Appendix: A Nodewise Regression Approach to Estimating Large Portfolios

Laurent Callot [*] Mehmet Caner[†] A.Özlem Önder [‡] Esra Ulasan [§][¶]

October 16, 2019

In this section we provide proofs.

Before going through the main steps, the following norm inequality is used in all of the proofs. Take a $p \times p$ generic matrix: M, and a generic $p \times 1$ vector x. Note that M_j' represents $1 \times p$, jth row vector in M, and M_j is $p \times 1$ vector (i.e. transpose of M_j', or column version of M_j')

$$\begin{aligned} \|Mx\|_1 &= |M_1'x| + |M_2'x| + \cdots + |M_p'x| \\ &\leq \|M_1\|_1\|x\|_\infty + \|M_2\|_1\|x\|_\infty + \cdots + \|M_p\|_1\|x\|_\infty \\ &= [\sum_{j=1}^{p} \|M_j\|_1]\|x\|_\infty \\ &\leq p \max_j \|M_j\|_1\|x\|_\infty, \end{aligned} \tag{A.1}$$

where we use Holders inequality to get each inequality.

Next to use one of the results in Chang et al. (2019), we need the following notation. These are defined at the beginning of the Appendix at Chang et al. (2019). Let $\rho_1 = 2\Xi_2^{-1} + \Xi_1^{-1}$, $\rho_2^{-1} = 2\Xi_3^{-1} + \Xi_1^{-1}$,$\rho_3^{-1} = \Xi_1^{-1} + \Xi_2^{-1} + \Xi_3^{-1}$, $\rho_4^{-1} = \max\{\rho_2^{-1}, \rho_3^{-1}\} + \Xi_1^{-1}$. Then $\zeta = min\{\rho_1, \rho_2, \rho_3, \rho_4\}$. So they restrict the β mixing behaviour of data. In addition we assume, $0 \leq \zeta \leq 1$, which shapes the relation between n and $logp$ in the next statement

[*]Amazon Research, Berlin, email: lcallot@amazon.com.

[†]Department of Economics, and Department of Agriculture and Resource Economics, North Carolina State University, email: mcaner@ncsu.edu

[‡]Department of Economics, Ege University, email: ozlem.onder@ege.edu.tr.

[§]Department of Economics, Ege University, email: esra.ulasan@ege.edu.tr

[¶]We thank NYU-Stern Statistics, SOFIE 2017. We also thank editor Todd Clark, an associate editor, and two anonymous referees for comments that substantially improved the paper. We thank Professor Anna O. Soter and Nancy House for professional text editing.

of the Theorem. To give an example: with $\Xi_1 = 1, \Xi_2 = 2, \Xi_3 = 2$, which provides $\rho_1^{-1} = 2, \rho_2^{-1} = 2, \rho_3^{-1} = 2, \rho_4^{-1} = 3$. These imply $\zeta = 1/3$.

Next we provide a result that will be useful for subsequent proofs.

Theorem A.1. *Under Assumptions 1-3, with* $logp = o(n)$

(i).

$$\max_{1 \le j \le p} \|\hat{\Theta}_j - \Theta_j\|_1 = O_p(\bar{s}\sqrt{\frac{logp}{n}}).$$

(ii).

$$\|\hat{\mu} - \mu\|_\infty = O_p(\sqrt{\frac{logp}{n}}).$$

(iii).

$$\max_{1 \le j \le p} \|\Theta_j\|_1 = O(\sqrt{\bar{s}}).$$

Remark. Note that proof is based on Chang et al. (2019) and Caner and Kock (2018). This is basically extending the iid proof in (B.45)-(B.55) in Caner and Kock (2018) to stationary β mixing case benefiting from Lemma 1 and equation (23) of Chang et al. (2019).

Proof of Theorem A.1. First, to ease the notation a bit, we take the constant $0 \le \zeta < 2$ in Lemma 1 of Chang et al. (2019) between $0 \le \zeta \le 1$ so that we have $logp = o(n)$. Constant ζ is related to β mixing condition.

(i). Note that Chang et al. (2019) provide results in l_1 norm of the nodewise regression estimates in time series context for the first time in the literature. Their section 3 and equations (5)-(8), via their Assumptions 1-3 justify nodewise regression. They also demean the data by time series average. Specifically, nodewise regression estimate $\hat{\alpha}_j$ in their equation (7) of their paper is equivalent to negative of ours: $\hat{\alpha}_j = -(\hat{\gamma}_{j1}, \cdots, 1, \cdots, \hat{\gamma}_{jp})'$ where 1 is in j th position in $p \times 1$ vector, $\hat{\gamma}_j$ is defined in equation (2.1). Then equation (23) of Chang et al. (2019) provides, under Assumptions 1-3, and $logp = o(n)$

$$\max_{1 \le j \le p} \|\hat{\alpha}_j - \alpha_j\|_1 = \max_{1 \le j \le p} \|\hat{\gamma}_j - \gamma_j\|_1 = O_p\left(\bar{s}\sqrt{\frac{logp}{n}}\right). \tag{A.2}$$

To derive the limit for the rows of the nodewise regression based estimate for the precision matrix, we need more than (A.2). A key result that we need to derive is the limit of the term $\hat{\tau}^2$ in equation (2.2). This will help us understand the behavior of the denominator in (2.3). By (5) and p.6 of Chang et al. (2019), for $j = 1, \cdots, p$

$$r_j^* = \mathbf{r}^*{}_{-j}\gamma_j + \eta_j, \tag{A.3}$$

and

$$E\eta_j'\mathbf{r}^*{}_{-j} = 0. \tag{A.4}$$

Also Lemma 1 of Chang et al. (2019) provide the following results, under Assumptions 1-3, with $logp = o(n)$,

$$\max_{1\leq j\leq p}\|\eta_j'\mathbf{r}^*{}_{-j}/n\|_\infty = O_p(\sqrt{\frac{logp}{n}}), \tag{A.5}$$

and we define $\tau_j^2 = E\eta_j'\eta_j$:

$$\max_{1\leq j\leq p}|\frac{\eta_j'\eta_j}{n} - \tau_j^2| = O_p(\sqrt{\frac{logp}{n}}). \tag{A.6}$$

These are the fourth and second results in the statement of Lemma 1 of Chang et al. (2019), respectively. Now we provide a formula for the estimator $\hat{\tau}_j^2$ from (C.101) of Caner and Kock (2018).

$$\hat{\tau}_j^2 = \frac{(r_j^* - \mathbf{r}^*_{-j}\hat{\gamma}_j)'r_j^*}{n}.$$

Next using (A.3) in the formula above, and by triangle inequality or by (B.45) of Caner and Kock (2018),

$$\begin{aligned}\max_{1\leq j\leq p}|\hat{\tau}_j^2 - \tau_j^2| &\leq \max_{1\leq j\leq p}|\frac{\eta_j'\eta_j}{n} - \tau_j^2| + \max_{1\leq j\leq p}|\eta_j'\mathbf{r}^*{}_{-j}(\hat{\gamma}_j - \gamma_j)/n| \\ &+ \max_{1\leq j\leq p}|\eta_j'\mathbf{r}^*{}_{-j}\gamma_j/n| + \max_{1\leq j\leq p}|\gamma_j'\frac{\mathbf{r}^{*\prime}{}_{-j}\mathbf{r}^*{}_{-j}}{n}(\hat{\gamma}_j - \gamma_j)|. \end{aligned} \tag{A.7}$$

Consider the second term in (A.7)

$$\begin{aligned}\max_{1\leq j\leq p}|\eta_j'\mathbf{r}^*{}_{-j}(\hat{\gamma}_j - \gamma_j)/n| &\leq \max_{1\leq j\leq p}\|\eta_j'\mathbf{r}^*{}_{-j}/n\|_\infty \max_{1\leq j\leq p}\|\hat{\gamma}_j - \gamma_j\|_1 \\ &= O_p(\sqrt{\frac{logp}{n}})O_p(\bar{s}\sqrt{\frac{logp}{n}}) = O_p(\bar{s}\frac{logp}{n}), \end{aligned} \tag{A.8}$$

by (A.5) and (A.2). Given (A.3)(A.4), via Assumption 2, and using the analysis in (B.48) of Caner and Kock (2018) we have

$$\|\gamma_j\|_1 = O(\sqrt{\bar{s}}). \tag{A.9}$$

Now we consider the third term in (A.7)

$$\begin{aligned}\max_{1\leq j\leq p}|\eta_j'\mathbf{r}^*{}_{-j}\gamma_j/n| &\leq \max_{1\leq j\leq p}\|\eta_j'\mathbf{r}^*{}_{-j}/n\|_\infty \max_{1\leq j\leq p}\|\gamma_j\|_1 \\ &= O_p(\sqrt{\frac{logp}{n}})O_p(\sqrt{\bar{s}}) = O_p(\sqrt{\bar{s}\frac{logp}{n}}), \end{aligned} \tag{A.10}$$

by (A.5) and (A.9). Next consider the fourth term in (A.7). Before that there is a simplification due to first order conditions in nodewise regression as in (B.49)-(B.50) of Caner and

Kock (2018)

$$\max_{1\leq j\leq p}\|\frac{\mathbf{r}^{*\prime}_{-j}\mathbf{r}^{*}_{-j}}{n}(\hat{\gamma}_j-\gamma_j)\|_\infty\leq\max_{1\leq j\leq p}\|\mathbf{r}^{*\prime}_{-j}\eta_j/n\|_\infty+\max_{1\leq j\leq p}\lambda_j.$$

Then using the above inequality

$$\begin{aligned}\max_{1\leq j\leq p}|\gamma_j'\frac{\mathbf{r}^{*\prime}_{-j}\mathbf{r}^{*}_{-j}}{n}(\hat{\gamma}_j-\gamma_j)| &\leq \max_{1\leq j\leq p}\|\gamma_j\|_1\max_{1\leq j\leq p}\|\frac{\mathbf{r}^{*\prime}_{-j}\mathbf{r}^{*}_{-j}}{n}(\hat{\gamma}_j-\gamma_j)\|_\infty \\ &= O(\sqrt{\bar{s}})O_p(\sqrt{\frac{logp}{n}})=O_p(\sqrt{\bar{s}\frac{logp}{n}}),\end{aligned}\tag{A.11}$$

where we use (A.9) and (A.5) with the algebraic analysis in (B.50) of Caner and Kock (2018) and with $\lambda_j=O(\sqrt{\frac{logp}{n}})$ for all $j=1,\cdots,p$. Combine (A.6) (A.8)-(A.11) in (A.7) to have

$$\max_{1\leq j\leq p}|\hat{\tau}_j^2-\tau_j^2|=O_p(\sqrt{\bar{s}\frac{logp}{n}}).\tag{A.12}$$

After this, given the rates in (A.2) and (A.12), via the definition of $\hat{\Theta}_j$, by repeating the same analysis in (B.51)-(B.53) of Caner and Kock (2018), we have

$$\max_{1\leq j\leq p}\|\hat{\Theta}_j-\Theta_j\|_1=O_p(\bar{s}\sqrt{\frac{logp}{n}}).$$

Q.E.D.

(ii). The analysis in Lemma 1 of Chang et al. (2019) provides the desired result.**Q.E.D.**

(iii). Given (A.3)-(A.4), Assumption 2, and the analysis in (B.55) of Caner and Kock (2018) provides the result. **Q.E.D.**

The following Lemma A.1 is useful for the proof of Theorem 3.1. We define $\hat{A}=1_p'\hat{\Theta}1_p/p$, also note that $A=1_p'\Theta 1_p/p$, where the population quantity $\Theta=\Sigma^{-1}$.

Lemma A.1. *Under Assumptions 1-4, uniformly in* $j\in\{1,\cdots,p\}$,

$$|\hat{A}-A|=o_p(1).$$

Proof of Lemma A.1. First, see that

$$\hat{A}-A=(1_p'\hat{\Theta}1_p-1_p'\Theta 1_p)/p=(1_p'(\hat{\Theta}-\Theta)1_p)/p.\tag{A.13}$$

Now consider the the right side of (A.13)

$$\begin{aligned} |1_p'(\hat{\Theta} - \Theta)1_p|/p &\leq \|(\hat{\Theta} - \Theta)1_p\|_1 \|1_p\|_\infty/p \\ &\leq \max_{1 \leq j \leq p} \|\hat{\Theta}_j - \Theta_j\|_1 \\ &= O_p(\bar{s}\sqrt{\text{log}p/n}) = o_p(1), \end{aligned} \tag{A.14}$$

where Holders inequality is used in the first inequality, and (A.1) is used for the second inequality and the last equality is obtained by using Theorem A.1(i), and imposing Assumption 4. **Q.E.D.**

Before the proof of main theorem, below, we show a property of $\hat{\Theta}$. Take a $p \times 1$ vector $\delta \neq 0$ (all constants) then following exactly the same proof as in Lemma A.1 above

$$|\delta'\hat{\Theta}\delta/p - \delta'\Sigma^{-1}\delta/p| = o_p(1), \tag{A.15}$$

uniformly in $j = 1, \cdots p$, and $\hat{\Theta} = (\hat{\Theta}_1', \cdots, \hat{\Theta}_j', \cdots, \hat{\Theta}_p')$. We provide a proof by contradiction. Assume that there exists $\delta \neq 0$ such that $\delta'\hat{\Theta}\delta/p \leq 0$, then (A.15) shows that we should have had $\delta'\Sigma^{-1}\delta/p \leq 0$, with probability approaching one, but by Assumption $\Theta = \Sigma^{-1}$ is positive definite then $\delta'\hat{\Theta}\delta \leq 0$ is not possible with probability approaching one.

Proof of Theorem 3.1. We consider

$$\left| \frac{\hat{A}^{-1}}{A^{-1}} - 1 \right| = \frac{|A - \hat{A}|}{|\hat{A}|}. \tag{A.16}$$

First, use Assumption 2 to have, (where $C_0 = Eigmin(\Sigma^{-1}) > 0, C_0$ is a positive constant, and it represents the minimal eigenvalue of $\Theta = \Sigma^{-1}$)

$$A = 1_p'\Sigma^{-1}1_p/p \geq C_0 > 0,$$

which shows

$$A \geq C_0 > 0. \tag{A.17}$$

By Lemma A.1 and its proof we have $|\hat{A} - A| = O_p(\bar{s}\sqrt{logp/n}) = o_p(1)$. Then use this last equation for the numerator in (A.16)

$$\left| \frac{\hat{A}^{-1}}{A^{-1}} - 1 \right| = \frac{o_p(1)}{(o_p(1) + C_0)} = o_p(1), \tag{A.18}$$

where the denominator is bounded away from zero by A being bounded away from zero as shown in (A.17). Also use $\hat{A} = A + o_p(1)$ by Lemma A.1 to get the denominator's rate and the result. **Q.E.D.**

Before the next Lemma, we define $\hat{B} = 1_p'\hat{\Theta}\hat{\mu}/p$, and $B = 1_p'\Theta\mu/p$.

Lemma A.2. *Under Assumptions 1-4, uniformly in* $j \in \{1, \cdots, p\}$

$$|\hat{B} - B| = o_p(1).$$

Proof of Lemma A.2. We can decompose $\hat{B}$ by simple addition and subtraction into

$$\begin{aligned} \hat{B} - B &= [1_p'(\hat{\Theta} - \Theta)(\hat{\mu} - \mu)]/p && \text{(A.19)} \\ &+ [1_p'(\hat{\Theta} - \Theta)\mu]/p && \text{(A.20)} \\ &+ [1_p'\Theta(\hat{\mu} - \mu)]/p && \text{(A.21)} \end{aligned}$$

Now we analyze each of the terms above. Since $\hat{\mu} = n^{-1} \sum_{t=1}^{n} r_t$,

$$\begin{aligned} |1_p'(\hat{\Theta} - \Theta)(\hat{\mu} - \mu)|/p &\leq \|(\hat{\Theta} - \Theta)1_p\|_1 \|\hat{\mu} - \mu\|_\infty / p \\ &\leq [\max_{1 \leq j \leq p} \|\hat{\Theta}_j - \Theta_j\|_1] \|\hat{\mu} - \mu\|_\infty \\ &= O_p(\bar{s}\sqrt{\text{log}p/n}) O_p(\sqrt{\text{log}p/n}), && \text{(A.22)} \end{aligned}$$

where we use Holder's inequality in the first inequality, and the norm inequality in (A.1) with $M = \hat{\Theta} - \Theta$, $x = 1_p$ in the second inequality above, and the rate is by Theorem A.1(i)-(ii).

So we consider (A.20) above. Since we assume in Section 2, $\|\mu\|_\infty \leq C < \infty$, where C is a positive constant.

$$\begin{aligned} |1_p'(\hat{\Theta} - \Theta)\mu|/p &\leq \|(\hat{\Theta} - \Theta)1_p\|_1 \|\mu\|_\infty / p \\ &\leq C[\max_{1 \leq j \leq p} \|\hat{\Theta}_j - \Theta_j\|_1] \\ &= C O_p(\bar{s}\sqrt{\text{log}p/n}), && \text{(A.23)} \end{aligned}$$

where we use Holder's inequality in the first inequality, and the norm inequality in (A.1) with $M = \hat{\Theta} - \Theta$, $x = 1_p$ in the second inequality above, and the rate is by Theorem A.1(i).

Now consider (A.21).

$$\begin{aligned} |1_p'\Theta(\hat{\mu} - \mu)|/p &\leq \|\Theta 1_p\|_1 \|\hat{\mu} - \mu\|_\infty / p \\ &\leq [\max_{1 \leq j \leq p} \|\Theta_j\|_1] \|\hat{\mu} - \mu\|_\infty \\ &= O(\sqrt{\bar{s}}) O_p(\sqrt{\text{log}p/n}), && \text{(A.24)} \end{aligned}$$

where we use Holder's inequality in the first inequality, and the norm inequality in (A.1) with $M = \Theta$, $x = 1_p$ in the second inequality above, and the rate is from Theorem A.1(ii)-(iii). Combine (A.22)(A.23)(A.24) in (A.19)-(A.21), and note that the rate is coming from (A.23). So use Assumption 4, $\bar{s}\sqrt{logp/n} = o(1)$ to have

$$|\hat{B} - B| = O_p(\bar{s}\sqrt{logp/n}) = o_p(1). \tag{A.25}$$

.**Q.E.D.**

Next, we show the uniform consistency of another term in the estimated optimal weights. Note that $D = \mu'\Theta\mu/p$, and its estimator is $\hat{D} = \hat{\mu}'\hat{\Theta}\hat{\mu}/p$.

Lemma A.3. *Under Assumptions 1-4, uniformly in* $j \in \{1, \cdots, p\}$

$$|\hat{D} - D| = o_p(1).$$

Proof of Lemma A.3. By simple addition and subtraction

$$\begin{aligned}
\hat{D} - D &= [(\hat{\mu}-\mu)'(\hat{\Theta}-\Theta)(\hat{\mu}-\mu)]/p && \text{(A.26)}\\
&+ [(\hat{\mu}-\mu)'\Theta(\hat{\mu}-\mu)]/p && \text{(A.27)}\\
&+ [2(\hat{\mu}-\mu)'\Theta\mu]/p && \text{(A.28)}\\
&+ [2\mu'(\hat{\Theta}-\Theta)(\hat{\mu}-\mu)]/p && \text{(A.29)}\\
&+ [\mu'(\hat{\Theta}-\Theta)\mu]/p. && \text{(A.30)}
\end{aligned}$$

We start with (A.26).

$$\begin{aligned}
|(\hat{\mu}-\mu)'(\hat{\Theta}-\Theta)(\hat{\mu}-\mu)|/p &\leq \|(\hat{\Theta}-\Theta)(\hat{\mu}-\mu)\|_1\|\hat{\mu}-\mu\|_\infty/p\\
&\leq [\|\hat{\mu}-\mu\|_\infty]^2[\max_j \|\hat{\Theta}_j - \Theta_j\|_1]\\
&= O_p(\text{log}p/n)O_p(\bar{s}\sqrt{\text{log}p/n})\\
&= O_p(\bar{s}(\text{log}p/n)^{3/2}), && \text{(A.31)}
\end{aligned}$$

where Holder's inequality is used for the first inequality above, and the inequality (A.1), with $M = \hat{\Theta} - \Theta$ and $x = \hat{\mu} - \mu$ for the second inequality above, and for the rates we use Theorem A.1(i)-(ii).

We continue with (A.27).

$$\begin{aligned}
|(\hat{\mu}-\mu)'(\Theta)(\hat{\mu}-\mu)|/p &\leq \|(\Theta)(\hat{\mu}-\mu)\|_1\|\hat{\mu}-\mu\|_\infty/p\\
&\leq [\|\hat{\mu}-\mu\|_\infty]^2[\max_j \|\Theta_j\|_1]\\
&= O_p(\text{log}p/n)O(\sqrt{\bar{s}})\\
&= O_p(\sqrt{\bar{s}}(\text{log}p/n)), && \text{(A.32)}
\end{aligned}$$

where Holder's inequality is used for the first inequality above, and the inequality (A.1), with $M = \Theta$ and $x = \hat{\mu} - \mu$ for the second inequality above, for the rates we use Theorem A.1(ii)-(iii).

Then we consider (A.28), with using $\|\mu\|_\infty \leq C$,

$$\begin{aligned}
|(\hat{\mu}-\mu)'(\Theta)(\mu)|/p &\leq \|(\Theta)(\hat{\mu}-\mu)\|_1 \|\mu\|_\infty /p \\
&\leq C[\|\hat{\mu}-\mu\|_\infty][\max_j \|\Theta_j\|_1] \\
&= O_p(\sqrt{\text{log} p/n}) O(\sqrt{\bar{s}}) \\
&= O_p(\sqrt{\bar{s}}\sqrt{logp/n}), \qquad (A.33)
\end{aligned}$$

where Holder's inequality is used for the first inequality above, and the inequality (A.1), with $M = \Theta$ and $x = \hat{\mu} - \mu$ for the second inequality above, for the rates we use Theorem A.1(ii)-(iii).

Then we consider (A.29).

$$\begin{aligned}
|(\mu)'(\hat{\Theta}-\Theta)(\hat{\mu}-\mu)|/p &\leq \|(\hat{\Theta}-\Theta)(\mu)\|_1 \|\hat{\mu}-\mu\|_\infty /p \\
&\leq \|\mu\|_\infty \max_j \|\hat{\Theta}_j - \Theta_j\|_1 \|\hat{\mu}-\mu\|_\infty \\
&\leq C[\max_j \|\hat{\Theta}_j - \Theta_j\|_1]\|(\hat{\mu}-\mu)\|_\infty \\
&= O_p(\bar{s}\sqrt{\text{log} p/n}) O_p(\sqrt{\text{log} p/n}) \\
&= O_p(\bar{s}\text{log} p/n), \qquad (A.34)
\end{aligned}$$

where Holder's inequality is used for the first inequality above, and the inequality (A.1), with $M = \hat{\Theta} - \Theta$ and $x = \mu$ for the second inequality above, and for the third inequality above we use $\|\mu\|_\infty \leq C$, and for the rates we use Theorem A.1(i)-(ii).

Then we consider (A.30),

$$\begin{aligned}
|(\mu)'(\hat{\Theta}-\Theta)(\mu)|/p &\leq \|(\hat{\Theta}-\Theta)(\mu)\|_1 \|\mu\|_\infty /p \\
&\leq [\|\mu\|_\infty]^2 \max_j \|\hat{\Theta}_j - \Theta_j\|_1 \\
&\leq C[\max_j \|\hat{\Theta}_j - \Theta_j\|_1] \\
&= O_p(\bar{s}\sqrt{logp/n}), \qquad (A.35)
\end{aligned}$$

where Holder's inequality is used for the first inequality above, and the inequality (A.1), with $M = \hat{\Theta} - \Theta$ and $x = \mu$ for the second inequality above, and for the third inequality above we use $\|\mu\|_\infty \leq C$, and for the rate we use Theorem A.1(i). Note that the last rate above in (A.35) derives our result, since it is the largest rate by Assumption 4.

Combine (A.31)-(A.35) in (A.26)-(A.30) and the rate in (A.35) to have

$$|\hat{D} - D| = O_p(\bar{s}\sqrt{logp/n}) = o_p(1). \qquad (A.36)$$

Q.E.D.

The following lemma establishes orders for the terms in the optimal weight, A, B, D. Note that both A, D are positive by Assumption 2, and uniformly bounded away from zero.

Lemma A.4. *Under Assumption 2*

$$A = O(1).$$

$$|B| = O(1).$$

$$D = O(1).$$

Proof of Lemma A.4. We do the proof for term $D = \mu'\Theta\mu/p$. The proof for $A = 1_p'\Theta 1_p/p$ is the same.

$$D = \mu'\Theta\mu/p \leq Eigmax(\Theta)\|\mu\|_2^2/p = O(1),$$

where we use the fact that maximum of μ_j is a constant as assumed in Section 2, and the maximal eigenvalue of $\Theta = \Sigma^{-1}$ is finite by Assumption 2. For term B, the proof can be obtained by using Cauchy-Schwartz inequality first and the using the same analysis for terms A and D.**Q.E.D.**

Next we need the following technical lemma, that provides the limit and the rate for the denominator in optimal portfolio.

Lemma A.5. *Under Assumptions 1-4, uniformly over j in $\lambda_j = O(\sqrt{logp/n})$*

$$|(\hat{A}\hat{D} - \hat{B}^2) - (AD - B^2)| = o_p(1).$$

Proof of Lemma A.5. Note that by simple adding and subtracting

$$\hat{A}\hat{D} - \hat{B}^2 = [(\hat{A} - A) + A][(\hat{D} - D) + D] - [(\hat{B} - B) + B]^2.$$

Then using this last expression and simplifying, A, D being both positive

$$\begin{aligned} |(\hat{A}\hat{D} - \hat{B}^2) - (AD - B^2)| &\leq \{|\hat{A} - A||\hat{D} - D| + |\hat{A} - A|D \\ &+ A|\hat{D} - D| + (\hat{B} - B)^2 + 2|B||\hat{B} - B|\} \\ &= O_p(\bar{s}\sqrt{logp/n}) = o_p(1), \end{aligned} \tag{A.37}$$

where we use (A.14)(A.25)(A.36), Lemma A.4, and Assumption 4: $\bar{s}\sqrt{logp/n} = o(1)$.**Q.E.D.**

Proof of Theorem 3.2. Now we define notation to help us in the proof here. First set

$$\hat{x} = \hat{A}\rho_1^2 - 2\hat{B}\rho_1 + \hat{D}. \tag{A.38}$$

$$x = A\rho_1^2 - 2B\rho_1 + D. \tag{A.39}$$

$$\hat{y} = \hat{A}\hat{D} - \hat{B}^2. \tag{A.40}$$

$$y = AD - B^2. \tag{A.41}$$

Then we can write the estimate of the optimal portfolio variance as

$$\hat{\Psi}_{OPV} = p^{-1}[\frac{\hat{x}}{\hat{y}}],$$

and the optimal portfolio variance is

$$\Psi_{OPV} = p^{-1}[\frac{x}{y}].$$

To start the main part of the proof we need a rate for a limit fraction: y/x. Note that the fraction is positive by Assumptions $AD - B^2 \geq C_1 > 0, A\rho_1^2 - 2B\rho_1 + D \geq C_1 > 0$.

$$\begin{aligned} \frac{y}{x} &= \frac{AD - B^2}{A\rho_1^2 - 2B\rho_1 + D} \leq \frac{AD}{A\rho_1^2 - 2B\rho_1 + D} \\ &= O(1), \end{aligned} \tag{A.42}$$

where we use $B^2 > 0$ and the assumption $A\rho_1^2 - 2B\rho_1 + D \geq C_1 > 0$ and Lemma A.4.

So we can setup the problem as, by adding and subtracting xy from the numerator, and $y/x > 0$ by assumption, and use (A.42) for the second equality below

$$\begin{aligned} \left| \frac{\hat{\Psi}_{OPV} - \Psi_{OPV}}{\Psi_{OPV}} \right| &= \left| \frac{\hat{x}}{\hat{y}} - \frac{x}{y} \right| \frac{y}{x} \\ &= O(1) \left| \frac{\hat{x}}{\hat{y}} - \frac{x}{y} \right| \\ &= \left| \frac{\hat{x}y - xy + xy - x\hat{y}}{\hat{y}y} \right| O(1) \\ &= \left| \frac{(\hat{x} - x)y + x(y - \hat{y})}{\hat{y}y} \right| O(1). \end{aligned} \tag{A.43}$$

We consider each term in the numerator in (A.43). Via Lemma A.1-A.3, and ρ_1 being bounded, and (A.14)(A.25)(A.36)

$$\begin{aligned} |\hat{x} - x| &= |\hat{A}\rho_1^2 - 2\hat{B}\rho_1 + \hat{D} - (A\rho_1^2 - 2B\rho_1 + D)| \\ &\leq \{|\hat{A} - A|\rho_1^2 + 2|\hat{B} - B|\rho_1 + |\hat{D} - D|\} \\ &= O_p(\bar{s}\sqrt{logp/n}) = o_p(1), \end{aligned} \tag{A.44}$$

where we use Assumption 4 in the rate above. Now analyze the following term in the numerator

$$y = AD - B^2 \leq AD = O(1), \tag{A.45}$$

where we use $B^2 > 0$ in the inequality, and Lemma A.4 for the rate result, which is the final equality above in (A.45). Next, consider the following in the numerator

$$x = A\rho_1^2 - 2B\rho_1 + D \leq (A\rho_1^2 + 2|B|\rho_1 + D) = O(1), \tag{A.46}$$

where we use A, D being positive, and Lemma A.4, with ρ_1 being uniformly bounded away from infinity. Then Lemma A.5 and (A.37) provides

$$|\hat{y} - y| = |\hat{A}\hat{D} - \hat{B}^2 - (AD - B^2)| = O_p(\bar{s}\sqrt{logp/n}) = o_p(1). \tag{A.47}$$

So the numerator in (A.43) is,

$$|\hat{x} - x|y + x|y - \hat{y}| = O_p(\bar{s}\sqrt{logp/n}) = o_p(1), \tag{A.48}$$

where we use (A.44)-(A.47) and $x > 0, y > 0$.

We consider the denominator in (A.43)

$$\begin{aligned} |\hat{y}y| &= \big|[(\hat{y} - y) + y]y\big| \\ &= \big(o_p(1) + y\big)\, y \\ &\geq (o_p(1) + C_1)C_1 > 0, \end{aligned} \tag{A.49}$$

where we add and subtract y in the first equality, and use Lemma A.5 in the second equality, and $y = (AD - B^2) \geq C_1 > 0$, and C_1 is a positive constant by assumption. Next, combine (A.48)(A.49) in (A.43) with Assumption 4 to have

$$\left| \frac{\hat{\Psi}_{OPV} - \Psi_{OPV}}{\Psi_{OPV}} \right| \leq \frac{O_p(\bar{s}\sqrt{logp/n})}{C_1^2 + o_p(1)} = \frac{o_p(1)}{C_1^2 + o_p(1)} = o_p(1).$$

Q.E.D.

Proof of Theorem 3.3. Using (3.3)-(3.4) and via adding and subtracting $A\Theta 1_p/p$ from the numerator below

$$\begin{aligned} \hat{w}_u - w_u &= \frac{[(A\hat{\Theta}1_p) - (\hat{A}\Theta 1_p)]/p}{(\hat{A}A)} \\ &= \frac{[(A\hat{\Theta}1_p) - (A\Theta 1_p) + (A\Theta 1_p) - (\hat{A}\Theta 1_p)]/p}{(\hat{A}A)}. \end{aligned}$$

Using the above result

$$\|\hat{w}_u - w_u\|_1 \leq \frac{A\frac{\|(\hat{\Theta}-\Theta)1_p\|_1}{p} + |A - \hat{A}|\frac{\|\Theta 1_p\|_1}{p}}{|\hat{A}|A}. \tag{A.50}$$

Then in (A.50) consider the numerator. To that effect, analyze the terms below.

$$
\begin{aligned}
\|(\hat{\Theta} - \Theta)1_p\|_1/p &\le \max_{1\le j\le p} \|\hat{\Theta}_j - \Theta_j\|_1 \\
&= O_p(\bar{s}\sqrt{\text{log}p/n}),
\end{aligned} \tag{A.51}
$$

where we use (A.1) for the inequality, and Theorem A.1(i) for the rate result in (A.51). Now we analyze

$$
\|\Theta 1_p\|_1/p \le \max_{1\le j\le p} \|\Theta_j\|_1 = O(\sqrt{\bar{s}}), \tag{A.52}
$$

where the inequality is obtained by (A.1), and the rate is by Theorem A.1(iii).
Via Lemma A.4, $A = O(1)$, also by (A.13),(A.14)

$$
|\hat{A} - A| = O_p(\bar{s}\sqrt{logp/n}). \tag{A.53}
$$

By (A.51)-(A.53)

$$
\begin{aligned}
A\frac{\|(\hat{\Theta} - \Theta)1_p\|_1}{p} + |A - \hat{A}|\frac{\|\Theta 1_p\|_1}{p} &= O(1)O_p(\bar{s}\sqrt{logp/n}) \\
&+ O_p(\bar{s}\sqrt{logp/n})O(\sqrt{\bar{s}}) \\
&= O_p((\bar{s})^{3/2}\sqrt{logp/n}) = o_p(1),
\end{aligned} \tag{A.54}
$$

where we use sparsity assumption $(\bar{s})^{3/2}\sqrt{logp/n} = o(1)$ in the last step. Then for the denominator in (A.50) from (A.16)-(A.18) we have, for $C_0 > 0$, is a positive constant,

$$
|\hat{A}|A \ge (o_p(1) + C_0)C_0. \tag{A.55}
$$

Now combine (A.54)(A.55) in (A.50) to have the desired result. **Q.E.D.**

Proof of Theorem 3.4. Denote $w^* = [\Delta_1(\Theta 1_p/p) + \Delta_2(\Theta\mu/p)]$, where

$$
\Delta_1 = \frac{D - \rho_1 B}{AD - B^2},
$$

$$
\Delta_2 = \frac{\rho_1 A - B}{AD - B^2}.
$$

Next, denote $\hat{w} = [\hat{\Delta}_1(\hat{\Theta}1_p/p) + \hat{\Delta}_2(\hat{\Theta}\hat{\mu}/p)]$, where $\hat{\Delta}_1, \hat{\Delta}_2$ represent estimators for Δ_1, Δ_2 respectively. We get $\hat{\Delta}_1, \hat{\Delta}_2$ by replacing A, B, D, in the formula for Δ_1, Δ_2 with their estimators shown in above Theorems. Next, by adding and subtracting

$$
\begin{aligned}
\hat{w} - w^* &= [\hat{\Delta}_1(\hat{\Theta}1_p/p) + \hat{\Delta}_2(\hat{\Theta}\hat{\mu}/p) - \Delta_1(\Theta 1_p/p) - \Delta_2(\Theta\mu/p)] \\
&= [(\hat{\Delta}_1 - \Delta_1) + \Delta_1][(\hat{\Theta} - \Theta) + \Theta]1_p/p \\
&+ [(\hat{\Delta}_2 - \Delta_2) + \Delta_2][(\hat{\Theta} - \Theta) + \Theta][(\hat{\mu} - \mu) + \mu]/p \\
&- [\Delta_1(\Theta 1_p/p) - \Delta_2(\Theta\mu/p)] \\
&= (\hat{\Delta}_1 - \Delta_1)(\hat{\Theta} - \Theta)1_p/p + (\hat{\Delta}_1 - \Delta_1)\Theta 1_p/p \\
&+ \Delta_1(\hat{\Theta} - \Theta)1_p/p + (\hat{\Delta}_2 - \Delta_2)(\hat{\Theta} - \Theta)(\hat{\mu} - \mu)/p \\
&+ (\hat{\Delta}_2 - \Delta_2)\Theta(\hat{\mu} - \mu)/p + (\hat{\Delta}_2 - \Delta_2)(\hat{\Theta} - \Theta)\mu/p \\
&+ (\hat{\Delta}_2 - \Delta_2)\Theta\mu/p + \Delta_2(\hat{\Theta} - \Theta)(\hat{\mu} - \mu)/p \\
&+ \Delta_2\Theta(\hat{\mu} - \mu)/p + \Delta_2(\hat{\Theta} - \Theta)\mu/p.
\end{aligned} \tag{A.56}
$$

Using (A.56), and since $\hat{\Delta}_1, \Delta_1, \hat{\Delta}_2, \Delta_2$ are all scalars,

$$
\begin{aligned}
\|\hat{w} - w^*\|_1 &\leq |(\hat{\Delta}_1 - \Delta_1)|\|(\hat{\Theta} - \Theta)1_p\|_1/p + |(\hat{\Delta}_1 - \Delta_1)|\|\Theta 1_p\|_1/p \\
&+ |\Delta_1|\|(\hat{\Theta} - \Theta)1_p\|_1/p + |(\hat{\Delta}_2 - \Delta_2)|\|(\hat{\Theta} - \Theta)(\hat{\mu} - \mu)\|_1/p \\
&+ |(\hat{\Delta}_2 - \Delta_2)|\|\Theta(\hat{\mu} - \mu)\|_1/p + |(\hat{\Delta}_2 - \Delta_2)|\|(\hat{\Theta} - \Theta)\mu\|_1/p \\
&+ |(\hat{\Delta}_2 - \Delta_2)|\|\Theta\mu\|_1/p + |\Delta_2|\|(\hat{\Theta} - \Theta)(\hat{\mu} - \mu)\|_1/p \\
&+ |\Delta_2|\|\Theta(\hat{\mu} - \mu)\|_1/p + |\Delta_2|\|(\hat{\Theta} - \Theta)\mu\|_1/p.
\end{aligned} \tag{A.57}
$$

We consider each term above. But rather than analyzing them one by one, we analyze common elements and then determine the order of each term on the right side of (A.57). Using the definitions of $\hat{y}, y$ in (A.40)(A.41) respectively, and adding and subtracting $y(D - \rho_1 B)$ respectively from the numerator, with ρ_1 being bounded, $y > 0$ by assumption

$$
\begin{aligned}
|\hat{\Delta}_1 - \Delta_1| &= \left| \frac{y(\hat{D} - \rho_1\hat{B}) - \hat{y}(D - \rho_1 B)}{\hat{y}y} \right| \\
&= \left| \frac{y(\hat{D} - \rho_1\hat{B}) - y(D - \rho_1 B) + y(D - \rho_1 B) - \hat{y}(D - \rho_1 B)}{\hat{y}y} \right| \\
&\leq \frac{y|(\hat{D} - D)| + y\rho_1|(\hat{B} - B)| + |(y - \hat{y})||(D - \rho_1 B)|}{|\hat{y}y|}.
\end{aligned} \tag{A.58}
$$

Now we analyze each term in the numerator. By Lemma A.4, with $y > 0$ by assumption

$$
y = AD - B^2 \leq AD = O(1). \tag{A.59}
$$

Next, by (A.25)(A.36)

$$
y|\hat{D} - D| + y|\rho_1||\hat{B} - B| = O_p(\bar{s}\sqrt{logp/n}). \tag{A.60}
$$

Then

$$|y-\hat{y}||D-\rho_1 B| \leq \left(|\hat{A}\hat{D}-\hat{B}^2-(AD-B^2)|\right)\left(D+|\rho_1||B|\right) = O_p(\bar{s}\sqrt{logp/n}), \quad (A.61)$$

where we use $y, \hat{y}$ definitions in the inequality, and to get the rate Lemma A.4 with (A.37) is used. Combine now (A.60)(A.61) in the numerator in (A.58) to have

$$y|\hat{D}-D| + y|\rho_1||\hat{B}-B| + |y-\hat{y}||D-\rho_1 B| = O_p(\bar{s}\sqrt{logp/n}). \quad (A.62)$$

Then combine (A.49)(A.62) to have

$$|\hat{\Delta}_1 - \Delta_1| = O_p(\bar{s}\sqrt{logp/n}). \quad (A.63)$$

Exactly following the same way we derive

$$|\hat{\Delta}_2 - \Delta_2| = O_p(\bar{s}\sqrt{logp/n}). \quad (A.64)$$

Consider, by using $AD - B^2 \geq C_1 > 0$ by assumption

$$\begin{aligned} |\Delta_1| &= \left|\frac{D-\rho_1 B}{AD-B^2}\right| \leq \left|\frac{D-\rho_1 B}{C_1}\right| \\ &\leq \frac{|D| + |\rho_1||B|}{C_1} = O(1), \quad (A.65) \end{aligned}$$

where we use Lemma A.4 to have $D = O(1), |B| = O(1)$, and ρ_1 being bounded. In the same way we obtain

$$|\Delta_2| = O(1). \quad (A.66)$$

Next, we consider the following term:

$$\begin{aligned} \|(\hat{\Theta}-\Theta)(\hat{\mu}-\mu)\|_1/p &\leq \|(\hat{\mu}-\mu)\|_\infty \max_{1\leq j\leq p} \|\hat{\Theta}_j - \Theta_j\|_1 \\ &= O_p(\bar{s}\mathrm{log}p/n), \quad (A.67) \end{aligned}$$

where we use (A.1) for the first inequality, and the rate is derived from Theorem A.1(i). Then given $\|\mu\|_\infty \leq C$

$$\begin{aligned} \|(\hat{\Theta}-\Theta)\mu\|_1/p &\leq C \max_{1\leq j\leq p} \|\hat{\Theta}_j - \Theta_j\|_1 \\ &= O_p(\bar{s}\sqrt{\mathrm{log}p/n}), \quad (A.68) \end{aligned}$$

where we use (A.1) for the first inequality, and the rate is derived from Theorem A.1(i). Note that

$$\|\Theta\mu\|_1/p = O(\sqrt{\bar{s}}), \quad (A.69)$$

where we use the same analysis in (A.52).

Next,

$$\begin{aligned}\|\Theta(\hat{\mu}-\mu)\|_1/p &\leq \|\hat{\mu}-\mu\|_\infty \max_{1\leq j\leq p}\|\Theta_j\|_1 \\ &= O_p(\sqrt{\text{log}p/n})O(\sqrt{\bar{s}}) \\ &= O_p(\sqrt{\bar{s}}\sqrt{\text{log}p/n}), \end{aligned} \tag{A.70}$$

where we use (A.1) for the first inequality and Theorem A.1(ii)-(iii) for rates.

Use (A.51)(A.52), (A.63)-(A.70) in (A.57) to have

$$\begin{aligned}\|\hat{w}-w^*\|_1 &= O_p((\bar{s})^2\text{log}p/n) + O_p((\bar{s})^{3/2}\sqrt{logp/n}) \\ &+ O_p(\bar{s}\sqrt{logp/n}) + O_p((\bar{s})^2(\text{log}p/n)^{3/2}) \\ &+ O_p((\bar{s})^{3/2}\text{log}p/n) + O_p((\bar{s})^2(\text{log}p/n)) \\ &+ O_p((\bar{s})^{3/2}\sqrt{logp/n}) + O_p((\bar{s})(\text{log}p/n)) \\ &+ O_p(\sqrt{\bar{s}}\sqrt{logp/n}) + O_p((\bar{s})\sqrt{logp/n}) \\ &= O_p((\bar{s})^{3/2}\sqrt{logp/n}) = o_p(1), \end{aligned} \tag{A.71}$$

where we use the fact that $(\bar{s})^{3/2}\sqrt{logp/n}$ is the slowest rate of convergence on the right hand side terms. Then by $\bar{s}^{3/2}\sqrt{logp/n} = o(1)$ we have the last result.**Q.E.D**

Proof of Theorem 3.5. We consider

$$|\hat{w}_u'(\hat{\Sigma}-\Sigma)\hat{w}_u| \leq \|\hat{w}_u\|_1^2\|\hat{\Sigma}-\Sigma\|_\infty. \tag{A.72}$$

In (A.72) we analyze each right side term. First,

$$\|\hat{w}_u\|_1 \leq \|\hat{w}_u - w_u\|_1 + \|w_u\|_1. \tag{A.73}$$

Then from the definition of global minimum variance portfolio

$$\|w_u\|_1 = \frac{\|\Theta 1_p\|_1/p}{1_p'\Theta 1_p/p}. \tag{A.74}$$

Apply (A.17)(A.52) in (A.74) to have

$$\|w_u\|_1 \leq \frac{O(\sqrt{\bar{s}})}{C_0} = O(\sqrt{\bar{s}}). \tag{A.75}$$

Then use (A.75) and Theorem 3.3 in (A.73) to have

$$\|\hat{w}_u\|_1 = O_p((\bar{s})^{3/2}\sqrt{logp/n}) + O(\sqrt{\bar{s}}) = o_p(1) + O(\sqrt{\bar{s}}) = O_p(\sqrt{\bar{s}}), \tag{A.76}$$

where we use Assumption that $(\bar{s}^{3/2}\sqrt{logp/n}) = o(1)$ in the second equality.

Then use Lemma 1 of Chang et al. (2019) to have $\|\hat{\Sigma}-\Sigma\|_\infty = O_p(\sqrt{logp/n})$ and (A.76) in (A.72) to have

$$|\hat{w}_u'(\hat{\Sigma} - \Sigma)\hat{w}_u| \leq O_p(\bar{s}) O_p(\sqrt{logp/n}) = o_p(1), \tag{A.77}$$

where we use Assumption that $(\bar{s})^{3/2}\sqrt{logp/n} = o(1)$ in the second equality.**Q.E.D.**

Proof of Theorem 3.6. We consider

$$|\hat{w}'(\hat{\Sigma} - \Sigma)\hat{w}| \leq \|\hat{w}\|_1^2 \|\hat{\Sigma} - \Sigma\|_\infty. \tag{A.78}$$

In (A.78) we analyze each right side term. First,

$$\|\hat{w}\|_1 \leq \|\hat{w} - w^*\|_1 + \|w^*\|_1. \tag{A.79}$$

Then from the definition of Markowitz portfolio

$$\begin{aligned} \|w^*\|_1 &\leq \frac{|D - \rho_1 B|}{AD - B^2}\|\Theta 1_p\|_1/p + \frac{\rho_1 A - B}{AD - B^2}\|\Theta\mu\|_1/p \\ &\leq \frac{|D| + |\rho_1||B||}{AD - B^2}\|\Theta 1_p\|_1/p + \frac{|\rho_1||A| + |B|}{AD - B^2}\|\Theta\mu\|_1/p. \end{aligned} \tag{A.80}$$

On the right side of (A.80) above, we use the analysis in (A.52), (A.69) for $\|\Theta 1_p\|_1/p$, $\|\Theta\mu\|_1/p$, and ρ_1 is bounded, and by Lemma A.4 with assumption $AD - B^2 \geq C_1 > 0$, to have

$$\|w^*\|_1 \leq \frac{O(1)}{C_1} O(\sqrt{\bar{s}}) + \frac{O(1)}{C_1} O(\sqrt{\bar{s}}) = O(\sqrt{\bar{s}}). \tag{A.81}$$

The rest of proof follows exactly as in the proof of Theorem 3.5, given the result in Theorem 3.4 to be used in (A.79).**Q.E.D.**

Supplementary Tables

In this part, we show extra tables that are robustness checks for the Tables in the main text. The monthly Table 1 (Supplementary) covers a subset of Table 1 in main text. It starts before the recession of 2008, at August 2006. We provide 4 and 5 year out of sample forecast. We see Nodewise based estimator has the best SR among the others. Specially at 5 year out of sample forecast, the results are striking, with transaction costs, Nodewise has Sharpe Ratio of 0.4587 in Global Minimum Variance Portfolio and 0.3766 in Markowitz portfolio compared to 0.1705 and 0.1254 of POET, and 0.2550,-0.2900 of Ledoit-Wolf based estimator respectively. In daily data we consider periods June 1 2017 to May 31 2018, and July 7 2016 to May 31 2018 in Table 2 (Supplementary). We analyze out of sample forecasts for 252 and 126 days respectively. Number of assets, $p = 442$. So $p < n$ in first part of Table 2 and $p > n$ in the second part of the same Table 2. Both POET and Nodewise do well in terms of SR with transaction costs, depending on time period.

	Global Minimum Portfolio				Markowitz Portfolio			
	Return	Variance	Sharpe	Turnover	Return	Variance	Sharpe	Turnover
In-Sample: Aug 2006-Apr 2014, Out-Of-Sample: May 2014-Apr 2018, $n_I = 93$, $n - n_I = 48$								
without TC								
POET	0.0109	0.00091	0.3622	0.0664	0.0092	0.00095	0.2994	0.1768
NodeWise	0.0102	0.00064	0.4003	0.1651	0.0087	0.00069	0.3313	0.2466
Ledoit-Wolf	0.0098	0.00361	0.1645	0.0812	-0.0300	0.01134	-0.2822	0.1608
with TC								
POET	0.0111	0.00093	0.3636	-	0.0091	0.00097	0.2945	-
NodeWise	0.0101	0.00066	0.3957	-	0.0085	0.00071	0.3192	-
Ledoit-Wolf	0.0098	0.00369	0.1628	-	-0.0309	0.01160	-0.2873	-
In-Sample: Aug 2006-Apr 2013, Out-Of-Sample: May 2013-Apr 2018 $n_I = 81$, $n - n_I = 60$								
without TC								
POET	0.0086	0.00241	0.1753	0.1786	0.0070	0.00275	0.1334	0.2867
NodeWise	0.0118	0.00066	0.4627	0.1624	0.0107	0.00076	0.3876	0.2572
Ledoit-Wolf	0.0164	0.00414	0.2555	0.0797	-0.0367	0.01652	-0.2861	0.1967
with TC								
POET	0.0084	0.00245	0.1705	-	0.0066	0.00280	0.1254	-
NodeWise	0.0118	0.00067	0.4587	-	0.0105	0.00077	0.3766	-
Ledoit-Wolf	0.0165	0.00421	0.2550	-	-0.0376	0.01683	-0.2900	-

Table 1: Monthly Returns-Variance-Sharpe Ratio-Turnover

	Global Minimum Portfolio				Markowitz Portfolio			
	Return	Variance	Sharpe	Turnover	Return	Variance	Sharpe	Turnover
In-Sample: Jun 1 2017-Nov 28 2017, Out-Of-Sample: Nov 29 2017-May 31 2018, $n_I = 126$, $n - n_I = 126$								
without TC								
POET	1.218e-04	7.940e-05	0.0136	0.4532	1.684e-04	7.619e-05	0.0192	0.4224
NodeWise	4.065e-04	8.188e-05	0.0449	0.1542	3.762e-04	7.987e-05	0.0420	0.1753
Ledoit-Wolf	3.479e-04	4.651e-05	0.0510	0.3991	3.955e-04	4.681e-05	0.0578	0.3992
with TC								
POET	-2.633e-04	8.206e-05	-0.0290	-	-1.856e-04	7.850e-05	-0.0209	-
NodeWise	3.192e-04	8.218e-05	0.0352	-	2.678e-04	8.019e-05	0.0299	-
Ledoit-Wolf	2.383e-06	4.711e-05	0.0003	-	5.006e-05	4.762e-05	0.0072	-
In-Sample: Jul 7 2016-Mar 1 2018, Out-Of-Sample: Mar 2 2018-May 31 2018 $n_I = 437$, $n - n_I = 63$								
without TC								
POET	0.000312	7.650e-05	0.0356	0.0262	0.000272	7.108e-05	0.03227	0.0445
NodeWise	0.000302	7.642e-05	0.0345	0.0956	0.000294	7.376e-05	0.03427	0.1056
Ledoit-Wolf	-0.000242	4.061e-05	-0.0380	0.4483	-0.000168	4.071e-05	-0.0263	0.4550
with TC								
POET	0.000443	7.621e-05	0.0508	-	0.000386	7.077e-05	0.0459	-
NodeWise	0.000370	7.596e-05	0.0424	-	0.000353	7.331e-05	0.0412	-
Ledoit-Wolf	-0.000587	4.103e-05	-0.0916	-	-0.000511	4.117e-05	-0.0797	-

Table 2: Daily Returns-Variance-Sharpe Ratio-Turnover

References

Caner, M. and A. Kock (2018): "Asymptotically Honest Confidence Regions for High Dimensional Parameters by the Desparsified Conservative Lasso," *Journal of Econometrics*, 203, 143–168.

Chang, J., Y. Qiu, Q. Yao, and T. Zou (2019): "Confidence regions for entries of a large precision matrix," *Journal of Econometrics, Forthcoming.*