OPERATIONS RESEARCH
DOI 10.1287/opre.1080.0566ec
pp. ec32–ec39

e-companion

ONLY AVAILABLE IN ELECTRONIC FORM

Electronic Companion—“Portfolio Selection with Robust Estimation” by Victor DeMiguel and Francisco J. Nogales, *Operations Research*, DOI 10.1287/opre.1080.0566.

Material for on-line appendix

Appendix A: Proofs of Theorems 1 and 2

A.1. Proof of Theorem 1

To derive the result we introduce a small perturbation in the return CDF. In particular, we replace $F(R)$ by $(1-h)F(R)+h\Delta_{\hat{r}}$ in (23)-(25). For equation (23) we obtain:

$$0=\int \psi(\mathrm{w}^\top R-m)d((1-h)F+h\Delta_{\hat{r}})(R). \tag{37}$$

By the assumptions made on the function ψ, we know that:

$$0=\int \psi(\mathrm{w}^\top R-m)dF(R)+h\int \psi(\mathrm{w}^\top R-m)d(\Delta_{\hat{r}}-F)(R). \tag{38}$$

Differentiating this expression with respect to h yields

$$0=\frac{\partial}{\partial h}\int \psi(\mathrm{w}^\top R-m)dF(R)+\int \psi(\mathrm{w}^\top R-m)d(\Delta_{\hat{r}}-F)(R)+ h\frac{\partial}{\partial h}\int \psi(\mathrm{w}^\top R-m)d(\Delta_{\hat{r}}-F)(R). \tag{39}$$

By the assumptions made on the functions $\rho(r)$ and $\psi(r)$, the order of the differentiation and integration operators in the expression above can be exchanged. Therefore,

$$0=\int \frac{\partial}{\partial h}\psi(\mathrm{w}^\top R-m)dF(R)+\int \psi(\mathrm{w}^\top R-m)d(\Delta_{\hat{r}}-F)(R)+ +h\int \frac{\partial}{\partial h}\psi(\mathrm{w}^\top R-m)d(\Delta_{\hat{r}}-F)(R). \tag{40}$$

To evaluate (40), it is important to note that the estimators w and m are implicitly defined by equations (23)-(25) as a function of the empirical distribution, which is $(1-h)F(R)+h\Delta_{\hat{r}}$ in (40). Thus, when differentiating (40) with respect to h, we need to apply the chain rule to the estimators w and m, which are a function of h. In particular, applying the chain rule and setting $h=0$ gives

$$0=\left(\int \psi'(\mathrm{w}^\top R-m)R^\top dF(R)\right)\mathrm{IF}_{\mathrm{W}}-\left(\int \psi'(\mathrm{w}^\top R-m)dF(R)\right)\mathrm{IF}_m+ +\int \psi(\mathrm{w}^\top R-m)d(\Delta_{\hat{r}})(R)-\int \psi(\mathrm{w}^\top R-m)dF(R). \tag{41}$$

Due to the shifting property of the delta, the third term of (41) is equal to $\psi(\mathrm{w}^\top\hat{r}-m)$. Moreover, by the first first-order optimality condition (23), the fourth term is zero. Therefore:

$$0=E(\psi'(\mathrm{w}^\top R-m)R^\top)\mathrm{IF}_{\mathrm{W}}-E(\psi'(\mathrm{w}^\top R-m))\mathrm{IF}_m+\psi(\mathrm{w}^\top\hat{r}-m). \tag{42}$$

Applying the same argument to the second and third first-order optimality conditions (24) and (25) yields:

$$0=E(\psi'(\mathrm{w}^\top R-m)RR^\top)\mathrm{IF}_{\mathrm{W}}-E(\psi'(\mathrm{w}^\top R-m)R)\mathrm{IF}_m-\mathrm{IF}_\lambda e-\lambda e+\psi(\mathrm{w}^\top\hat{r}-m)\hat{r} \tag{43}$$

and

$$0=e^\top\mathrm{IF}_{\mathrm{W}}. \tag{44}$$

The result in (26) then follows from (42), (43), and (44).

A.2. Proof of Proposition 1

Straightforward manipulation of the linear system in (26) yields the result.

A.3. Proof of Theorem 2

To derive the S-estimator IFs, we first introduce a small perturbation in the return CDF. In particular, we replace $F(R)$ by $(1-h)F + h\Delta_{\hat{r}}$ in (30)-(34). Then, we differentiate (30)-(34) with respect to h.

First equation. From (30) we obtain

$$0 = \frac{\partial}{\partial h}\int \frac{\nu}{s}\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)dF(R) + \int \frac{\nu}{s}\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)d(\Delta_{\hat{r}} - F)(R) + h\frac{\partial}{\partial h}\int \frac{\nu}{s}\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)d(\Delta_{\hat{r}} - F)(R). \quad (45)$$

By the assumptions made on the functions $\rho(r)$ and $\psi(r)$, the order of the differentiation and integration operators in the expression above can be exchanged. Also, note that the estimators m, s, w, and ν are implicitly defined by (30)-(34) as a function of the empirical distribution, which is $(1-h)F + h\Delta_{\hat{r}}$ in (45). Thus, when differentiating (45) with respect to h, we need to apply the chain rule to the estimators m, s, w, and ν, which are a function of h. Then, by applying the chain rule and setting $h = 0$, yields

$$\begin{aligned}0 = \int \Big(\Big(\frac{1}{s}IF_\nu - \frac{\nu}{s^2}IF_s\Big)\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big) + \frac{\nu}{s}\Big(-\frac{1}{s}\psi'\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)IF_m \\ - \frac{1}{s}\psi'\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)IF_s + \frac{1}{s}\psi'\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)R^\top IF_{\mathrm{w}}\Big)\Big)dF(R) \\ - \int \frac{\nu}{s}\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)dF(R) + \int \frac{\nu}{s}\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)d\Delta_{\hat{r}}(R).\end{aligned}$$

Due to the shifting property of the delta, the last term in the expression above is equal to $\frac{\nu}{s}\psi_{\hat{z}}$. Moreover, by the first first-order optimality condition (30), the third term is zero. Therefore:

$$-\frac{\nu}{s^2}E(\psi'_Z)IF_m - \frac{\nu}{s^2}\big(E(\psi_Z) + E(\psi'_Z Z)\big)IF_s + \frac{\nu}{s^2}E(\psi'_Z R^\top)IF_{\mathrm{w}} + \frac{1}{s}E(\psi_Z)IF_\nu = -\frac{\nu}{s}\psi_{\hat{z}}. \quad (46)$$

Second equation. From (31) we obtain

$$0 = \frac{\partial}{\partial h}\int \frac{\nu}{s}\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)dF(R) + \int \frac{\nu}{s}\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)d(\Delta_{\hat{r}} - F)(R).$$

By the assumptions made on the functions $\rho(r)$ and $\psi(r)$, the order of the differentiation and integration operators in the expression above can be exchanged. This, together with application of the chain rule and setting $h = 0$ give

$$\begin{aligned}-\frac{\nu}{s^2}\big(E(\psi_Z) + E(\psi'_Z Z)\big)IF_m - \frac{\nu}{s^2}\big(2E(\psi_Z Z) + E(\psi'_Z Z^2)\big)IF_s \\ + \frac{\nu}{s^2}\big(E(\psi'_Z Z R^\top) + E(\psi_Z R^\top)\big)IF_{\mathrm{w}} + \frac{1}{s}E(\psi_Z Z)IF_\nu = -\frac{\nu}{s}\psi_{\hat{z}}\hat{z} - 1. \quad (47)\end{aligned}$$

Third equation. From (32) we obtain

$$0 = \frac{\partial}{\partial h}\int -\frac{\nu}{s}\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)RdF(R) - \frac{\partial}{\partial h}\lambda e - \int \frac{\nu}{s}\psi\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)Rd(\Delta_{\hat{r}} - F)(R).$$

By the assumptions made on the functions $\rho(r)$ and $\psi(r)$, the order of the differentiation and integration operators in the expression above can be exchanged. This, together with applying the chain rule and setting $h=0$ give

$$\frac{\nu}{s^2}E(\psi'_Z R)IF_m + \frac{\nu}{s^2}\big(E(\psi_Z R) + E(\psi'_Z ZR)\big)IF_s - \frac{\nu}{s^2}E(\psi'_Z RR^\top)IF_{\mathrm{W}} \\ - \frac{1}{s}E(\psi_Z R)IF_\nu + IF_\lambda e = \frac{\nu}{s}\psi_{\hat{z}}\hat{r} + \lambda e. \quad (48)$$

Fourth equation. From (33) we obtain

$$0 = \frac{\partial}{\partial h}\int \rho\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)dF(R) + \int \rho\Big(\frac{\mathrm{w}^\top R - m}{s}\Big)d(\Delta_{\hat{r}} - F)(R).$$

By the assumptions made on the functions $\rho(r)$ and $\psi(r)$, the order of the differentiation and integration operators in the expression above can be exchanged. This, together with applying the chain rule and setting $h=0$ give

$$-\frac{1}{s}E(\psi_Z)IF_m - \frac{1}{s}E(\psi_Z Z)IF_s + \frac{1}{s}E(\psi_Z R^\top)IF_{\mathrm{W}} = -\rho_{\hat{z}} + K. \quad (49)$$

Fifth equation. From (34) we obtain

$$0 = e^\top \mathrm{IF}_{\mathrm{W}}. \quad (50)$$

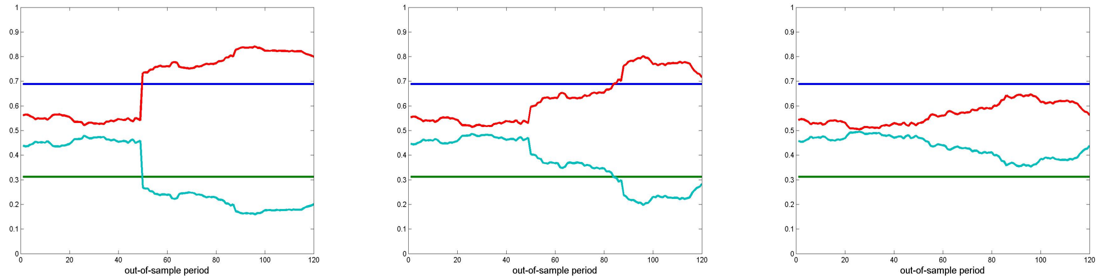
The result in (35) follows from (46)–(50).

A.4. Proof of Proposition 2

First, we show that the vector b in (35) is bounded. To see this, note that Tukey's biweight function $\rho(r)$ is bounded for all r. Moreover, its first derivative $\psi(r)$ is zero for all values of r with a large absolute value. Therefore, $\psi(r)r$ is also bounded for all r. The result follows from the invertibility of the matrix in (35).

Figure 3 Minimum-variance and M- and S-portfolio weights for the example in Section 2.

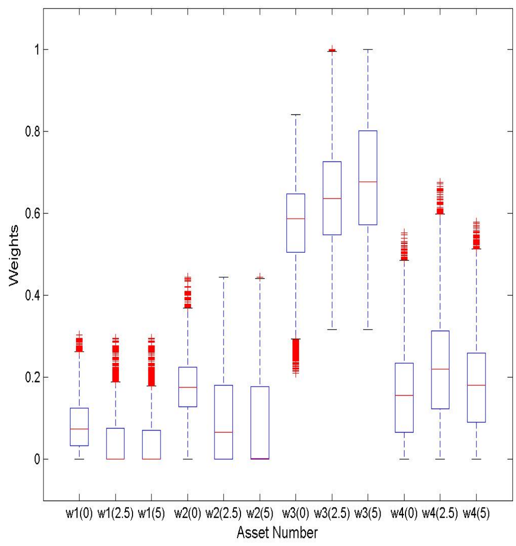
This figure depicts the weights of the 120 estimated minimum-variance, minimum M-risk and minimum S-risk portfolios for the two-asset example in Section 2. The figure also depicts the "true" minimum-variance portfolio weights obtained from the true asset-return distribution used to generate the time series of asset returns for the example. The "true" portfolio weights are constant for all 120 periods and thus they are represented by straight horizontal lines. Panel (a) depicts the minimum-variance portfolio weights, Panel (b) depicts the minimum M-risk portfolio weights, and Panel (c) depicts the minimum S-risk portfolio weights.



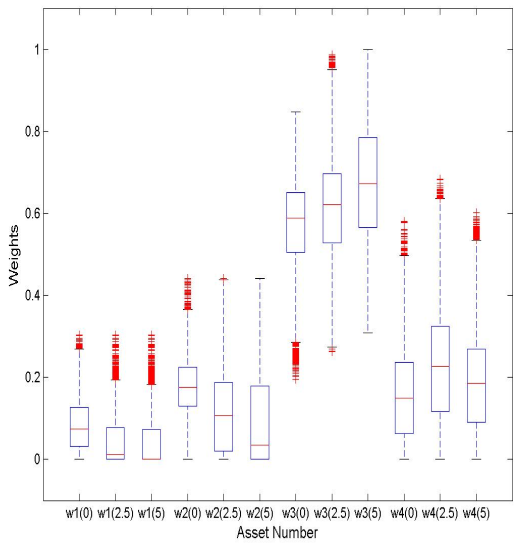
(a)Min-var portfolio weights (b)M-risk portfolio weights (c)S-risk portfolio weights

Figure 4 Boxplots of constrained portfolio weights for simulated data

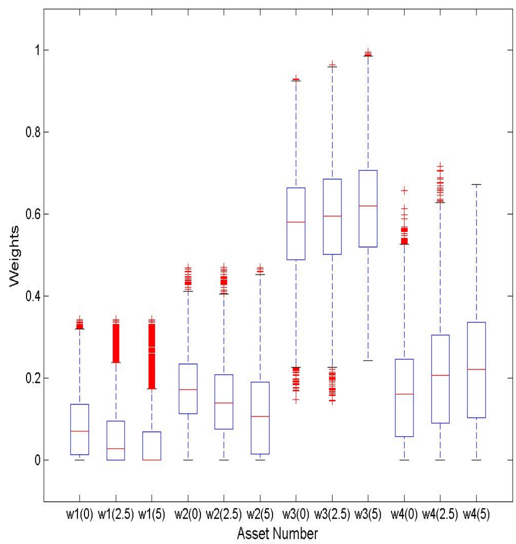
This figure gives the boxplots of the portfolio weights for the shortselling constrained policies and for the simulated datasets with 0, 2.5, and 5% of the data deviating from normality. Each of the four panels gives the boxplots for one of the following four unconstrained policies: minimum-variance (Min-var), 2-step robust minimum-variance (2-Var), minimum M-risk portfolio with Huber's loss function (M-Hub), and minimum S-risk portfolio with Tukey's biweight function (S-Tuk). Each panel contains 12 boxplots corresponding to each of the four assets and each of the tree degrees of deviation from normality. Each boxplot is labelled as $wk(h)$, where $k = 1, 2, 3, 4$ is the asset number and $h = 0, 2.5, 5$ is the proportion of the return data deviating from normality. The boxplots for the portfolio weights for different values of h are given side-by-side. Finally, the box for each portfolio weight has lines at the 25, 50 and 75 percentile values of the portfolio weights. The whiskers are lines extending from each end of the boxes to show the extent of the rest of the data. Extreme portfolio weights that have values beyond the whiskers are also depicted.



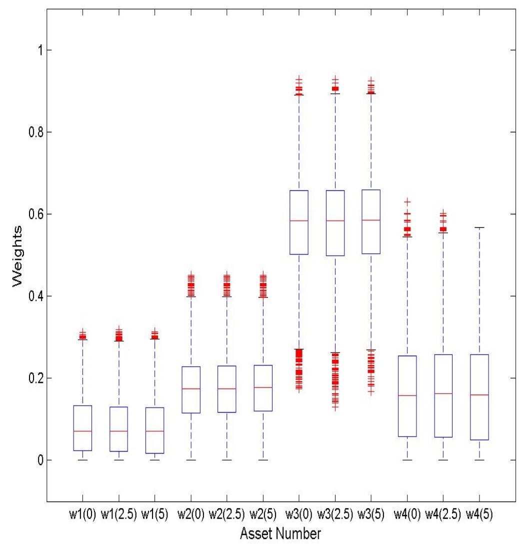
(a)Min-var



(b)2-Var



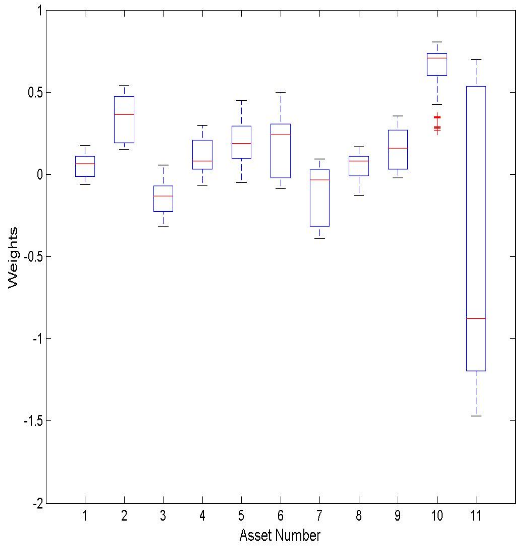
(c)M-Hub



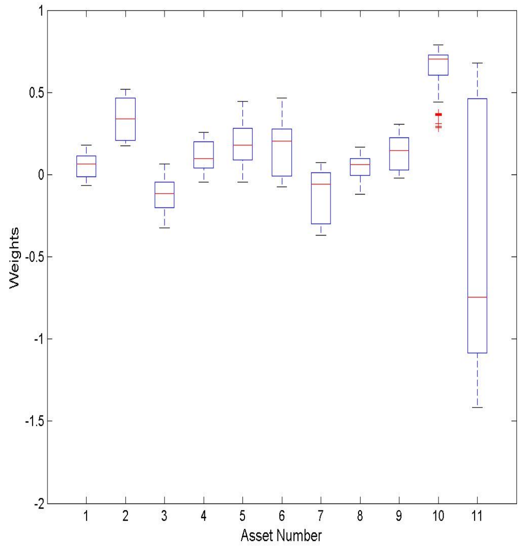
(d)S-Tuk

Figure 5 Boxplots of unconstrained portfolio weights for ten S&P sectors and market

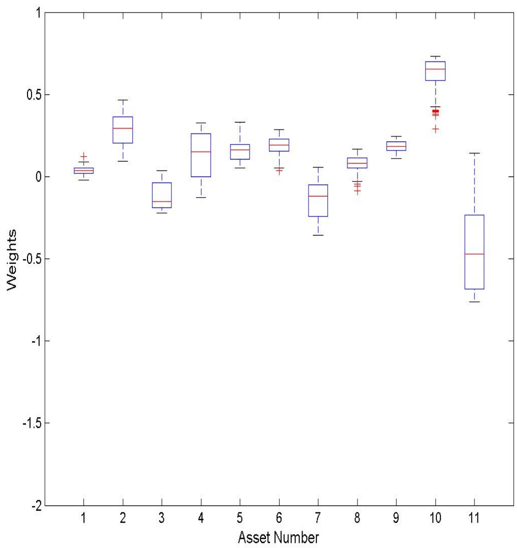
This figure gives the boxplots of the portfolio weights for the unconstrained policies and for the empirical dataset corresponding to the ten S&P sector-tracking portfolios and the market. Each of the four panels gives the boxplots for one of the following four unconstrained policies: minimum-variance (Min-var), two-step robust minimum-variance (2-Var), minimum M-risk portfolio with Huber's loss function (M-Hub) and minimum S-risk portfolio with Tukey's biweight function (S-Tuk). Each of the four panels gives the boxplots of the portfolio weights for each of the 11 assets. The box for each portfolio weight has lines at the 25, 50 and 75 percentile values of the portfolio weights. The whiskers are lines extending from each end of the boxes to show the extent of the rest of the data. Extreme portfolio weights that have values beyond the whiskers are also depicted.



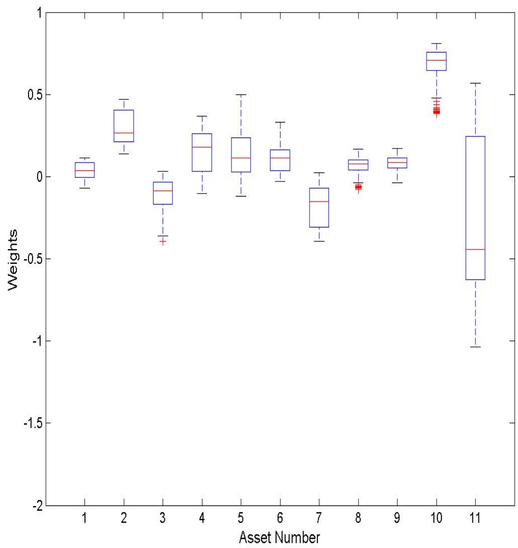
(a)Min-var



(b)2-Var



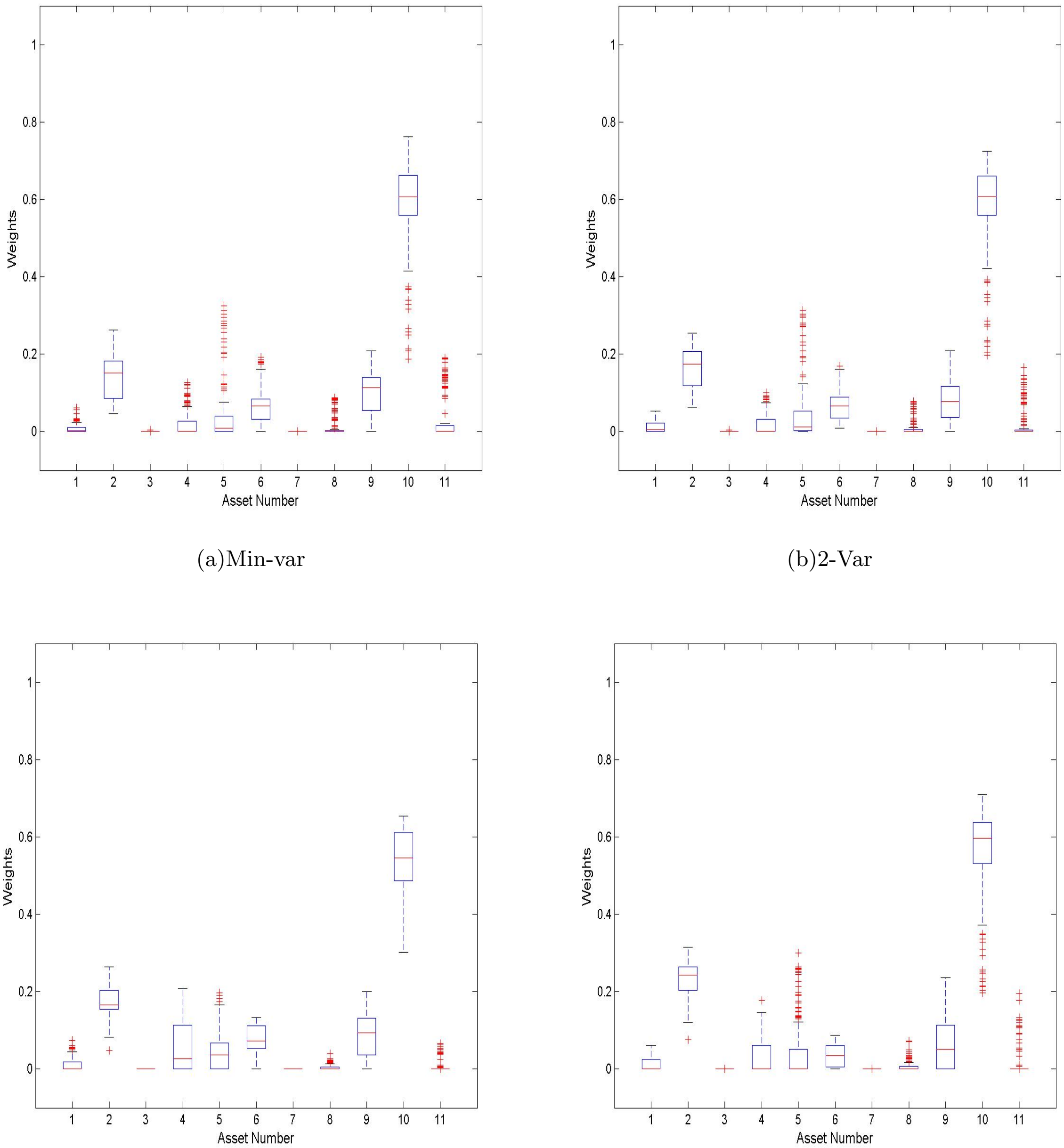
(c)M-Hub



(d)S-Tuk

Figure 6 Boxplots of constrained portfolio weights for ten S&P sectors and market

This figure gives the boxplots of the portfolio weights for the shortselling constrained policies and for the empirical dataset corresponding to the ten S&P sector-tracking portfolios and the market. Each of the four panels gives the boxplots for one of the following four unconstrained policies: minimum-variance (Min-var), two-step robust minimum-variance (2-Var), minimum M-risk portfolio with Huber's loss function (M-Hub) and minimum S-risk portfolio with Tukey's biweight function (S-Tuk). Each of the four panels gives the boxplots of the portfolio weights for each of the 11 assets. The box for each portfolio weight has lines at the 25, 50 and 75 percentile values of the portfolio weights. The whiskers are lines extending from each end of the boxes to show the extent of the rest of the data. Extreme portfolio weights that have values beyond the whiskers are also depicted.



(a)Min-var (b)2-Var

(c)M-Hub (d)S-Tuk

Table 3 Simulated dataset with positive and negative deviations from normality

This table reports the out-of-sample mean return of the different portfolios on dates when the asset returns deviate from normality positively (the returns are well above their average) and on dates when the asset returns deviate from normality negatively (the returns are well below their average). The table also reports the out-of-sample Sharpe ratios of the different portfolios on all dates (including dates when the return follows the normal distribution as well as dates when the returns deviate from normality positively and negatively). The simulated asset-return data contain 5% of returns deviating from normality, which include returns deviating positively and negatively.

	Unconstrained policies						Constrained policies					
Statistic	Mean-var	Min-var	2-Mean	2-Var	M-Hub	S-Tuk	Mean-var	Min-var	2-Mean	2-Var	M-Hub	S-Tuk
Mean (neg)	-0.03125	-0.19306	-0.06542	-0.20600	-0.22046	-0.23376	-0.07669	-0.22132	-0.11525	-0.22499	-0.22831	-0.23393
Mean (pos)	0.40432	0.20702	0.40423	0.21985	0.23488	0.24626	0.29333	0.23114	0.29092	0.23511	0.23768	0.24572
Sharpe rat. (tot)	0.10898	0.17992	0.09884	0.18074	0.18697	0.18495	0.19882	0.18432	0.19033	0.18432	0.18590	0.18517

Table 4 Sharpe ratios of the ten S&P sector portfolios and the market.

The table reports the Sharpe ratio of each of the eleven assets that compose the S&P sectors dataset as well as the p-value of the difference between each Sharpe ratio and the Sharpe ratio of the constrained minimum-variance portfolio.

Assets	Sharpe ratio	pVal
Energy	0.0884	(0.48)
Material	0.0666	(0.39)
Industrials	0.0958	(0.42)
Cons-Discr	0.1691	(0.16)
Cons-Staples	0.2246	(0.03)
Helth-Care	0.2180	(0.04)
Financials	0.2220	(0.01)
Inf-Tech	0.1533	(0.21)
Telecom	0.0483	(0.33)
Utilities	0.0573	(0.31)
S&P500	0.1444	(0.19)