On the optimal combination of naive and mean-variance portfolio strategies

Nathan Lassance
UCLouvain (BE), LFIN

Rodolphe Vanderveken
UCLouvain (BE), LFIN

Frédéric Vrins
UCLouvain (BE), LFIN

IFABS2022
September 2022

Outline

Motivation

Contributions

Optimal vs. constrained combination

Mixed strategy & empirical results

Takeaways

Motivation

Efficient mean-variance portfolios

- Markowitz (1952) defines efficient mean-variance portfolios:

$$\boldsymbol{w}^\star = \operatorname*{argmax}_{\boldsymbol{w}} \underbrace{\boldsymbol{w}^\top \boldsymbol{\mu}}_{\text{portfolio mean}} - \frac{\gamma}{2} \overbrace{\boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}}^{\text{portfolio risk}} = \frac{1}{\gamma} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$$

  - $\boldsymbol{w}$: vector of weights on risky assets ($1 - \mathbf{1}'\boldsymbol{w}$ is invested in the risk-free asset)
  - $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$: vector of mean returns and covariance matrix of returns
  - γ: risk-aversion coefficient

Efficient mean-variance portfolios

- Markowitz (1952) defines efficient mean-variance portfolios:

$$\boldsymbol{w}^\star = \underset{\boldsymbol{w}}{\text{argmax}} \; \underbrace{\boldsymbol{w}^\top \boldsymbol{\mu}}_{\text{portfolio mean}} - \frac{\gamma}{2} \overbrace{\boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}}^{\text{portfolio risk}} = \frac{1}{\gamma} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$$

  - $\boldsymbol{w}$: vector of weights on risky assets ($1 - \mathbf{1}'\boldsymbol{w}$ is invested in the risk-free asset)
  - $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$: vector of mean returns and covariance matrix of returns
  - γ: risk-aversion coefficient

- Challenge: Investors must estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ and rely on

$$\widehat{\boldsymbol{w}}^\star = \frac{1}{\gamma} \widehat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\mu}}$$

Estimation risk and poor performance

- Problem: $\widehat{\boldsymbol{w}}^\star$ performs poorly out-of-sample. The naive equally weighted portfolio ($\boldsymbol{w}_{ew} = \mathbf{1}/N$) often outperforms it.
- See DeMiguel et al. (2009, RFS).
- This fact questions the added value of portfolio optimization.

How to deal with estimation risk?

- Two extreme alternatives: plug-in or naive

How to deal with estimation risk?

- Two extreme alternatives: plug-in or naive
- Plug-in: relying on $\widehat{\boldsymbol{w}}^{\star}$. It is theoretically optimal, but subject to heavy estimation risk
    - Approaches exist to limit estimation errors on parameters (e.g. Ledoit & Wolf shrinkage)

How to deal with estimation risk?

- Two extreme alternatives: plug-in or naive
- Plug-in: relying on $\widehat{\boldsymbol{w}}^\star$. It is theoretically optimal, but subject to heavy estimation risk
  - Approaches exist to limit estimation errors on parameters (e.g. Ledoit & Wolf shrinkage)
- Naive: relying on $\boldsymbol{w}_{ew}$. It is insensitive to estimation risk but it is also suboptimal!

How to deal with estimation risk?

- Two extreme alternatives: plug-in or naive
- Plug-in: relying on $\widehat{\boldsymbol{w}}^{\star}$. It is theoretically optimal, but subject to heavy estimation risk
  - Approaches exist to limit estimation errors on parameters (e.g. Ledoit & Wolf shrinkage)
- Naive: relying on $\boldsymbol{w}_{ew}$. It is insensitive to estimation risk but it is also suboptimal!
- Let's try to find the middle ground ...

Combination of naive and mean-variance portfolio strategies

- Tu and Zhou (2011, JFE) introduce a portfolio combination:

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^{\star} + \kappa_2 \boldsymbol{w}_{ew} \quad \text{with} \quad \boldsymbol{c} := \kappa_1 + \kappa_2 = 1.$$

- $\boldsymbol{\kappa} = (\kappa_1, \kappa_2)$ are the combination coefficients.

Combination of naive and mean-variance portfolio strategies

- Tu and Zhou (2011, JFE) introduce a portfolio combination:

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^\star + \kappa_2 \boldsymbol{w}_{ew} \quad \text{with} \quad \boldsymbol{c} := \kappa_1 + \kappa_2 = 1.$$

- $\boldsymbol{\kappa} = (\kappa_1, \kappa_2)$ are the combination coefficients.

- As in Kan And Zhou (2007, JFQA), they are found by maximizing the expected out-of-sample utility (***EU***)

$$\boldsymbol{\kappa}^{tz} = \underset{\boldsymbol{\kappa}}{\text{argmax}} \ \mathbb{E} \underbrace{\left[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})' \boldsymbol{\mu} - \frac{\gamma}{2} \hat{\boldsymbol{w}}(\boldsymbol{\kappa})' \boldsymbol{\Sigma} \hat{\boldsymbol{w}}(\boldsymbol{\kappa}) \right]}_{\text{out-of-sample utility}} \quad \text{s.t.} \quad \boldsymbol{c}$$

$$= \underset{\boldsymbol{\kappa}}{\text{argmax}} \ \boldsymbol{EU} \quad \text{s.t.} \quad \boldsymbol{c}$$

Combination of naive and mean-variance portfolio strategies

- Tu and Zhou (2011, JFE) introduce a portfolio combination:

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^{\star} + \kappa_2 \boldsymbol{w}_{ew} \quad \text{with} \quad \boldsymbol{c} := \kappa_1 + \kappa_2 = 1.$$

- $\boldsymbol{\kappa} = (\kappa_1, \kappa_2)$ are the combination coefficients.

- As in Kan And Zhou (2007, JFQA), they are found by maximizing the expected out-of-sample utility ($\boldsymbol{EU}$)

$$\boldsymbol{\kappa}^{tz} = \underset{\boldsymbol{\kappa}}{\operatorname{argmax}} \; \underbrace{\mathbb{E}\left[\hat{\boldsymbol{w}}(\boldsymbol{\kappa})'\boldsymbol{\mu} - \frac{\gamma}{2}\hat{\boldsymbol{w}}(\boldsymbol{\kappa})'\boldsymbol{\Sigma}\hat{\boldsymbol{w}}(\boldsymbol{\kappa})\right]}_{\text{out-of-sample utility}} \quad \text{s.t.} \quad \boldsymbol{c}$$

$$= \underset{\boldsymbol{\kappa}}{\operatorname{argmax}} \; \boldsymbol{EU} \quad \text{s.t.} \quad \boldsymbol{c}$$

- This combined rule significantly outperforms $\widehat{\boldsymbol{w}}^{\star}$ and may outperform $\boldsymbol{w}_{ew}$.

This seems all well and good, but ...

- Tu and Zhou (2011) force combination coefficients to sum to one:

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^\star + \kappa_2 \boldsymbol{w}_{ew} \quad \text{with} \quad \boldsymbol{c}$$

- The constraint is needed in combinations of fully invested portfolios, because it is equivalent to the full investment constraint

This seems all well and good, but ...

- Tu and Zhou (2011) force combination coefficients to sum to one:

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^\star + \kappa_2 \boldsymbol{w}_{ew} \quad \text{with} \quad \boldsymbol{c}$$

- The constraint is needed in combinations of fully invested portfolios, because it is equivalent to the full investment constraint

- Problem: $\hat{\boldsymbol{w}}^\star$ is not a fully invested portfolio: $\mathbf{1}'\widehat{\boldsymbol{w}}^\star \neq 1$.
    - It invests $\pi_{rf} = 1 - \mathbf{1}'\hat{\boldsymbol{w}}^\star$ in the risk-free asset

This seems all well and good, but ...

- Tu and Zhou (2011) force combination coefficients to sum to one:

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^\star + \kappa_2 \boldsymbol{w}_{ew} \quad \text{with} \quad \boldsymbol{c}$$

- The constraint is needed in combinations of fully invested portfolios, because it is equivalent to the full investment constraint

- Problem: $\hat{\boldsymbol{w}}^\star$ is not a fully invested portfolio: $\mathbf{1}'\widehat{\boldsymbol{w}}^\star \neq 1$.
  - It invests $\pi_{rf} = 1 - \mathbf{1}'\hat{\boldsymbol{w}}^\star$ in the risk-free asset

- We actually invest in three portfolios: the sample tangent, the risk-free asset, and the equally-weighted portfolio...

This seems all well and good, but ...

- Tu and Zhou (2011) force combination coefficients to sum to one:

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^\star + \kappa_2 \boldsymbol{w}_{ew} \quad \text{with} \quad \boldsymbol{c}$$

- The constraint is needed in combinations of fully invested portfolios, because it is equivalent to the full investment constraint

- Problem: $\hat{\boldsymbol{w}}^\star$ is not a fully invested portfolio: $\mathbf{1}'\widehat{\boldsymbol{w}}^\star \neq 1$.
  - It invests $\pi_{rf} = 1 - \mathbf{1}'\hat{\boldsymbol{w}}^\star$ in the risk-free asset

- We actually invest in three portfolios: the sample tangent, the risk-free asset, and the equally-weighted portfolio...

- ... But under $\boldsymbol{c}$, we only have one coefficient to control the weight allocated to each one of them !

This paper

- We propose to relax the constraint on combination coefficients:

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^\star + \kappa_2 \boldsymbol{w}_{ew} \quad \text{~~with~~} \quad \cancel{\boldsymbol{c}}$$

- We find the optimal combination coefficients that maximize the expected out-of-sample utility of the portfolio combination:

$$\boldsymbol{\kappa}^{opt} = \underset{\boldsymbol{\kappa}}{\text{argmax}} \quad \boldsymbol{EU} \quad \text{~~s.t.~~} \quad \cancel{\boldsymbol{c}}$$

Contributions

Contributions

Theoretical results:

1. We derive the optimal combination coefficients

Contributions

Theoretical results:

1. We derive the optimal combination coefficients
2. We highlight analytically the key benefits of relaxing the constraint

Contributions

Theoretical results:

1. We derive the optimal combination coefficients
2. We highlight analytically the key benefits of relaxing the constraint
3. *The wrong constraint helps*: the $\boldsymbol{c}$ constraint is helpful in some cases

Contributions

Theoretical results:

1. We derive the optimal combination coefficients
2. We highlight analytically the key benefits of relaxing the constraint
3. *The wrong constraint helps*: the $\boldsymbol{c}$ constraint is helpful in some cases
4. We introduce a mixed strategy combining both strategies

Contributions

Theoretical results:

1. We derive the optimal combination coefficients
2. We highlight analytically the key benefits of relaxing the constraint
3. *The wrong constraint helps*: the $\boldsymbol{c}$ constraint is helpful in some cases
4. We introduce a mixed strategy combining both strategies

Empirical results:

1. The optimal strategy always delivers a positive net utility

Contributions

Theoretical results:

1. We derive the optimal combination coefficients
2. We highlight analytically the key benefits of relaxing the constraint
3. *The wrong constraint helps*: the $\boldsymbol{c}$ constraint is helpful in some cases
4. We introduce a mixed strategy combining both strategies

Empirical results:

1. The optimal strategy always delivers a positive net utility
2. The constrained strategy may deliver negative net utility

Contributions

Theoretical results:

1. We derive the optimal combination coefficients
2. We highlight analytically the key benefits of relaxing the constraint
3. *The wrong constraint helps*: the $\boldsymbol{c}$ constraint is helpful in some cases
4. We introduce a mixed strategy combining both strategies

Empirical results:

1. The optimal strategy always delivers a positive net utility
2. The constrained strategy may deliver negative net utility
3. The mixed strategy accomplishes its mission

Optimal vs. constrained combination

Benefits of relaxing the constraint

Key benefits of relaxing the constraint;

1. Optimal exposition to the three funds

Benefits of relaxing the constraint

Key benefits of relaxing the constraint;

1. Optimal exposition to the three funds
2. Outperformance relative to the risk-free asset

Benefits of relaxing the constraint

Key benefits of relaxing the constraint;

1. Optimal exposition to the three funds
2. Outperformance relative to the risk-free asset
3. Less extreme portfolio weights

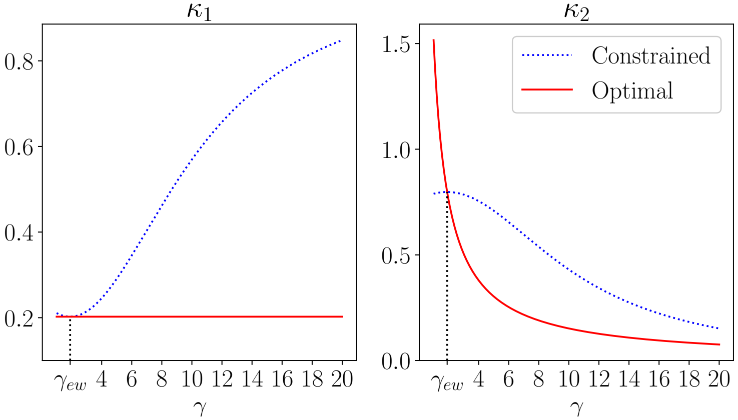
Benefits of relaxing the constraint

Key benefits of relaxing the constraint;

1. Optimal exposition to the three funds
2. Outperformance relative to the risk-free asset
3. Less extreme portfolio weights
4. Outperformance in expected out-of-sample utility (EU)
5. Outperformance relative to the two-fund rule
6. Outperformance in expected out-of-sample Sharpe Ratio

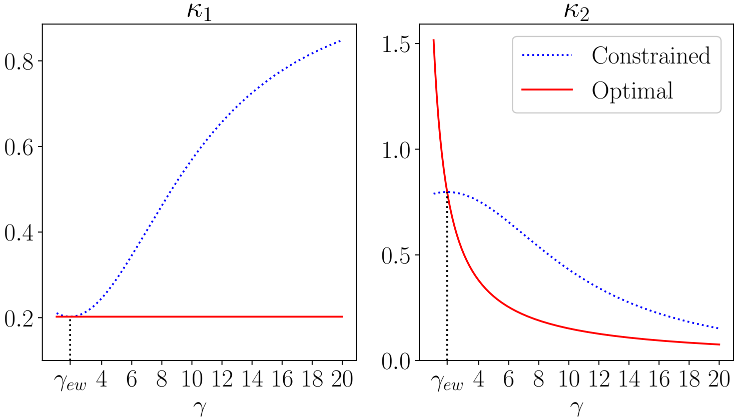
Combination coefficients

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^\star + \kappa_2 \boldsymbol{w}_{ew}$$



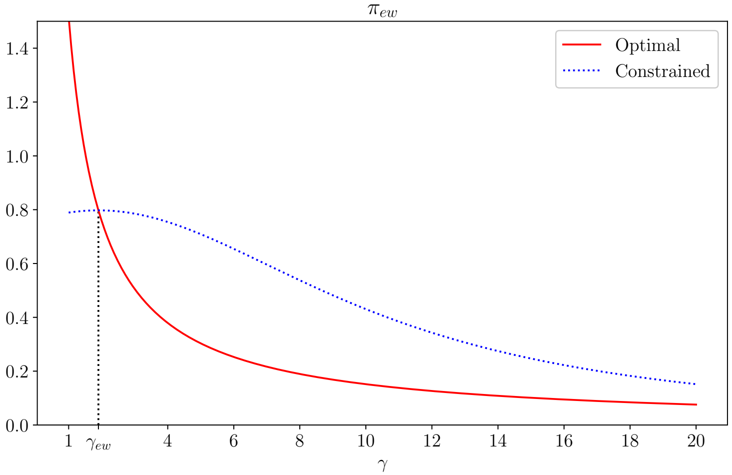
Combination coefficients

$$\hat{\boldsymbol{w}}(\boldsymbol{\kappa}) = \kappa_1 \hat{\boldsymbol{w}}^\star + \kappa_2 \boldsymbol{w}_{ew}$$



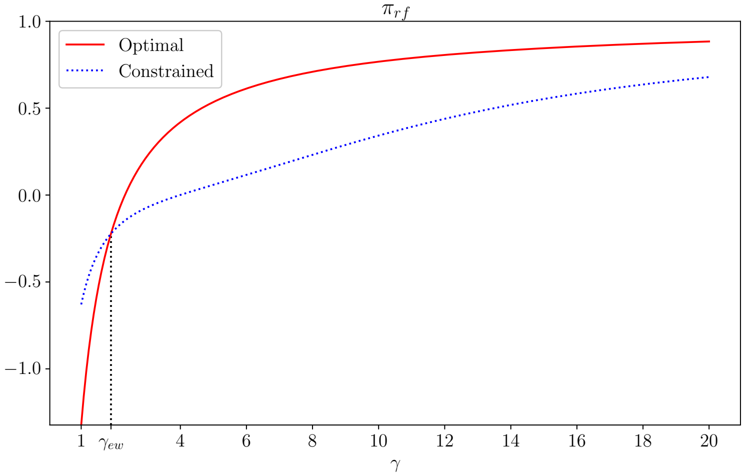
- Key facts: $\kappa_1^{tz} \geq \kappa_1^{opt}$ and equality iff $\gamma = \gamma_{ew} = \mu_{ew}/\sigma_{ew}^2$
- We use the 25 portfolios of stocks sorted on size and book-to-market (25SBTM) dataset spanning 07/26 to 12/21.

Benefit 1: Optimal exposure to the equally-weighted portfolio



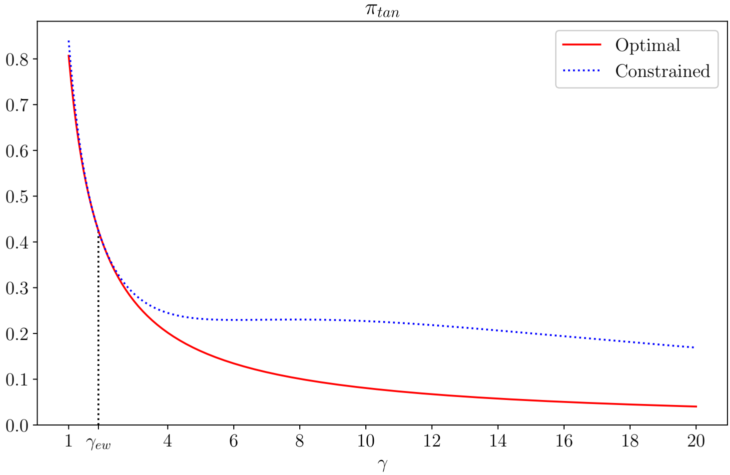
When $\gamma = 10$, $T = 120$, and $N = 25$, $\pi_{ew}^{tz} = 43\%$ and $\pi_{ew}^{opt} = 15\%$

Benefit 1: Optimal exposure to the risk-free asset



When $\gamma = 10$, $T = 120$, and $N = 25$, $\pi_{rf}^{tz} = 34\%$ and $\pi_{rf}^{opt} = 77\%$

Benefit 1: Optimal exposure to the tangent portfolio



When $\gamma = 10$, $T = 120$, and $N = 25$, $\pi_{tan}^{tz} = 23\%$ and $\pi_{tan}^{opt} = 8\%$

Benefit 2: Outperformance relative to the risk-free asset

Problem: Under mild conditions, the constrained combination can underperform the risk-free asset for investors with a risk-aversion higher than a given threshold: γ_{neg}[1]

[1] $\gamma_{neg} = 5.41$ with the 25SBTM dataset.

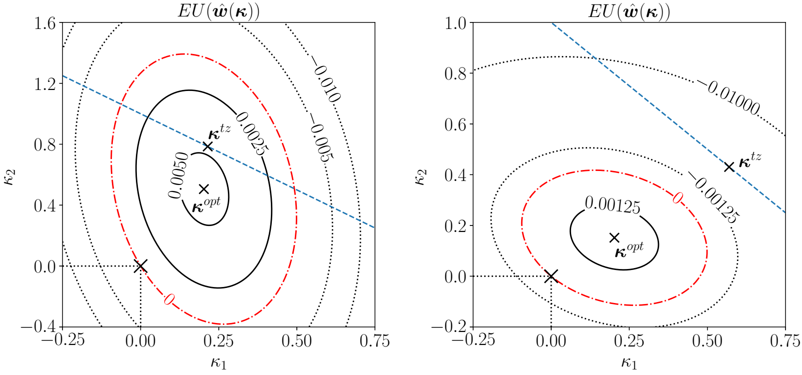
Benefit 2: Outperformance relative to the risk-free asset

Problem: Under mild conditions, the constrained combination can underperform the risk-free asset for investors with a risk-aversion higher than a given threshold: γ_{neg}[1]

Intuition: The combination should outperform each of its components, but it might not be possible because of the constraint.

[1] $\gamma_{neg} = 5.41$ with the 25SBTM dataset.

Benefit 2: Outperformance relative to the risk-free asset

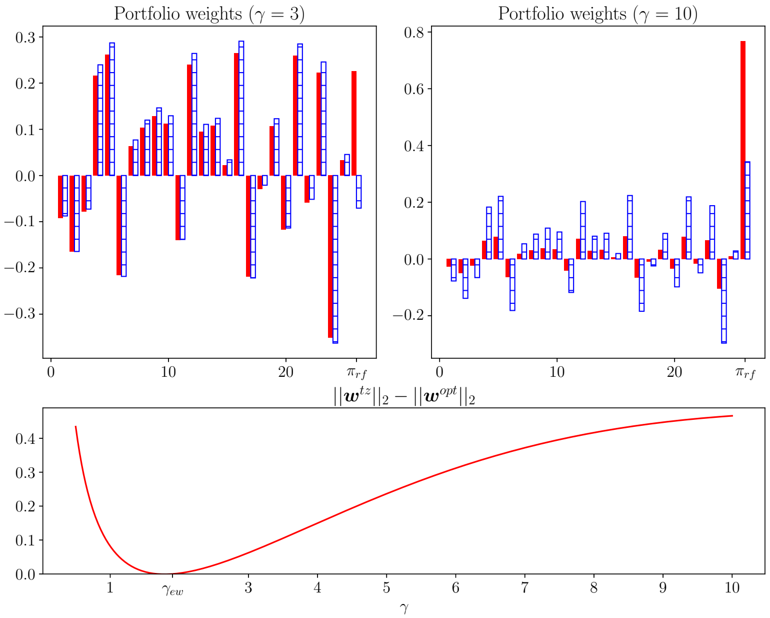


(a) $\gamma = 3 < \gamma_{neg}$ (b) $\gamma = 10 > \gamma_{neg}$

Benefit 3: Less extreme weight allocation

- Problem: The weights allocated to the risky assets by the constrained strategy are extreme.
- Why it matters: practical implementation and transaction costs

Benefit 3: Less extreme weight allocation



Mixed strategy & empirical results

Mixing $\hat{w}^{opt}$ and $\hat{w}^{tz}$

- Practical concern: $\hat{\kappa}_2^{opt}$ is unbounded and very sensitive to estimation errors in $\hat{\boldsymbol{\mu}}$ while $\hat{\boldsymbol{\kappa}}^{tz}$ is bounded in $[0, 1]$

Mixing $\hat{w}^{opt}$ and $\hat{w}^{tz}$

- Practical concern: $\hat{\kappa}_2^{opt}$ is unbounded and very sensitive to estimation errors in $\hat{\boldsymbol{\mu}}$ while $\hat{\boldsymbol{\kappa}}^{tz}$ is bounded in $[0, 1]$
- Estimation errors have a substantial impact on out-of-sample performance

Mixing $\hat{w}^{opt}$ and $\hat{w}^{tz}$

- Practical concern: $\hat{\kappa}_2^{opt}$ is unbounded and very sensitive to estimation errors in $\hat{\boldsymbol{\mu}}$ while $\hat{\boldsymbol{\kappa}}^{tz}$ is bounded in $[0, 1]$
- Estimation errors have a substantial impact on out-of-sample performance
- When theoretical gain is small, *the wrong constraint helps* (cf. Jagannathan and Ma (2003)).

Mixing $\hat{\boldsymbol{w}}^{opt}$ and $\hat{\boldsymbol{w}}^{tz}$

- Practical concern: $\hat{\kappa}_2^{opt}$ is unbounded and very sensitive to estimation errors in $\hat{\boldsymbol{\mu}}$ while $\hat{\boldsymbol{\kappa}}^{tz}$ is bounded in $[0,1]$
- Estimation errors have a substantial impact on out-of-sample performance
- When theoretical gain is small, *the wrong constraint helps* (cf. Jagannathan and Ma (2003)).
- We define a symmetric interval around γ_{ew}: $[\gamma_{ew} \pm \epsilon]$. In this interval, $\hat{\boldsymbol{w}}^{tz}$ is preferred to $\hat{\boldsymbol{w}}^{opt}$:

$$\hat{\boldsymbol{w}}^{opt,tz} = \begin{cases} \hat{\boldsymbol{w}}^{tz} & \text{if } \ \gamma \in [\gamma_{ew} \pm \epsilon], \\ \hat{\boldsymbol{w}}^{opt} & \text{otherwise.} \end{cases}$$

Empirical results

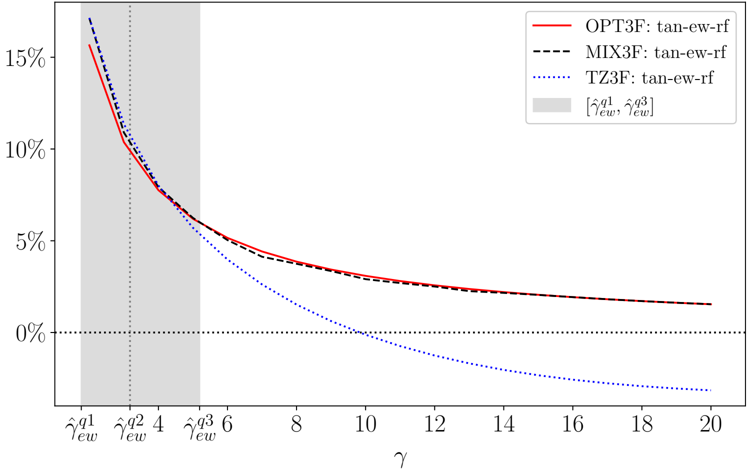
Dataset	N	Time period	Abbr.
25 portfolios formed on size and book-to-market	25	07/1926 - 12/2021	25SBTM
10 momentum portfolios	10	01/1927 - 12/2021	10MOM
25 portfolios formed on OP[2] and investment	25	07/1963 - 12/2021	25OPINV
48 industry portfolios	48	07/1969 - 12/2021	48IND

Table 1: List of datasets used in the empirical analysis (monthly frequency)

- Different rolling window sizes
- Various estimators for $\hat{\boldsymbol{\Sigma}}$: sample, linear shrinkage, nonlinear shrinkage (Ledoit & Wolf (2004, 2017))
- Monthly rebalancing and proportional transaction costs of 10bps

[2] Operating profitability.

Empirical results



Takeaways

Takeaways

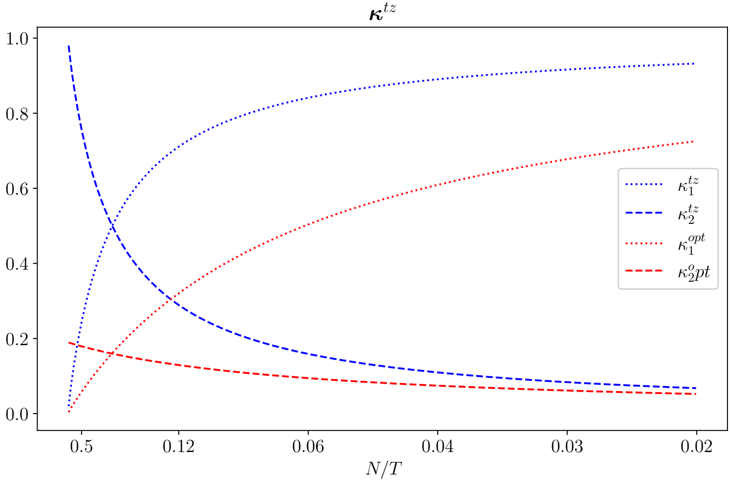
1. We derive the optimal combination coefficients
2. We highlight analytically the key benefits of relaxing the constraint
3. *The wrong constraint helps*: the $\boldsymbol{c}$ constraint is helpful in some cases
4. We introduce a mixed strategy combining both strategies
5. The optimal strategy always delivers a positive net utility
6. The constrained strategy may deliver negative net utility
7. The mixed strategy accomplishes its mission

Thank you!

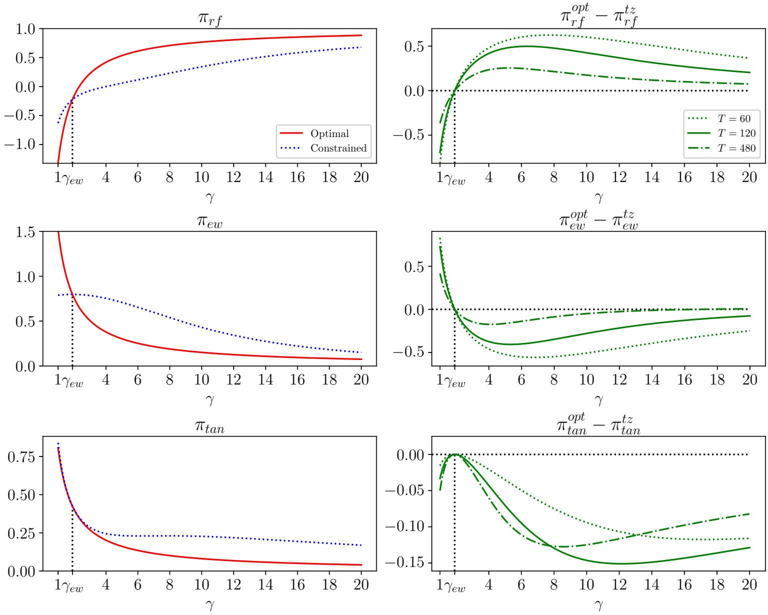
Please send comments to

rodolphe.vanderveken@uclouvain.be

Impact of estimation risk



Optimal exposure to the three funds + estimation risk



Combination coefficients

The constrained combination coefficients are

$$\kappa_1^{tz} = \frac{\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2}{\psi^2 + \sigma_{ew}^2(\gamma - \gamma_{ew})^2 + d} \in [0,1] \quad \text{and} \quad \kappa_2^{tz} = 1 - \kappa_1^{tz} \in [0,1].$$

The optimal combination coefficients are

$$\kappa_1^{opt} = \frac{\psi^2}{\psi^2 + d} \in [0,1] \quad \text{and} \quad \kappa_2^{opt} = \frac{\gamma_{ew}}{\gamma}(1 - \kappa_1^{opt}) \in \mathbb{R},$$

d is a parameter that increases with estimation risk N/T

γ_{neg} threshold definition

$$\gamma > \gamma_{ew}\left(1 + \sqrt{1 + \frac{\theta^4/\theta_{ew}^2}{d - \theta^2}}\right) = \gamma_{neg}.$$

Notation

- Sample mean-variance portfolio (SMV) (from sample of size T)

$$\hat{\boldsymbol{w}}^{\star} = \frac{1}{\gamma} \widehat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\mu}}$$

Notation

- Sample mean-variance portfolio (SMV) (from sample of size T)

$$\hat{\boldsymbol{w}}^\star = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}$$

- We define $\mu_{ew} = \boldsymbol{w}_{ew}'\boldsymbol{\mu}$ and $\sigma_{ew}^2 = \boldsymbol{w}_{ew}'\boldsymbol{\Sigma}\boldsymbol{w}_{ew}$. Moreover, we define $\gamma_{ew} = \mu_{ew}/\sigma_{ew}^2$

Notation

- Sample mean-variance portfolio (SMV) (from sample of size T)

$$\hat{\boldsymbol{w}}^{\star} = \frac{1}{\gamma}\widehat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}$$

- We define $\mu_{ew} = \boldsymbol{w}_{ew}'\boldsymbol{\mu}$ and $\sigma_{ew}^2 = \boldsymbol{w}_{ew}'\boldsymbol{\Sigma}\boldsymbol{w}_{ew}$. Moreover, we define $\gamma_{ew} = \mu_{ew}/\sigma_{ew}^2$

- We define $\theta^2 = \boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ as the maximum squared Sharpe ratio, $\theta_{ew}^2 = \mu_{ew}^2/\sigma_{ew}^2$ as the squared Sharpe ratio of the EW portfolio, and ψ^2 as the difference between the two:

$$\psi^2 = \theta^2 - \theta_{ew}^2 \geq 0.$$

Sample estimators

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{R_t}, \quad \widehat{\boldsymbol{\Sigma}} = \frac{1}{T-N-2}\sum_{t=1}^{T}(\boldsymbol{R_t} - \hat{\mu})(\boldsymbol{R_t} - \hat{\mu})'.$$

Efficient frontier - Short reminder