oc	2.16

Out of Sample Simulations

T=400 N=10 db=1		T=400 N=40 db=1		T=400 N=100 db=1	
	$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$		$\|\|\widehat{\Sigma}_t - \Sigma_t^{true}\|\|$
TV-n-h=.7	0.3	TV-e-h=.8	0.3	TV-n-h=.7	0.34
TV-e-h=.8	0.3	TV-n-h=.7	0.31	TV-e-h=.8	0.35
TV-n-h=.7-LW	0.3	TV-n-h=.7-LW	0.31	TV-n-h=.7-LW	0.35
TV-f-h=.8	0.32	TV-n-h=.6	0.33	TV-f-h=.8	0.38
TV-f-h=.7	0.34	TV-n-h=.6-LW	0.33	TV-n-h=.6	0.39
TV-n-h=.6	0.34	TV-f-h=.7	0.34	TV-f-h=.7	0.39
TV-n-h=.6-LW	0.35	TV-f-h=.8	0.34	TV-n-h=.6-LW	0.39
TV-e-h=.9	0.36	TV-e-h=.7	0.36	TV-e-h=.9	0.42
TV-e-h=.7	0.37	TV-n-oc	0.38	TV-e-h=.7	0.42
TV-n-h=.8	0.39	TV-n-CAI-h-oc	0.38	TV-n-CAI-h-oc	0.44
TV-n-oc	0.4	TV-e-h=.9	0.38	TV-n-oc	0.44
TV-n-CAI-h-oc	0.41	TV-n-LW-oc	0.4	TV-n-h=.8	0.45
TV-n-LW-oc	0.42	TV-n-LW(I)-oc	0.4	TV-n-LW(I)-oc	0.47
TV-n-LW(I)-oc	0.42	TV-n-h=.8	0.42	TV-n-LW-oc	0.47
CAI-h	0.91	LW	0.96	LW	0.96
LW	0.92	CAI-h	0.98	sample estimate	1
sample estimate	1	sample estimate	1	CAI-h	1.03
LW-oc	1.8	LW-oc	2.2	LW-oc	2.32

Data Description

Portfolio	Time Period	out of sample observations
5 industry portfolios	03/63 to 12/13 (609 obs)	from 01/06 to 12/13 (95 obs)
10 industry portfolios	03/63 to 12/13 (609 obs)	from 01/06 to 12/13 (95 obs)
17 industry portfolios	03/63 to 12/13 (609 obs)	from 01/06 to 12/13 (95 obs)
30 industry portfolios	03/63 to 12/13 (609 obs)	from 01/06 to 12/13 (95 obs)
6 size and book to market portfolios	03/63 to 12/13 (609 obs)	from 01/06 to 12/13 (95 obs)
25 size and book to market portfolios	03/63 to 12/13 (609 obs)	from 01/06 to 12/13 (95 obs)
The 5 S&P500 stocks with the highest average capitalization	03/93 to 12/13 (219 obs)	from 01/09 to 12/13 (60 obs)
The 10 S&P500 stocks with the highest average capitalization	03/93 to 12/13 (219 obs)	from 01/09 to 12/13 (60 obs)
The 20 S&P500 stocks with the highest average capitalization	03/93 to 12/13 (219 obs)	from 01/09 to 12/13 (60 obs)
The 30 S&P500 stocks with the highest average capitalization	03/93 to 12/13 (219 obs)	from 01/09 to 12/13 (60 obs)
The 40 S&P500 stocks with the highest average capitalization	03/93 to 12/13 (219 obs)	from 01/09 to 12/13 (60 obs)

Performance Criteria :

$$var = \frac{1}{T_n} \sum_{\tau=T_0}^{T} \left(w_{\theta,\tau|\tau-1} R_\tau - \frac{1}{T_n} \sum_{\tau=T_0}^{T} w_{\theta,\tau|\tau-1} R_\tau \right)^2 \text{ and } SR = \sum_{\tau=T_0}^{T} w_{\theta,\tau|\tau-1} R_\tau \Big/ var$$

$T_0 = o(T)$, $T_n := T - T_0 + 1$. $T_n = 60$ observations for the S&P500 stocks, $T_n = 95$ for the remaining portfolios

A summary of the empirical results

variance													
N	5	10	12	17	30	48	6	25	5	10	20	30	40
T	960	960	960	960	960	960	963	963	250	250	250	250	250
TV methods	10	10	10	10	10	7	10	9	10	10	10	9	9
Fixed methods	0	0	0	0	0	3	0	1	0	0	0	1	1
TV CL	2	3	2	0	0	0	3	2	3	0	0	0	0
TV FAN	0	2	0	1	0	1	1	0	0	0	8	1	3
TV LW	2	1	2	3	4	4	1	3	0	1	2	1	3
TV sample covariance	3	3	5	3	3	1	2	0	4	6	0	3	1
TV method with h=.5	1	0	0	0	0	0	3	0	1	0	0	0	0
TV method with h=.6	4	5	1	3	1	0	0	3	1	1	2	0	0
TV method with h=.7	0	0	2	6	3	1	2	5	0	2	1	0	1
TV method with h=.8	0	3	5	0	6	1	0	0	1	2	4	3	3
TV method with h=.9	0	0	0	0	0	3	0	0	1	0	1	2	3
TV CV methods	5	2	2	1	0	2	5	1	6	5	2	4	2
TV shrink the inverse	2	0	0	0	0	0	2	0	0	2	0	1	0
TV LW 1 factor	1	1	1	3	3	1	1	4	3	1	0	3	2

A summary of the empirical results

sharp ratio

N	5	10	12	17	30	48	6	25	5	10	20	30	40
T	960	960	960	960	960	960	963	963	250	250	250	250	250
TV methods	7	9	9	6	3	5	10	9	10	10	10	10	10
Fixed methods	3	1	1	4	7	5	0	1	0	0	0	0	0
TV CL	3	7	4	0	2	1	5	4	5	6	9	9	9
TV FAN	3	2	5	6	0	4	2	5	2	0	0	0	1
TV LW	0	0	0	0	0	0	0	0	0	1	1	0	0
TV sample covariance	1	0	0	0	1	0	3	0	3	3	0	1	0
TV method with h=.5	1	0	1	1	0	0	2	2	6	4	4	3	1
TV method with h=.6	1	1	1	0	0	0	0	0	3	2	2	2	3
TV method with h=.7	2	3	0	1	0	0	4	1	0	1	2	2	2
TV method with h=.8	0	3	6	4	0	2	3	5	0	0	1	1	0
TV method with h=.9	0	0	0	0	3	1	0	0	0	0	1	2	3
TV CV methods	3	2	1	0	0	2	1	1	1	3	0	0	1
TV shrink the inverse	0	0	0	0	0	0	0	0	0	0	0	0	0
TV LW 1 factor	0	0	0	0	0	0	0	0	0	0	0	0	0

A summary of the empirical results

turnover

N	5	10	12	17	30	48	6	25	5	10	20	30	40
T	960	960	960	960	960	960	963	963	250	250	250	250	250
TV methods	3	3	4	4	5	6	1	4	5	3	5	5	7
Fixed methods	7	7	6	6	5	4	9	6	5	7	5	5	3
TV CL	0	0	1	1	2	0	0	1	1	0	0	0	0
TV FAN	0	0	0	0	0	0	0	0	0	0	2	2	4
TV LW	1	1	1	1	1	3	0	1	1	1	1	1	1
TV sample covariance	0	0	0	0	0	1	0	0	1	0	0	0	0
TV method with h=.5	0	0	0	0	0	0	0	0	0	0	0	0	0
TV method with h=.6	0	0	0	0	0	0	0	0	0	0	0	0	0
TV method with h=.7	0	0	0	0	0	0	0	0	0	0	0	0	0
TV method with h=.8	0	0	0	0	0	1	0	0	0	0	1	1	1
TV method with h=.9	2	2	3	3	4	4	0	3	4	2	3	3	5
TV CV methods	0	0	0	0	0	0	0	0	0	0	0	0	0
TV shrink the inverse	0	0	0	0	0	0	0	0	0	0	0	0	0
TV LW 1 factor	1	1	1	1	1	1	0	1	1	1	1	1	1

A summary of the empirical results

certainty equivalent

N	5	10	12	17	30	48	6	25	5	10	20	30	40
T	960	960	960	960	960	960	963	963	250	250	250	250	250
TV methods	10	10	10	10	10	7	10	9	10	10	10	7	9
Fixed methods	0	0	0	0	0	3	0	1	0	0	0	3	1
TV CL	2	3	2	0	0	1	2	4	2	0	0	0	0
TV FAN	0	2	0	1	1	2	2	0	0	0	7	2	4
TV LW	1	0	2	2	3	2	1	2	0	1	2	1	2
TV sample covariance	4	5	5	4	3	1	2	0	5	6	1	2	1
TV method with h=.5	1	0	0	0	0	0	2	0	2	0	0	0	0
TV method with h=.6	3	3	1	3	1	0	0	3	1	1	2	0	1
TV method with h=.7	1	2	3	4	3	0	3	5	0	2	1	0	1
TV method with h=.8	0	3	5	2	6	1	0	1	1	2	4	4	3
TV method with h=.9	0	0	0	0	0	4	0	0	0	0	1	2	3
TV CV methods	5	2	1	1	0	2	5	0	6	5	2	1	1
TV shrink the inverse	2	0	0	0	0	0	2	0	0	2	0	0	0
TV LW 1 factor	1	0	1	3	3	1	1	3	3	1	0	2	2

Out of Sample Empirical Results

N=5, Industry Portfolios			N=10, Industry Portfolios			N=17, Industry Portfolios		
	var	SR		var	SR		var	SR
TV-e-h=.6	0.68	1.11	TV-f-h=.8	0.705	1.144	TV-e-h=.8	0.686	1.073
TV-f-h=.5	0.681	1.119	TV-f-oc	0.705	0.989	TV-f-oc	0.696	1.083
TV-e-ov	0.685	1.11	TV-e-h=.8	0.708	1.059	TV-n-h=.7	0.697	1.057
TV-n-ov	0.685	1.081	TV-n-h=.7	0.712	1.031	TV-f-h=.8	0.71	1.134
TV-n-h=.5	0.685	1.081	TV-n-h=.6	0.768	0.831	TV-n-h=.6-LW	0.727	0.918
TV-n-LW-ov	0.692	1.075	TV-n-CAI-ov	0.777	0.813	TV-n-LW-ov	0.749	0.876
TV-n-LW(I)-ov	0.693	1.068	TV-n-LW-ov	0.828	0.691	TV-n-LW(I)-ov	0.763	0.865
TV-ok-ov	0.694	1.078	TV-n-ov	0.837	0.735	TV-n-ov	0.768	0.919
TV-n-oc	0.728	0.72	TV-n-LW(I)-ov	0.839	0.717	TV-n-LW(I)-oc	0.845	0.679
TV-ok-oc	0.745	0.785	TV-ok-oc	0.844	0.704	TV-n-LW-oc	0.86	0.653
TV-n-LW-oc	0.777	0.643	TV-n-LW-oc	0.844	0.671	TV-n-oc	0.863	0.84
TV-n-LW(I)-oc	0.789	0.636	TV-n-LW(I)-oc	0.847	0.666	TV-n-CAI-ov	0.876	0.807
LW-oc	0.951	0.734	TV-n-oc	0.848	0.75	TV-ok-oc	0.904	0.767
LW	0.992	0.972	LW-oc	0.955	0.777	LW-oc	0.908	0.782
CAI-h-ov	1	1	LW	0.987	0.978	LW	0.986	0.985
CAI-h-oc	1	1	CAI-h-ov	1	1	CAI-h	1	1
CAI-h	1	1	CAI-h	1	1	sample estimate	1	1
sample estimate	1	1	sample estimate	1	1	LW-ov	1.02	0.712
LW-ov	1.071	0.641	LW-ov	1.072	0.621	CAI-h-ov	1.051	0.918
true	1.249	0.621	CAI-h-oc	1.409	0.889	CAI-h-oc	1.95	0.103
TV-n-CAI-h-oc	1.835	0.648	TV-n-CAI-h-oc	1.496	0.5	true	2.06	0.57
TV-n-CAI-ov	2.103	0.404	true	1.503	0.63	TV-n-CAI-h-oc	8.868	0.177

Out of Sample Empirical Results

N=30, Industry Portfolios			N=6, size-book to market portfolios			N=25, size-book to market portfolios		
	var	SR		var	SR		var	SR
TV-f-h=.8	0.749	0.881	TV-n-h=.5	0.777	1.009	TV-n-LW(I)-ov	0.946	0.99
TV-n-h=.7	0.771	0.743	TV-n-h=.6	0.78	1.054	TV-n-h=.6-LW	0.953	0.933
TV-e-h=.8	0.773	0.718	TV-e-h=.7	0.782	0.955	TV-n-LW-ov	0.977	0.781
TV-e-h=.9	0.784	0.873	TV-e-h=.6	0.789	1.04	TV-e-h=.8	0.99	0.834
TV-n-h=.8-LW	0.793	0.885	TV-f-ov	0.79	0.975	TV-n-h=.7	0.995	0.874
TV-n-LW-ov	0.95	0.689	TV-n-ov	0.792	1.028	CAI-h-oc	0.998	1.194
TV-n-LW(I)-ov	0.971	0.696	TV-n-LW-ov	0.794	1.025	CAI-h	1	1
LW	0.994	0.988	TV-n-LW(I)-ov	0.804	1.015	sample estimate	1	1
sample estimate	1	1	TV-n-oc	0.833	0.855	CAI-h-ov	1	1.093
CAI-h	1	1	TV-n-LW(I)-oc	0.851	0.729	TV-n-CAI-ov	1.006	1.162
CAI-h-ov	1	1	TV-n-LW-oc	0.851	0.72	TV-e-ov	1.007	0.797
TV-n-ov	1.011	0.599	TV-ok-oc	0.924	0.947	LW	1.009	0.953
LW-oc	1.109	0.712	sample estimate	1	1	TV-ok-ov	1.009	0.752
LW-ov	1.114	0.725	CAI-h-ov	1	1	TV-n-ov	1.014	0.862
TV-n-LW(I)-oc	1.219	0.496	CAI-h-oc	1	1	LW-ov	1.067	0.78
TV-n-LW-oc	1.245	0.503	CAI-h	1	1	LW-oc	1.211	0.592
TV-n-oc	1.302	0.484	LW-ov	1.002	1.002	TV-n-oc	1.532	0.565
TV-ok-oc	1.379	0.469	LW	1.049	0.935	TV-n-LW-oc	1.576	0.28
true	2.559	0.58	LW-oc	1.128	0.798	TV-n-LW(I)-oc	1.585	0.285
CAI-h-oc	2.902	0.091	true	1.467	0.793	TV-ok-oc	1.647	0.861
TV-n-CAI-h-oc	3.06	0.698	TV-n-CAI-ov	1.624	1.163	true	1.869	0.877
TV-n-CAI-ov	8.664	0.61	TV-n-CAI-h-oc	NaN	NaN	TV-n-CAI-h-oc	2.059	0.782

Out of Sample Empirical Results

N=5, SP500			N=10, SP500			N=20, SP500		
	var	SR		var	SR		var	SR
TV-n-h=.6-LW	0.65	1.03	TV-f-h=.7	0.7	1.48	TV-n-LW(I)-ov	0.82	1.47
TV-n-LW-ov	0.69	0.98	TV-f-ov	0.72	1.5	LW	0.82	1.8
TV-n-h=.5-LW	0.7	1.31	TV-ok-ov	0.72	1.5	TV-n-h=.95-LW	0.82	1.64
TV-e-oc	0.7	0.76	TV-e-h=.8	0.74	1.13	TV-n-h=.9-LW	0.83	1.54
TV-f-h=.7	0.71	0.89	TV-f-h=.8	0.78	0.81	TV-n-h=.8-LW	0.85	1.28
TV-n-h=.8	0.71	0.86	TV-n-LW(I)-ov	0.85	1.09	TV-n-LW-ov	0.88	1.13
TV-n-LW-oc	0.72	0.67	TV-n-LW-ov	0.87	1.05	TV-f-h=.95	0.93	1.06
TV-n-oc	0.73	0.61	TV-n-LW(I)-oc	0.94	1.32	LW-ov	0.93	1.22
TV-n-ov	0.74	0.79	TV-n-ov	0.96	0.98	TV-n-h=.6-LW	0.93	1.23
TV-n-CAI-ov	0.74	0.78	LW-ov	1	1.01	TV-n-ov	0.98	1.03
TV-ok-oc	0.74	0.61	sample estimate	1	1	LW-oc	1	1
TV-n-LW(I)-ov	0.83	0.8	LW-oc	1	1	sample estimate	1	1
TV-n-LW(I)-oc	0.9	0.49	CAI-h-oc	1	1	CAI-h-oc	1	0.98
CAI-h-oc	0.96	1.02	LW	1.04	1.48	CAI-h-ov	1.02	0.97
sample estimate	1	1	TV-n-LW-oc	1.16	0.91	TV-n-CAI-ov	1.14	0.48
CAI-h-ov	1.2	1.34	TV-n-oc	1.4	0.64	TV-n-LW(I)-oc	1.36	1.35
LW-oc	1.35	0.66	TV-ok-oc	1.52	0.64	TV-n-LW-oc	1.49	1.02
LW-ov	2.04	0.6	TV-n-CAI-ov	1.76	-0.2	true	1.64	2.59
LW	2.47	0.82	true	2.44	2.19	TV-n-oc	1.84	1.1
true	5.09	0.88	CAI-h-ov	2.76	-0.58	TV-ok-oc	3.26	0.96
CAI-h	64.97	0.35	CAI-h	40748.46	1.92	CAI-h	12.39	-0.92
TV-n-CAI-h-oc	NaN	NaN	TV-n-CAI-h-oc	NaN	NaN	TV-n-CAI-h-oc	NaN	NaN

Out of Sample Empirical Results

N=30, SP500			N=40, SP500		
TV-n-h=.9-LW	0.9	1.29	TV-n-h=.9-LW	0.89	1.31
TV-n-h=.95-LW	0.91	1.39	TV-n-h=.95-LW	0.89	1.41
LW	0.91	1.56	LW	0.92	1.57
TV-n-h=.8-LW	0.92	0.94	TV-n-h=.8-LW	0.92	0.92
TV-f-h=.95	0.96	0.98	LW-ov	0.96	1.07
LW-ov	0.97	1.14	TV-n-LW(I)-ov	1	1.23
TV-n-h=.95	0.99	0.92	TV-n-h=.95	1	0.9
LW-oc	1	1	LW-oc	1	1
sample estimate	1	1	sample estimate	1	1
CAI-h-oc	1.02	0.96	TV-n-h=.9	1.01	0.82
CAI-h-ov	1.13	0.82	CAI-h-oc	1.03	1.02
TV-n-LW-ov	1.36	0.1	TV-n-ov	1.07	1.08
TV-n-LW(I)-ov	1.37	0.28	CAI-h-ov	1.22	0.82
TV-n-ov	1.49	0.14	TV-n-LW-ov	1.23	0.41
true	1.71	2.38	TV-n-CAI-ov	1.42	0.19
TV-n-CAI-ov	1.94	0.18	true	1.8	2.57
TV-n-LW-oc	2.16	0	TV-n-LW(I)-oc	2.51	0.4
TV-n-LW(I)-oc	2.23	0.06	TV-n-LW-oc	2.68	0.08
TV-ok-oc	2.48	0.03	TV-ok-oc	3.38	0.19
TV-n-oc	2.99	-0.03	TV-n-oc	3.5	0.26
TV-n-CAI-h-oc	38.32	1.96	CAI-h	41.51	-1.35
CAI-h	596.25	1.33	TV-n-CAI-h-oc	288.56	-0.84