The impact of model risk on dynamic portfolio selection under multi-period mean-standard-deviation criterion

Spiridon Penev[1] Pavel Shevchenko[2] Wei Wu[1]

[1]The University of New South Wales, Sydney
Australia

[2]Macquarie University
Sydney, Australia

Sydney, 8th December 2017

Outline

- We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.
- Uncertainty is measured by the Kullback-Leibler divergence (more generally by α-divergence).
- We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.
- As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.
- Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution$\rightarrow$ quantify model risk.

- We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.
- Uncertainty is measured by the Kullback-Leibler divergence (more generally by α-divergence).
- We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.
- As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.
- Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution$\rightarrow$ quantify model risk.

- We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.
- Uncertainty is measured by the Kullback-Leibler divergence (more generally by α-divergence).
- We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.
- As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.
- Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution$\rightarrow$ quantify model risk.

- We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.
- Uncertainty is measured by the Kullback-Leibler divergence (more generally by α-divergence).
- We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.
- As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.
- Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution$\rightarrow$ quantify model risk.

- We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.
- Uncertainty is measured by the Kullback-Leibler divergence (more generally by α-divergence).
- We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.
- As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.
- Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution$\rightarrow$ quantify model risk.

Which criterion should an investor choose to optimize portfolio wealth?

- Markowitz (1952): $\max_u (E(W) - \kappa * Var(W))$
- Others: Safety-first: $\min_u P(W \leq d)$, targeting particular wealth level: $\min_u E[(W - \hat{W})^2]$, etc.
- One more: Mean-st. deviation (MSD) criterion: $\max_u (E(W) - \kappa \sqrt{Var(W)})$.

Why the latter: for elliptically distributed returns, optimizing a risk measure form the whole class of Translation-invariant and positive-homogeneous risk measures (TIPH) is equivalent to optimizaing the MSD criterion (with a suitable “risk averse enough" $\kappa > 0$).

Which criterion should an investor choose to optimize portfolio wealth?

- Markowitz (1952): $\max_u (E(W) - \kappa * Var(W))$
- Others: Safety-first: $\min_u P(W \leq d)$, targeting particular wealth level: $\min_u E[(W - \hat{W})^2]$, etc.
- One more: Mean-st. deviation (MSD) criterion: $\max_u (E(W) - \kappa \sqrt{Var(W)})$.

Why the latter: for elliptically distributed returns, optimizing a risk measure form the whole class of Translation-invariant and positive-homogeneous risk measures (TIPH) is equivalent to optimizaing the MSD criterion (with a suitable “risk averse enough" $\kappa > 0$).

Which criterion should an investor choose to optimize portfolio wealth?

- Markowitz (1952): $\max_u(E(W) - \kappa * Var(W))$
- Others: Safety-first: $\min_u P(W \leq d)$, targeting particular wealth level: $\min_u E[(W - \hat{W})^2]$, etc.
- One more: Mean-st. deviation (MSD) criterion: $\max_u(E(W) - \kappa\sqrt{Var(W)})$.

Why the latter: for elliptically distributed returns, optimizing a risk measure form the whole class of Translation-invariant and positive-homogeneous risk measures (TIPH) is equivalent to optimizaing the MSD criterion (with a suitable “risk averse enough" $\kappa > 0$).

However: i) the joint distribution of the assets is unkonwn (e.g., ‘slightly deviating’ from a nominal multivariate normal).
ii) changes of the distribution over time (in multiple periods) (need dynamic approach).
The deviation can be measured by: KL divergence, α-divergence..

Scenario: Investor (decision-maker) is concerned about model accuracy and wants to safe-guard against possible worst-case events on the market. Try to quantify the intuition: “big divergence$\rightarrow$ significant impact on an optimal investment decision that is calculated under the nominal distribution". Ultimate goal: If distributional assumptions are violated only ‘slightly’$\rightarrow$ use the optimal investment strategy under the nominal model (since robust approach may deliver too pessimistic strategy). Ideally: ball of radius η_0 around the nominal model: stay with the nominal “inside", switch to robust “outside". Need also to quantify model risk from risk management perspective.

Outline

Using KL divergence: reasonable for short term re-balancing (daily or weekly). Big advantage: closed form worst case distribution is available$\rightarrow$allows us to get the optimal (time-consistent) strategy under the worst case distribution in a semi-analytical form. Our previous work (BGPW, *Automatica* 2016) delivers the optimal strategy under the nominal model$\rightarrow$ can compare performance under worst case scenario.

Problem Formulation $d > 1$ risky assets; fixed investment horizon $[0,N]$; return of each asset over the nth period $[n, n+1]$, $n = 0, ..., N-1$: as $\boldsymbol{r}_{n+1} = (r^1_{n+1}, ..., r^d_{n+1})^\intercal$, withe r^i_{n+1}, $i = 1, ..., d$. Filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_n), \mathbb{P})$ with the sample space Ω, the sigma-algebra $\mathscr{F}$, filtration $(\mathscr{F}_n)$, probab. measure $\mathbb{P}$, sigma-algebra $\mathscr{F}_n = \sigma(\boldsymbol{r}_m, 1 \le m \le n)$. Return vector $\boldsymbol{r}_{n+1}$: $\mathbb{E}(|\boldsymbol{r}_{n+1}|^2) < \infty$.
At time $n = 0, ..., N-1$: re-balance using strategy $\boldsymbol{u} = (\boldsymbol{u}_0, ..., \boldsymbol{u}_{N-1})^\intercal$, all $\boldsymbol{u}_n$, taking values in U:

$$U = \left\{ \boldsymbol{u} \in \mathbb{R}^d : 1^\intercal \boldsymbol{u} = 1, \text{ for } i = 1, ..., d \right\}.$$

For $m > 0$, we use $\mathscr{U}^m$ to denote the set of admissible sub-strategies $\boldsymbol{u}^m = (\boldsymbol{u}_n)_{n \ge m}$

Let W_n denote the wealth at time n ($n = 0, ..., N$). Assume: W_n and $\boldsymbol{r}_{n+1}$ are independent. During $[n, n+1]$, wealth changes:

$$W_{n+1} = W_n(\mathbf{1} + \boldsymbol{r}_{n+1})^\top \boldsymbol{u}_n = W_n \boldsymbol{R}_{n+1}^\top \boldsymbol{u}_n,$$

where $\boldsymbol{R}_{n+1} = \mathbf{1} + \boldsymbol{r}_{n+1}$. At any time m, aim: optimize $J_{m,x}(\boldsymbol{u}^m) = \mathbb{E}\left(\sum_{n=m}^{N-2} \mathscr{J}_{n,W_n}(W_{n+1}) + \mathscr{J}_{N-1,W_{N-1}}(W_N) | W_m = x \right)$, where

$$\mathscr{J}_{n,W_n}(W_{n+1}) = W_{n+1} - \kappa_n \sqrt{Var_{n,W_n}(W_{n+1})} = W_n \left(\boldsymbol{R}_{n+1}^\top \boldsymbol{u}_n - \kappa_n \sqrt{\boldsymbol{u}_n^\top \boldsymbol{\Sigma}_n \boldsymbol{u}_n} \right)$$

$$Var_{n,W_n}(W_{n+1}) = Var(W_{n+1} | W_n), \quad \text{and} \quad \boldsymbol{\Sigma}_n = Var(\boldsymbol{r}_{n+1}),$$

The above: multi-period selection criterion of MSD type; κ_n characterizes investor's risk aversion. Details: BGPW.

UNSW THE UNIVERSITY OF NEW SOUTH WALES

The value function of this control problem:

$$\mathscr{V}(m,x) = \sup_{\boldsymbol{u}^m \in \mathscr{U}^m} J_{m,x}(\boldsymbol{u}^m). \tag{1}$$

Use KL divergence $\mathscr{R}(\mathscr{E}) = \mathbb{E}\left(\mathscr{E} \log \mathscr{E}\right)$, where $\mathscr{E}$ is the ratio of the density of an alternative distribution to model distribution. For a given $\eta > 0$, a KL divergence ball is:

$$\mathscr{B}_\eta = \{\mathscr{E} : \mathscr{R}(\mathscr{E}) \leq \eta\}. \tag{2}$$

Next: robust version of the problem. Let balls

$$\mathscr{B}_{\eta_n} = \{\mathscr{E} : \mathscr{R}(\mathscr{E}) \leq \eta_n\}, \quad \text{where } n = 0, \ldots, N-1.$$

For starting time $m = 0, \ldots, N-1$, denote $\mathscr{E}^m = (\mathscr{E}_m, \ldots, \mathscr{E}_{N-1})$ such that each $\mathscr{E}_n \in \mathscr{B}_{\eta_n}$, where $n = m, \ldots, N-1$, by $\mathscr{B}^m$ ($\mathscr{E}_n$: density ratio of alternative to nominal distribution over $[n, n+1]$). Then: $V(m,x) = \sup_{\boldsymbol{u}^m \in \mathscr{U}^m} \inf_{\mathscr{E}^m \in \mathscr{B}^m} J_{m,x}(\mathscr{E}^m, \boldsymbol{u}^m)$,

$$J_{m,x}(\mathscr{E}^m, \boldsymbol{u}^m) = \mathbb{E}\Bigg(\mathscr{E}_m W_m \Big(\boldsymbol{R}_{m+1}^{\mathsf{T}} \boldsymbol{u}_m - \kappa_m \sqrt{\boldsymbol{u}_m^{\mathsf{T}} \boldsymbol{\Sigma}_m \boldsymbol{u}_m} \Big) + \sum_{n=m+1}^{N-1} e^{-\eta_n c_n \kappa_n} \mathscr{E}_n \mathscr{J}_{n,W_n}(W_{n+1}) \Big| W_m = x \Bigg),$$

Remark. c_n is a weighting to reflect the fact that a large uncertainty in the future should not impact too much the decision at the current stage. In addition, this also guarantees the existence an optimal solution. As $c_n \to 0 \to$ return to the non-robust case.

Semi-Analytical Optimal Solution under KL Divergence. We require strongly time consistent optimal robust strategy. It represents a robustified version of a strong time consistent optimal strategy inspired by Kang and Filar 2006 (see also BGPW).

Definition

A strategy $\boldsymbol{u}^{m,*} = (\boldsymbol{u}^*_m, \dots, \boldsymbol{u}^*_{N-1})$ is strongly time consistent optimal robust w.r.t. $J_{m,x}(\mathscr{E}^m, \boldsymbol{u}^m)$ if:

- **1**: Let $\mathscr{A}^m \subset \mathscr{U}^m$ be a set of strategies $\boldsymbol{u}^m = (\boldsymbol{v}, \boldsymbol{u}^*_{m+1}, \dots, \boldsymbol{u}^*_{N-1})$. Then $\exists \mathscr{E}^{m,*} \in \mathscr{B}^m$ s.t. $\sup_{\boldsymbol{u}^m \in \mathscr{A}^m} \inf_{\mathscr{E}^m \in \mathscr{B}^m} J_{m,x}(\mathscr{E}^m, \boldsymbol{u}^m) = J_{m,x}(\mathscr{E}^{m,*}, \boldsymbol{u}^{m,*})$.
- **2**: For $n = m+1, \dots, N-1, \exists \mathscr{E}^{n,*} \in \mathscr{B}^n$ s.t. $\sup_{\boldsymbol{u}^n \in \mathscr{U}^n} \inf_{\mathscr{E}^n \in \mathscr{B}^n} J_{n,x}(\mathscr{E}_n, \boldsymbol{u}^n) = J_{n,x}(\mathscr{E}^{n,*}, \boldsymbol{u}^{n,*})$.

If only **1** is satisfied: call it weakly time consistent optimal robust strategy w.r.t. $J_{n,x}(\cdot)$.

Semi-Analytical Optimal Solution under KL Divergence. We require strongly time consistent optimal robust strategy. It represents a robustified version of a strong time consistent optimal strategy inspired by Kang and Filar 2006 (see also BGPW).

Definition

A strategy $\boldsymbol{u}^{m,*} = (\boldsymbol{u}^*_m, \ldots, \boldsymbol{u}^*_{N-1})$ is strongly time consistent optimal robust w.r.t. $J_{m,x}(\mathscr{E}^m, \boldsymbol{u}^m)$ if:

- **1**: Let $\mathscr{A}^m \subset \mathscr{U}^m$ be a set of strategies $\boldsymbol{u}^m = (\boldsymbol{v}, \boldsymbol{u}^*_{m+1}, \ldots, \boldsymbol{u}^*_{N-1})$. Then $\exists \mathscr{E}^{m,*} \in \mathscr{B}^m$ s.t. $\sup_{\boldsymbol{u}^m \in \mathscr{A}^m} \inf_{\mathscr{E}^m \in \mathscr{B}^m} J_{m,x}(\mathscr{E}^m, \boldsymbol{u}^m) = J_{m,x}(\mathscr{E}^{m,*}, \boldsymbol{u}^{m,*})$.
- **2**: For $n = m+1, \ldots, N-1, \exists \mathscr{E}^{n,*} \in \mathscr{B}^n$ s.t. $\sup_{\boldsymbol{u}^n \in \mathscr{U}^n} \inf_{\mathscr{E}^n \in \mathscr{B}^n} J_{n,x}(\mathscr{E}_n, \boldsymbol{u}^n) = J_{n,x}(\mathscr{E}^{n,*}, \boldsymbol{u}^{n,*})$.

If only **1** is satisfied: call it weakly time consistent optimal robust strategy w.r.t. $J_{n,x}(\cdot)$.

Semi-Analytical Optimal Solution under KL Divergence. We require strongly time consistent optimal robust strategy. It represents a robustified version of a strong time consistent optimal strategy inspired by Kang and Filar 2006 (see also BGPW).

Definition

A strategy $\boldsymbol{u}^{m,*} = (\boldsymbol{u}^*_m, ..., \boldsymbol{u}^*_{N-1})$ is strongly time consistent optimal robust w.r.t. $J_{m,x}(\mathscr{E}^m, \boldsymbol{u}^m)$ if:

- **1**: Let $\mathscr{A}^m \subset \mathscr{U}^m$ be a set of strategies $\boldsymbol{u}^m = (\boldsymbol{v}, \boldsymbol{u}^*_{m+1}, ..., \boldsymbol{u}^*_{N-1})$. Then $\exists \mathscr{E}^{m,*} \in \mathscr{B}^m$ s.t. $\sup_{\boldsymbol{u}^m \in \mathscr{A}^m} \inf_{\mathscr{E}^m \in \mathscr{B}^m} J_{m,x}(\mathscr{E}^m, \boldsymbol{u}^m) = J_{m,x}(\mathscr{E}^{m,*}, \boldsymbol{u}^{m,*})$.
- **2**: For $n = m+1, ..., N-1, \exists \mathscr{E}^{n,*} \in \mathscr{B}^n$ s.t. $\sup_{\boldsymbol{u}^n \in \mathscr{U}^n} \inf_{\mathscr{E}^n \in \mathscr{B}^n} J_{n,x}(\mathscr{E}_n, \boldsymbol{u}^n) = J_{n,x}(\mathscr{E}^{n,*}, \boldsymbol{u}^{n,*})$.

If only **1** is satisfied: call it weakly time consistent optimal robust strategy w.r.t. $J_{n,x}(\cdot)$.

Loosely speaking: (1): Nash equilibrium. In game-theoretic terms: choose an arbitrary point in time m of the game (many players). Suppose that every player at moment in time $> m$ uses certain strategy. Then the optimal choice for the player at moment m is based on choices which use the knowledge that future players are using this particular strategy in the future (pseudo DPP). (2) the extra property that any sub-strategy of of a weakly time consistent robust strategy is also optimal for the corresponding subsequent time period.

Since the value function of the robust control problem is separable (in the sense that it can be written as a sum of expectations), we know: weakly time consistent optimal strategy, which can be found by period-wise optimization, is also a strongly time consistent optimal strategy.

Theorem. Suppose $(\boldsymbol{u}_m^*)$, $m = 0, 1, ..., N-1$ is a strategy where there exists a sequence (θ_m^*) s.t. $\mathbb{E}\Big(\exp\big(-\boldsymbol{R}_{m+1}^{\mathsf{T}}\boldsymbol{u}_m^*\frac{2}{\theta_m^*}\big)\Big) < \infty$, and

$$\boldsymbol{u}_m^* = \frac{S_m^*}{\kappa_m}\Big(\boldsymbol{\Sigma}_m^{-1}\boldsymbol{X}_m^* - \frac{b_m^*\boldsymbol{\Sigma}_m^{-1}\mathbb{1}}{a_m}\Big) + \frac{\boldsymbol{\Sigma}_m^{-1}\mathbb{1}}{a_m}, \quad (3)$$

$$S_m^* = \sqrt{\frac{\frac{1}{a_m}}{1 - \frac{h_m^*}{\kappa_m^2} + \frac{(b_m^*)^2}{\kappa_m^2 a_m}}} = \sqrt{\frac{\frac{1}{a_m}}{1 - \frac{1}{\kappa_m^2}g_m^*}}, \quad (4)$$

$$\boldsymbol{X}_m^* = \frac{\mathbb{E}\Big(\exp\big(-\boldsymbol{R}_{m+1}^{\mathsf{T}}\boldsymbol{u}_m^*\frac{1}{\theta_m^*}\big)\boldsymbol{R}_{m+1}\Big)}{\mathbb{E}\Big(\exp\big(-\boldsymbol{R}_{m+1}^{\mathsf{T}}\boldsymbol{u}_m^*\frac{1}{\theta_m^*}\big)\Big)}$$

$$+e^{-\eta_{m+1}c_{m+1}\kappa_{m+1}}G_{m+1}(\boldsymbol{u}_{m+1}^*, \theta_{m+1}^*)\mathbb{E}(\boldsymbol{R}_{m+1}),$$

$$\mathbb{E}\big(\mathscr{E}_m^*\log(\mathscr{E}_m^*)\big) = \eta_m,$$

where

$$g_m^* = h_m^* - \frac{(b_m^*)^2}{a_m}, \quad h_m^* = (\boldsymbol{X}_m^*)^T \boldsymbol{\Sigma}_m^{-1} \boldsymbol{X}_m^*, \quad a_m = 1^\mathsf{T} \boldsymbol{\Sigma}_m^{-1} 1,$$

$$b_m^* = 1^\mathsf{T} \boldsymbol{\Sigma}_m^{-1} \boldsymbol{X}_m^*, \quad \mathscr{E}_m^* = \frac{\exp\left(-\boldsymbol{R}_{m+1}^\mathsf{T} \boldsymbol{u}_m^* \frac{1}{\theta_m^*}\right)}{\mathbb{E}\Big(\exp\left(-\boldsymbol{R}_{m+1}^\mathsf{T} \boldsymbol{u}_m^* \frac{1}{\theta_m^*}\right)\Big)} \quad \mathbb{P}\text{-a.s.},$$

$G_m(\boldsymbol{u}_m^*, \theta_m^*) = -\theta_m^* \log \mathbb{E}\Big(\exp\left(-\boldsymbol{R}_{m+1}^\mathsf{T} \boldsymbol{u}_m^* \frac{1}{\theta_m^*}\right)\Big) +$
$e^{-\eta_{m+1} c_{m+1} \kappa_{m+1}} G_{m+1}(\boldsymbol{u}_{m+1}^*, \theta_{m+1}^*) \mathbb{E}\big(\boldsymbol{R}_{m+1}^\mathsf{T} \boldsymbol{u}_m^*\big) - \kappa_m S_m^* - \eta_m \theta_m^*,$
$G_N(\boldsymbol{u}_N^*, \theta_N^*) = 0.$
Then, $(\boldsymbol{u}_m^*)$ is optimal, and the value function is given by

$$V(m,x) = x G_m(\boldsymbol{u}_m^*, \theta_m^*),$$

where $x \in (0, \infty)$.

$d=3$ stocks: Navitas, Domino and Tabcorp, $N=5$. Historical daily prices collected: 1 Jan 2015 - 31 Dec 2015. The 261 daily returns calculated. Set $\kappa_n = 3$, $W_0 = 1$, returns are assumed to be i.i.d. over the investment horizon.

Comparison of Optimal Robust and Non-Robust Portfolio. For a nominal 3-dim. MVN, mean $\boldsymbol{\mu}$, cov. mat. $\boldsymbol{\Sigma}$, and alternative model: 3-dimensional MVN, mean $\bar{\boldsymbol{\mu}}$, cov. mat. $\bar{\boldsymbol{\Sigma}}$, KL:

$$\mathscr{R}(\mathscr{E}) = \frac{1}{2}\left(trace(\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{\Sigma}}) + (\boldsymbol{\mu}-\bar{\boldsymbol{\mu}})^{\intercal}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}-\bar{\boldsymbol{\mu}}) - d + \log\left(\frac{det(\boldsymbol{\Sigma})}{det(\bar{\boldsymbol{\Sigma}})}\right)\right).$$

For illustration$\rightarrow$ consider $\bar{\boldsymbol{\mu}} = \gamma \times \boldsymbol{\mu}$, for some $\gamma \in \mathbb{R}$, and $\bar{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}$.

In the worst case scenario for model disturbance, the alternative model is on the boundary of the KL div. ball (see Theorem)$\rightarrow$ the divergence between the two models is equal to η_n. Assume $\eta_n = \eta$ to be constant over the investment horizon. For simplicity, simply write η and c. Choose c_2, c_3, c_4, c_5 such that $(c_2\eta\kappa, c_3\eta\kappa, c_4\eta\kappa, c_5\eta\kappa) = (7.5, 8.0, 8.5, 9.0)$ (reflecting investor's risk tolerance for uncertainty of distribution) (in contrast to κ which is the risk aversion of the investor's preference for a fixed distribution). So: the investor will have its own freedom to choose the amount of penalization.

Suppose: we calculate under the worst case distribution. By generating data from it$\rightarrow$ compare the performance under the optimal robust and non-robust strategies for different η. The optimal robust strategy: calculated by using $500,000$ Monte Carlo simulations. Then, we simulate 500,000 daily return paths (over 5 days), and compare the performance over three reasonable metrics:

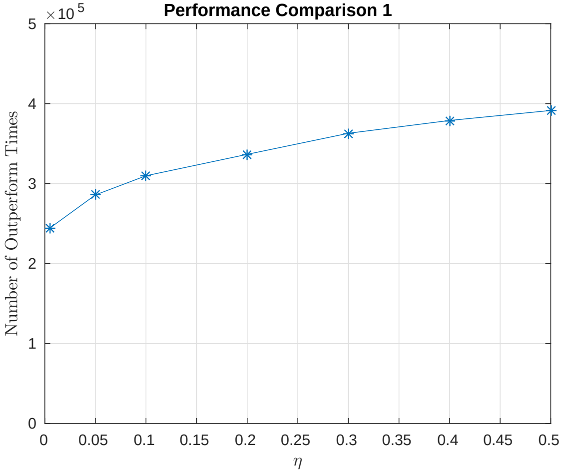

Figure: the number of times robust outperforms non-robust

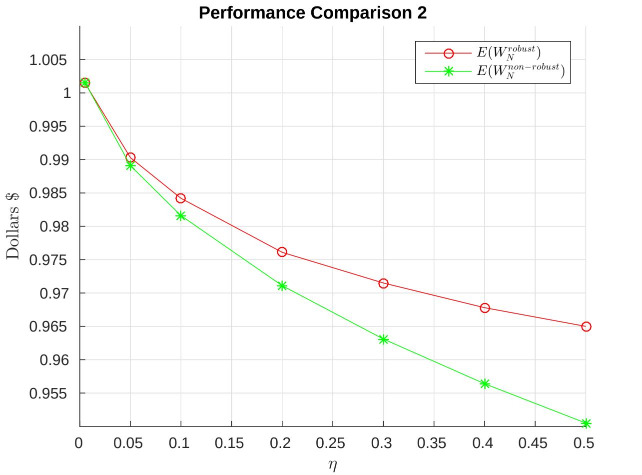

Figure: robust vs non-robust: expected terminal wealth

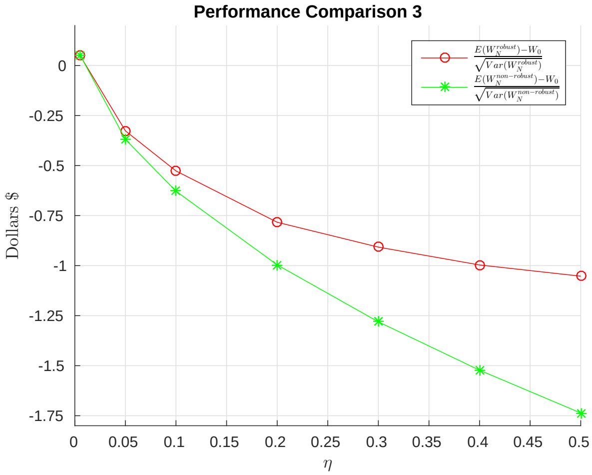

Figure: robust vs non-robust: ratio of the difference between the expected terminal wealth and the initial wealth to the standard deviation of the terminal wealth

Table: Performance of Robust and Non-Robust Optimal Solution (Shift of Mean)

γ	η	Times robust outperforms model	%	$\mathbb{E}(W_N)$ robust	$\mathbb{E}(W_N)$ non-robust	$\mathbb{E}(W_N)$ difference
0.2139	0.0050	244429	48.89%	1.0015	1.0016	-0.0001
-1.4859	0.0500	285828	57.17%	0.9903	0.9891	0.0012
-2.5156	0.1000	309814	61.96%	0.9842	0.9816	0.0026
-3.9718	0.2000	336583	67.32%	0.9761	0.9711	0.0050
-5.0892	0.3000	362909	72.58%	0.9715	0.9631	0.0084
-6.0312	0.4000	378952	75.79%	0.9678	0.9564	0.0114
-6.8611	0.5000	391459	78.29%	0.9650	0.9505	0.0145

Another case with a closed form formula for the KL div.: when both the nominal and alternative models are multivariate skew-normal. Given a nominal model $\boldsymbol{Y} \sim SN_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\xi})$ and an alternative model $\bar{\boldsymbol{Y}} \sim SN_d(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}, \bar{\boldsymbol{\xi}})$, then models is given by:

$$\mathscr{R}_{skew}(\mathscr{E}) =$$

$$\mathscr{R}(\mathscr{E}) + 2\sqrt{\frac{2}{\pi}}(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^{\intercal}\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{\Sigma}}^{\frac{1}{2}}\bar{\boldsymbol{\xi}} - \mathbb{E}\left(\log\left(2\Phi\left(\Xi_2 | 1 - \boldsymbol{\xi}^T\boldsymbol{\xi}\right)\right)\right)$$

$$+\mathbb{E}\left(\log\left(2\Phi\left(\Xi_1 | 1 - \bar{\boldsymbol{\xi}}^T\bar{\boldsymbol{\xi}}\right)\right)\right),$$

$$\Xi_1 \sim SN_1\left(0, \bar{\boldsymbol{\xi}}^{\intercal}\bar{\boldsymbol{\xi}}, \sqrt{\bar{\boldsymbol{\xi}}^{\intercal}\bar{\boldsymbol{\xi}}}\right),$$

$$\Xi_2 \sim SN_1\left(\boldsymbol{\xi}^T\boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}), \boldsymbol{\xi}^T\boldsymbol{\Sigma}^{-\frac{1}{2}}\bar{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\xi}, \frac{\boldsymbol{\xi}^T\boldsymbol{\Sigma}^{-\frac{1}{2}}\bar{\boldsymbol{\Sigma}}^{\frac{1}{2}}\bar{\boldsymbol{\xi}}}{\sqrt{\boldsymbol{\xi}^T\boldsymbol{\Sigma}^{-\frac{1}{2}}\bar{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\xi}}}\right).$$

For illustration: take a *d*-dimensional MVN with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ as the nominal. The worst case: *d*-dimensional multivariate skew-normal with location $\bar{\boldsymbol{\mu}}$, scale $\bar{\boldsymbol{\Sigma}}$ and a skewness $\bar{\boldsymbol{\xi}}$, such that $\bar{\boldsymbol{\mu}} = \boldsymbol{\mu}$, and $\bar{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}$. We choose $\bar{\boldsymbol{\xi}}$ such that the mean of the worst case distribution is changed to $\beta\%$ of the mean of the nominal. (Note that the location parameter $\boldsymbol{\mu}$, and the scale parameter $\boldsymbol{\Sigma}$ are *not* the mean and the covariance matrix of the multivariate skew-normal). The reason of choice: to show that the robust case is a strategy to safeguard against (the losses) due to a shift of the mean.

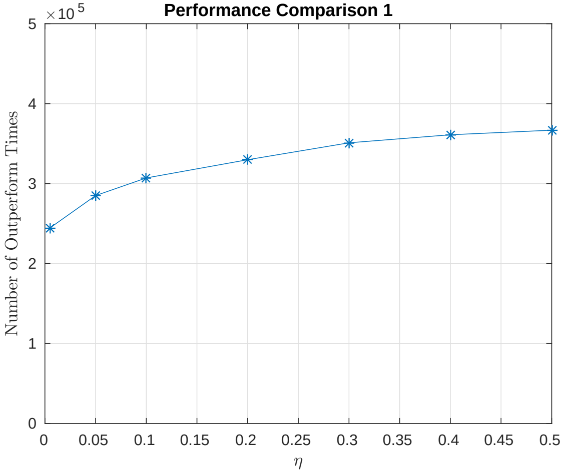

Figure: the number of times robust outperform non-robust

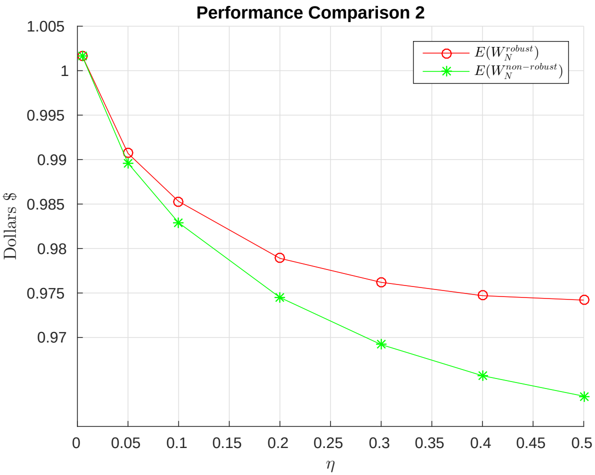

Figure: robust vs non-robust: expected terminal wealth

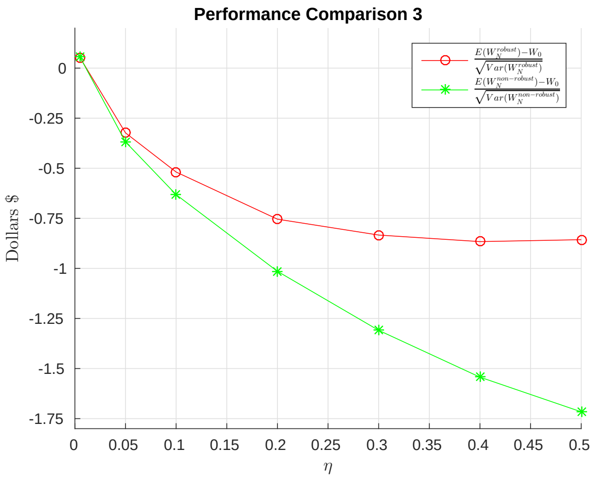

Figure: robust vs non-robust: ratio of the difference between the expected terminal wealth and the initial wealth to the standard deviation of the terminal wealth

Outline

Although in some cases it may be worth choosing the non-robust optimal strategy, we still need to quantify the amount of model risk involved in doing so. Let $\mathbb{Q}$ denote the probability measure of an alternative model (i.e. the empirical measure for the true distribution). The optimal portfolio is said to have a model risk of θ with a confidence level q if

$$\mathbb{Q}\left(W_N^{non-robust} - W_N^{robust} \leq -\theta\right) = 1 - q.$$

(In other words, model risk is the $(1-q)$th-quantile of the distribution of the difference between the terminal wealth under the non-robust strategy and the robust strategy.)

Divide the dataset: the first (dataset 1), is used to estimate the expected value and the covariance matrix of the nominal, which yields:

$$\tilde{\boldsymbol{\mu}} = \begin{pmatrix} 0.0009 \\ 0.0018 \\ 0.0014 \end{pmatrix}, \quad \tilde{\boldsymbol{\Sigma}} = \begin{pmatrix} 0.0004 & 0.0001 & 0.0001 \\ 0.0001 & 0.0004 & 0.0001 \\ 0.0001 & 0.0001 & 0.0003 \end{pmatrix} \quad (5)$$

For illustration$\rightarrow$ assume that the nominal is a 3-dimensional MVN with above mean vector and cov. matrix. Dataset 2 (60 data points) is used to evaluate a forecasted distribution of the incoming daily returns (our estimated alternative model to be used in the following 5 days). Use estimation procedure based on the *k*th-nearest-neighbor approach. Each time, a sample of 60 is generated from the MVN and the divergence between MNV and alternative model is estimated by using this sample, and dataset 2. Repeat 100,000 times.

Estimated divergence using kth-nearest-neighbor (WKV) estimator: $\hat{\mathscr{R}}(\mathscr{E}) = \frac{d}{S}\sum_{i=1}^{S}\log\left(\frac{Sy_k(i)}{(S-1)\tilde{y}_k(i)}\right)$,
$\tilde{y}_k(i)$: Eucl. distance of the kth-nearest-neighbor of $\tilde{\boldsymbol{Y}}_i$ in $(\tilde{\boldsymbol{Y}}_j)_{j\neq i}$, $y_k(i)$: Eucl. distance of the kth-nearest-neighbor of $\tilde{\boldsymbol{Y}}_i$ in $(\boldsymbol{Y}_i)$, $d=3$.

$$\hat{\mathscr{R}}(\mathscr{E}) \approx 0.4001.$$

By knowing the KL divergence, we use a bootstrapping to sample $100,000$ data points from dataset 2 to construct the distribution of $W_N^{non-robust} - W_N^{robust}$. The estimated model risk at $q = 95\%$ confidence level is 0.007, i.e., if the optimal non-robust strategy is applied but the optimal robust strategy turns to be more appropriate, then 95% of the time we would lose no more than 0.7 cents for every one dollar.

UNSW THE UNIVERSITY OF NEW SOUTH WALES

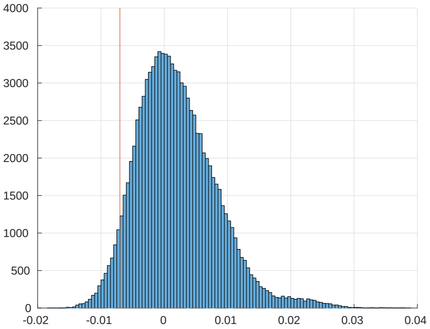

Figure: The distribution of $(W_N^{non-robust} - W_N^{robust})$

- Derived: semi-analytical form of optimal robust strategy for an investment portfolio with uncertainty of distribution of the returns.
- Applied our approach to several numerical examples to illustrate te effect.
- Defined model risk from risk management perspective and presented an algorithm for quantifying the model risk by using empirical data.

Thank you!

Acknowledge: ARC DP grant: DP160103489.

- Derived: semi-analytical form of optimal robust strategy for an investment portfolio with uncertainty of distribution of the returns.
- Applied our approach to several numerical examples to illustrate te effect.
- Defined model risk from risk management perspective and presented an algorithm for quantifying the model risk by using empirical data.

Thank you!

Acknowledge: ARC DP grant: DP160103489.

- Derived: semi-analytical form of optimal robust strategy for an investment portfolio with uncertainty of distribution of the returns.
- Applied our approach to several numerical examples to illustrate te effect.
- Defined model risk from risk management perspective and presented an algorithm for quantifying the model risk by using empirical data.

Thank you!

Acknowledge: ARC DP grant: DP160103489.

- H. Bannister, B. Goldys, S. Penev and W. Wu (2016), "Multiperiod mean-standard-deviation time consistent portfolio selection",*Automatica*,73, 15–26.
- P. Glasserman and X. B. Xu (2013), "Robust portfolio control with stochastic factor dynamics," *Oper. Res.*, 61, 874–893.
- P. Glasserman and X. B. Xu (2014),"Robust Risk Measurement and Model Risk," *Quant. Finance*, 14,29–58.
- J. C. Schneider and N. Schweizer (2015), "Robust measurement of (heavy-tailed) risks: Theory and implimentation,", *J. Econom. Dynam. Control*, 61, 183– 203.
- T. Breuer and I. Csiszár (2016),"Measuring distribution model risk," *Math. Finance*, 26, 395–411.

- H. Bannister, B. Goldys, S. Penev and W. Wu (2016), "Multiperiod mean-standard-deviation time consistent portfolio selection",*Automatica*,73, 15–26.
- P. Glasserman and X. B. Xu (2013), "Robust portfolio control with stochastic factor dynamics," *Oper. Res.*, 61, 874–893.
- P. Glasserman and X. B. Xu (2014),"Robust Risk Measurement and Model Risk," *Quant. Finance*, 14,29–58.
- J. C. Schneider and N. Schweizer (2015), "Robust measurement of (heavy-tailed) risks: Theory and implimentation,", *J. Econom. Dynam. Control*, 61, 183– 203.
- T. Breuer and I. Csiszár (2016),"Measuring distribution model risk," *Math. Finance*, 26, 395–411.

- H. Bannister, B. Goldys, S. Penev and W. Wu (2016), "Multiperiod mean-standard-deviation time consistent portfolio selection",*Automatica*,73, 15–26.
- P. Glasserman and X. B. Xu (2013), "Robust portfolio control with stochastic factor dynamics," *Oper. Res.*, 61, 874–893.
- P. Glasserman and X. B. Xu (2014),"Robust Risk Measurement and Model Risk," *Quant. Finance*, 14,29–58.
- J. C. Schneider and N. Schweizer (2015), "Robust measurement of (heavy-tailed) risks: Theory and implimentation,", *J. Econom. Dynam. Control*, 61, 183– 203.
- T. Breuer and I. Csiszár (2016),"Measuring distribution model risk," *Math. Finance*, 26, 395–411.

- H. Bannister, B. Goldys, S. Penev and W. Wu (2016), "Multiperiod mean-standard-deviation time consistent portfolio selection",*Automatica*,73, 15–26.
- P. Glasserman and X. B. Xu (2013), "Robust portfolio control with stochastic factor dynamics," *Oper. Res.*, 61, 874–893.
- P. Glasserman and X. B. Xu (2014),"Robust Risk Measurement and Model Risk," *Quant. Finance*, 14,29–58.
- J. C. Schneider and N. Schweizer (2015), "Robust measurement of (heavy-tailed) risks: Theory and implimentation,", *J. Econom. Dynam. Control*, 61, 183– 203.
- T. Breuer and I. Csiszár (2016),"Measuring distribution model risk," *Math. Finance*, 26, 395–411.

- H. Bannister, B. Goldys, S. Penev and W. Wu (2016), "Multiperiod mean-standard-deviation time consistent portfolio selection",*Automatica*,73, 15–26.
- P. Glasserman and X. B. Xu (2013), "Robust portfolio control with stochastic factor dynamics," *Oper. Res.*, 61, 874–893.
- P. Glasserman and X. B. Xu (2014),"Robust Risk Measurement and Model Risk," *Quant. Finance*, 14,29–58.
- J. C. Schneider and N. Schweizer (2015), "Robust measurement of (heavy-tailed) risks: Theory and implimentation,", *J. Econom. Dynam. Control*, 61, 183– 203.
- T. Breuer and I. Csiszár (2016),"Measuring distribution model risk," *Math. Finance*, 26, 395–411.