Consistency of robust portfolio estimates.

Optimization in Finance 2007 | Coimbra | 26.10.2007 | Ralf Werner

Hypo Real Estate HOLDING

Outline.

Portfolio optimization.
Traditional Markowitz model.

Markowitz framework.

- Set of feasible portfolios $X \subset \mathbb{R}^n$ (convex, compact), i.e. $x^T \mathbf{1} = 1$ for $x \in X$.
- Expected asset returns $\boldsymbol{r} = \mathbf{E}[R]$.
- Covariance matrix of asset returns $\boldsymbol{C} = \mathbf{Var}[R]$.
- Expected portfolio return $x^T \boldsymbol{r}$.
- Volatility (= risk) of portfolio $\sqrt{x^T \boldsymbol{C} x}$.

Markowitz portfolio optimization problem.

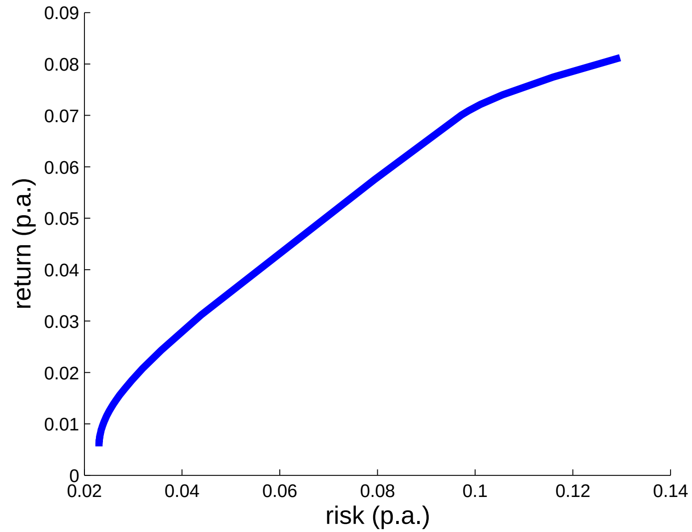
$$\min_{x \in X} \quad (1-\lambda)\sqrt{x^T \boldsymbol{C} x} + \lambda(-x^T \boldsymbol{r}) \tag{PO}$$

- Risk-return trade-off parameter λ (with $0 \leq \lambda \leq 1$).
- Optimal portfolio x^* depends on $\boldsymbol{r}$ and $\boldsymbol{C}$, i.e. $x^* = x^*_{\boldsymbol{r},\boldsymbol{C}}$.

Portfolio optimization.
Traditional Markowitz model.

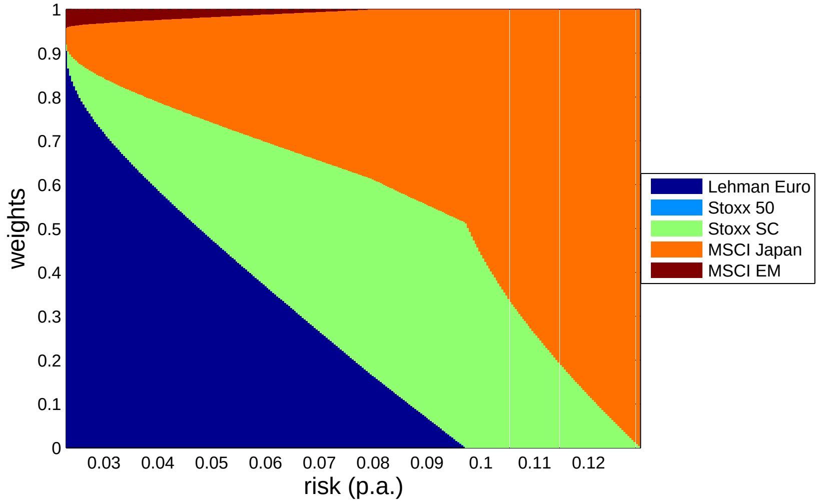
Illustration.

- The trade-off parameter λ is used to trace the **efficient frontier**.
- For $\lambda = 0$ we get the **minimum variance portfolio**.
- For $\lambda = 1$ we get the **maximum return portfolio**.



Remark.

- The calculation of the efficient frontier can also be formulated as a vector optimization problem.



Hypo Real Estate HOLDING

Portfolio optimization.
Traditional Markowitz model.

Market model.

- Elliptical model for asset returns $R \sim \mathcal{E}(r, C, \varphi)$ with density generator φ.
- Elliptical models contain the multivariate normal distribution as a special case.
- Elliptical models are still compatible with preference/utility theory.

Estimation of the input parameters r and C.

- We assume that S historical return realizations $R_1, \ldots, R_S$ (iid) are given.
- In the traditional Markowitz framework, maximum likelihood estimators for r and C are used to get the input data for (PO)

$$\hat{\mu}_S^{ML} = \frac{1}{S}\sum_{s=1}^{S} R_s, \qquad \hat{\Sigma}_S^{ML} = \frac{1}{S}\sum_{s=1}^{S}(R_s - \hat{\mu}_S^{ML})(R_s - \hat{\mu}_S^{ML})^T$$

- Other approaches like Bayes estimator, Black-Litterman estimators or robust estimators are also used frequently.

Portfolio optimization.
Estimation risk.

True and actual efficient portfolio.

- For market parameters (r, C), the **true efficient portfolio** is $x^*_{r,C}$.
- As only estimators (μ, Σ) are available, the **actual efficient portfolio** is $x^*_{\mu,\Sigma}$.
- The actual portfolio can be seen as an estimator for the true efficient portfolio.

True, actual and predicted risk and return.

	true	actual	predicted
expected return	${x^*_{r,C}}^T r$	${x^*_{\mu,\Sigma}}^T r$	${x^*_{\mu,\Sigma}}^T \mu$
risk	$\sqrt{{x^*_{r,C}}^T C x^*_{r,C}}$	$\sqrt{{x^*_{\mu,\Sigma}}^T C x^*_{\mu,\Sigma}}$	$\sqrt{{x^*_{\mu,\Sigma}}^T \Sigma x^*_{\mu,\Sigma}}$

Estimation risk.

- Estimation risk = true quantity – predicted quantity.

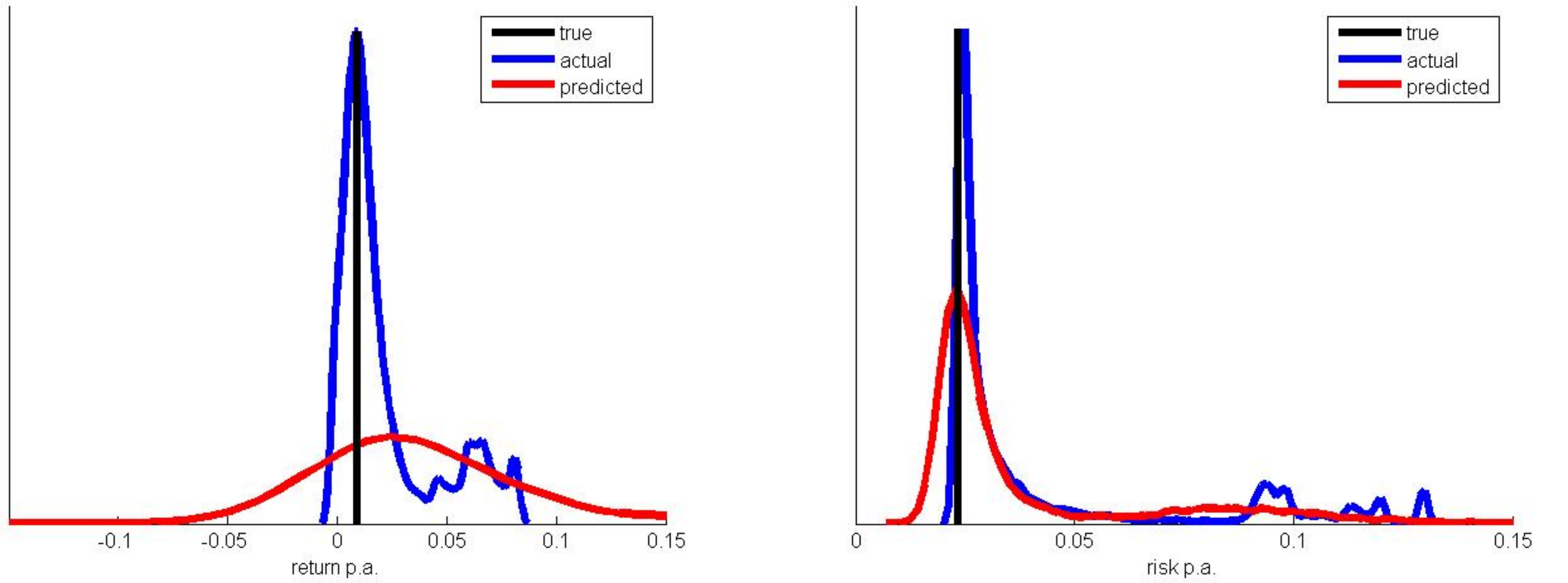
How big is this estimation risk?

Hypo Real Estate HOLDING

Portfolio optimization.
Estimation risk.

Example.

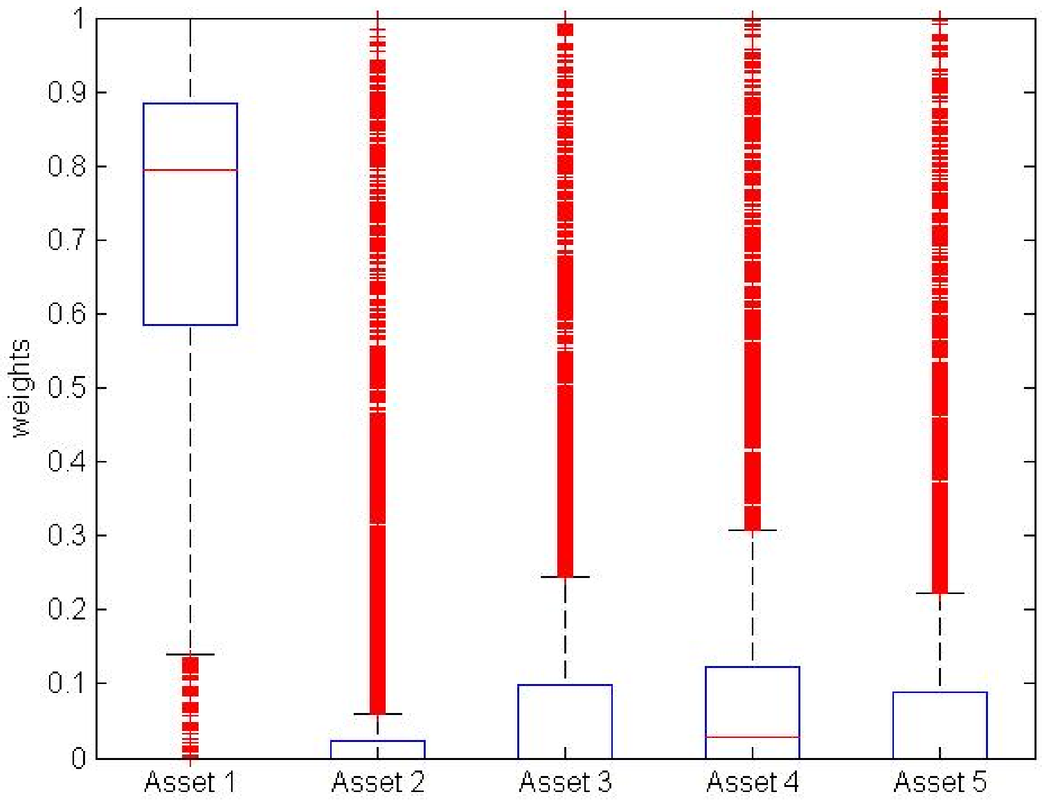
- The actual risk and return figures deviate from the optimal ones.
- The predicted return figures show significant deviations.



Portfolio optimization.
Estimation risk.

Example (cont'd).

- The weights vary strongly, sometimes even dramatically (i.e. the outliers).



Portfolio optimization.
Estimation risk.

Brief summary of known results.

- Estimation risk was an active research topic from late 70's until early 90's.
- Most popular papers: Barry, Jobson/Korkie, Bawa/Brown/Klein, ...
- Main (empirical) result: optimal portfolios strongly depend on input r and C.

Is estimation risk vanishing with increasing sample size S?

- Jobson/Korkie: if $S > 200$, estimation risk can be neglected.
- Random matrix theory: the ratio of S to n must be large.
- The appropriate notion from statistics is **consistency**.
- An even better property allowing for some quantitative estimate is **asymptotic normality**.

Portfolio optimization.
Consistency.

Definition.

A point estimator $\mathcal{Q}_{p,S}$ for a parameter p based on a sample of size S is called

- **unbiased**, if $\mathbf{E}[\mathcal{Q}_{p,S}] = p$,
- **strongly consistent**, if $\mathbf{P}\left[\lim_{S\to\infty} \mathcal{Q}_{p,S} = p\right] = 1$ (convergence almost surely),
- **(weakly) consistent**, if $\lim_{S\to\infty} \mathbf{P}\left[|\mathcal{Q}_{p,S} - p| > \varepsilon\right] = 0$ (convergence in probability).

Remarks.

- Almost sure convergence and convergence in probability remain valid after continuous transformations.
- The portfolio estimator is in general biased, even if unbiased estimators for the input data are applied.

Portfolio optimization.
Consistency.

Main results concerning consistency and asymptotic normality.

- Jobson/Korkie (1980): The optimal solution x^* is consistent and asymptotically normal, if R is normal. This result is derived from an analytical solution for x^* based on a special structure of X.
- Mori (2004): Extension to the case that X includes linear equalities.
- Lauprete (2002): Similar results to Mori, but with R being elliptic and a slightly different optimization problem.
- Jobson/Korkie (1980s) also characterized the distribution of x^* for small S.
- Okhrin/Schmid (2006): Extension of the Jobson/Korkie results to elliptical distributions.

Consistency for a general set X and R elliptic is still missing!

Portfolio optimization.
Consistency.

Theorem.

Assume that the following convex optimization problem (with convex, compact X)

$$\begin{aligned} \min_{x \in X} \quad & f(x,u) \\ \text{s.t.} \quad & g(x,u) \leq 0 \end{aligned}$$

has an unique optimal solution $x^*(u)$ in a neighborhood of $\hat{u}$. Then x^* is continuous at $\hat{u}$, if

- the objective f and the constraint g are continuous, and either
- the constraint g is not depending on u, or
- there exists a Slater point for $\hat{u}$, i.e. $\exists\, \hat{x} \in X$ such that $g(\hat{x}, \hat{u}) < 0$.

Corollary.

The optimal portfolio $x^*_{r,C}$ is continuously depending on (r, C).

Portfolio optimization.
Consistency.

Theorem (Schöttle, Werner – 2006).

Let μ and Σ be consistent estimators for r and C. Then the optimal solution $x^*_{\mu,\Sigma}$ is also a consistent estimator for $x^*_{r,C}$.

Remarks.

- The above result generalizes all existing results.
- The key to consistency is continuity of the solution of (PO) with respect to the parameters.
- Thus, the result can easily be generalized to the case that X depends (Hausdorff) continuously on r and C.
- Uniqueness of x^* follows from the strict convexity of $x \mapsto \sqrt{x^T C x}$ on X.
- For asymptotic normality, we need differentiability of x^* with respect to the input parameters r and C.

Outline.

1. Traditional portfolio optimization.

   1.1 Markowitz optimization

   1.2 Estimation risk

   1.3 Consistency

2. **Robust portfolio optimization.**

   2.1 Robust counterparts

   2.2 Estimation risk

   2.3 Consistency

Robust portfolio optimization.
Robust counterpart.

The need for robustification.

- Although stability is given from a mathematical point of view, dependence on the parameters is still unsatisfactorily high.
- In the last 15 years, several approaches were introduced to reduce the estimation error while keeping efficiency as high as possible.
  - Usage of more robust estimators (shrinkage, M-estimators, ...)
  - Michaud's resampling,
  - Stochastic optimization and scenario optimization,
  - **Robust counterpart**.
- Several empirical studies support the usage of robust approaches for small sample sizes S.
- Robustification usually decreases the estimation variance, but at the same time introduces a bias in the estimation.

Robust portfolio optimization.
Robust counterpart.

The robust counterpart.

Based on an uncertainty set U the robust counterpart is defined as

$$\begin{array}{ll} \min\limits_{x \in X} & f(x,u) \\ \text{s.t.} & g(x,u) \leq 0 \end{array} \qquad \begin{array}{ll} \min\limits_{x \in X} & \max\limits_{u \in U} f(x,u) \\ \text{s.t.} & \max\limits_{u \in U} g(x,u) \leq 0. \end{array}$$

Setting $F(x,U) := \max\limits_{u \in U} f(x,u)$ and $G(x,U) := \max\limits_{u \in U} g(x,u)$ this becomes

$$\begin{array}{ll} \min\limits_{x \in X} & F(x,U) \\ \text{s.t.} & G(x,U) \leq 0 \end{array}$$

Robust portfolio optimization

$$\min_{x \in X} \max_{(\mu,\Sigma) \in U} \quad (1-\lambda)\sqrt{x^T \Sigma x} + \lambda(-x^T \mu) \tag{RO}$$

Robust portfolio optimization.
Robust counterpart.

Important facts about the robust counterpart (Schöttle, Werner – 2006).

- It holds: f, g convex in $x \implies F$, G convex in x.
- It holds: f, g strictly convex on $X \implies F$, G strictly convex on X.
- It holds: f, g continuous in $u \implies F$, G continuous in U.
- Continuity in U is always understood in the Hausdorff sense.
- If the original problem has a Slater point and U is small enough, then the robust counterpart also possesses a Slater point.

Interpretation.

- The robust counterpart inherits all nice properties from the original problem.
- Instead of a real parameter $u \in \mathbb{R}^d$ a set $U \in 2^{\mathbb{R}^d}$ becomes the parameter.

Hypo Real Estate HOLDING

Robust portfolio optimization.
Robust counterpart.

Choice of the uncertainty set U.

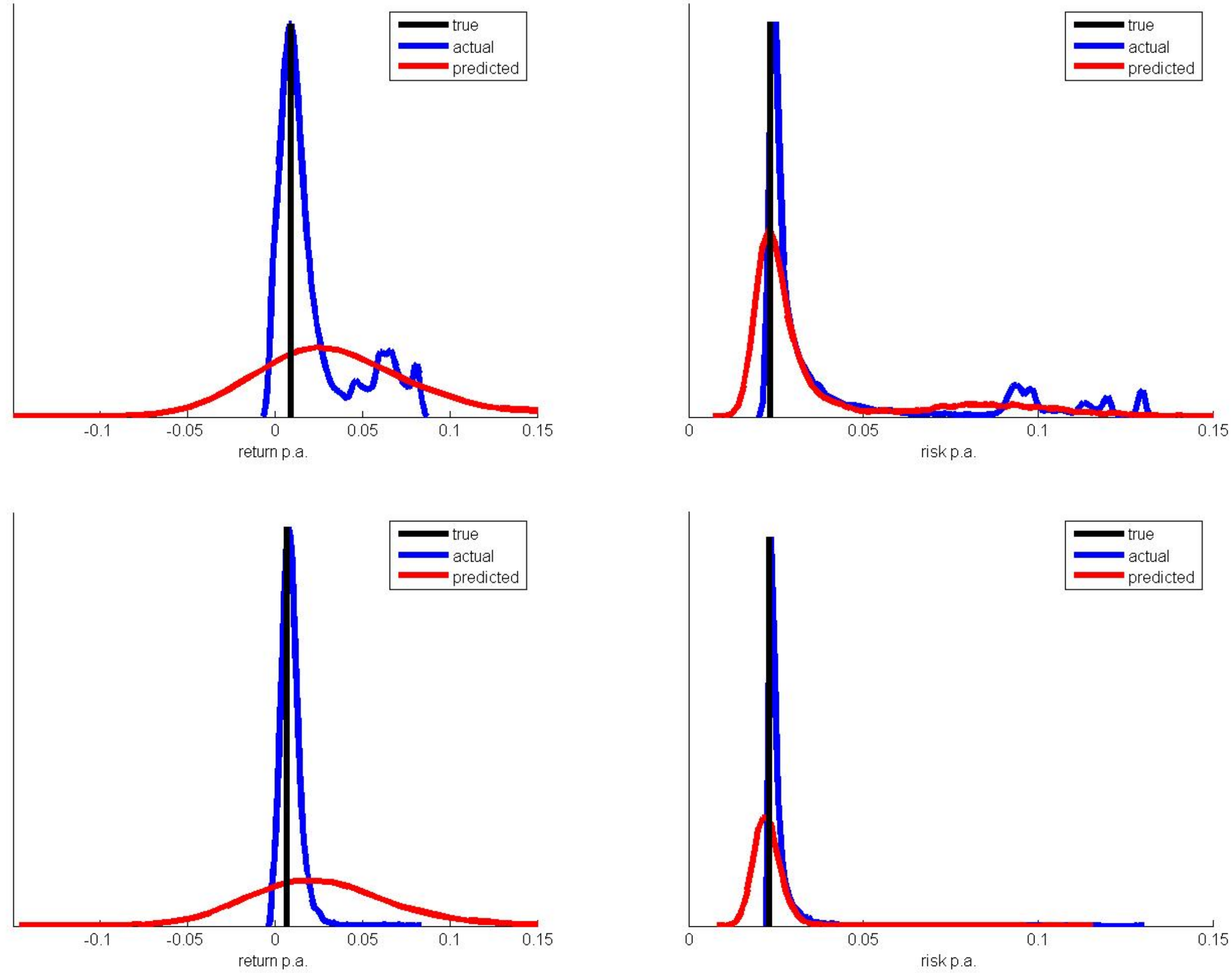
- Most obvious choice for U is the (joint) confidence ellipsoid centered at the estimates $\hat{\mu}$ and $\hat{\Sigma}$.
- In the special case of normally distributed returns and maximum likelihood estimators, the uncertainty set can be explicitely described by

$$U = \{(\mu, \Sigma) \mid S(\mu - \hat{\mu})^T \hat{\Sigma}^{-1}(\mu - \hat{\mu}) + \frac{S-1}{2} \|\hat{\Sigma}^{-\frac{1}{2}}(\Sigma - \hat{\Sigma})\hat{\Sigma}^{-\frac{1}{2}}\|_{\mathrm{tr}}^2 \leq \delta^2\}.$$

- Generalizations to elliptical distributions and other estimators are in general possible, but may involve numerical procedures (i.e. numerical integration, etc.).
- Other – mainly polyhedral – uncertainty sets have also been investigated in the literature.
- For small S the shape of U plays an important role, but for large S, only the size of U matters.

Robust portfolio optimization.
Estimation risk.

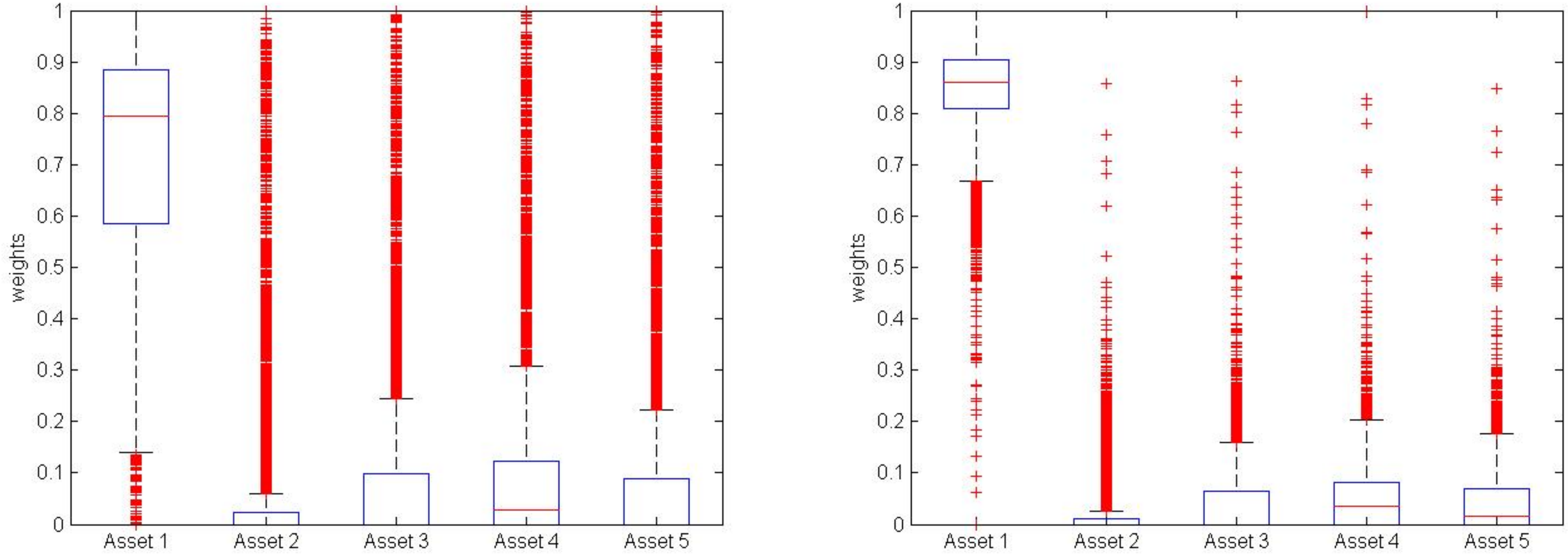
Robustifcation reduces estimation risk.



Hypo Real Estate HOLDING

Robust portfolio optimization.
Estimation risk.

Robustified portfolios are more stable.



- Stability in weights comes with a small bias in portfolio weights.

Robust portfolio optimization.
Consistent uncertainty sets.

Definition.

An uncertainty set U is called **strongly consistent** for the pair (r, C) if

$$H_d\big(U, \{(r, C)\}\big) \to 0 \quad \text{almost surely for } S \to \infty,$$

with $H_d(A, B)$ denoting the Hausdorff distance between the sets A and B.

Remarks.

- (Weak) consistency can be defined analogously (by convergence in probability).
- Consistent uncertainty sets are the natural analogon to consistent point estimates.
- Consistent uncertainty sets shrink to the real data.
- The uncertainty set from the previous example is strongly consistent.

Portfolio optimization.
Consistency.

Theorem.

Assume that the robust counterpart

$$\begin{aligned} \min_{x \in X} \quad & F(x, U) \\ \text{s.t.} \quad & G(x, U) \leq 0 \end{aligned}$$

has an unique optimal solution $x^*(\hat{U})$ in a neighborhood of $\hat{U}$. Then x^* is continuous at $\hat{U}$, if

- the objective F and the constraint G are continuous, and either
- the constraint G is not depending on U, or
- there exists a Slater point for $\hat{U}$.

Remark.

Not surprisingly, this is the same result as for the original problem.

Robust portfolio optimization.
Consistency.

Theorem (Schöttle, Werner – 2006).

Let U be a consistent uncertainty set for (r, C). Then the optimal solution x^*_U is a consistent estimator for $x^*_{r,C}$.

Remarks.

- The above result generalizes the result of the traditional framework.
- The key to consistency is continuity of the solution with respect to the uncertainty set.
- Thus, the result can be easily generalized to the case that X depends (Hausdorff) continuously on r and C.
- Uniqueness of x^* follows from the strict convexity of $x \mapsto \sqrt{x^T C x}$ on X.
- What about asymptotic normality?

Outline.

1. Traditional portfolio optimization.

    1.1 Markowitz optimization

    1.2 Estimation risk

    1.3 Consistency

2. **Resampled portfolio optimization.**

    2.1 Resampled portfolios

    2.2 Estimation risk

    2.3 Consistency

Resampled portfolio optimization.
Resampled asset returns and bootstrapped estimators.

Resampled asset returns.

- Fix resampling parameters $(r_{res}, C_{res}, \psi_{res})$ for **resampled asset returns**

$$R_{res} \sim \mathcal{E}(r_{res}, C_{res}, \psi_{res}).$$

The bootstrapped estimator distribution.

- Take S samples of R_{res} and use any **continuous and consistent estimator** to obtain μ_{res} and Σ_{res}.
- This induces the **bootstrapped distribution** $\mathcal{B}_S$ for μ_{res} and Σ_{res}:

$$(\mu_{res}, \Sigma_{res}) \sim \mathcal{B}_S(r_{res}, C_{res}, \psi_{res}).$$

Example.

- In Michaud's original setting: $R_{res} \sim \mathcal{N}(r_{res}, C_{res})$.
- Using the maximum likelihood estimators, the bootstrapped distribution is analytically given: $\mathcal{B}_S(r_{res}, C_{res}, \psi_{res}) = \mathcal{N}(r_{res}, \frac{1}{S}C_{res}) \otimes \mathcal{W}(\frac{1}{S}C_{res}, S-1)$.

Resampled portfolio optimization.
Resampled portfolios.

Resampled portfolios.

- Plug in μ_{res} and Σ_{res} in (PO) to obtain $x^*_{\mu_{res},\Sigma_{res}}$.
- Based on the distribution of the bootstrapped x^* the **resampled portfolio** is defined as:

$$y^*_{r_{res},C_{res}} := y^*_{r_{res},C_{res},S,\psi_{res}} := \mathbf{E}[x^*_{\mu_{res},\Sigma_{res}}] \quad \text{with} \quad (\mu_{res},\Sigma_{res}) \sim \mathcal{B}_S(r_{res},C_{res},\psi_{res})$$

Resampled portfolios – algorithmic view.

1. Fix resampling parameters and resample S asset returns.
2. Calculate estimators for risk and return.
3. Plug them into the portfolio problem and compute optimal portfolios.
4. Repeat the above K times and take the average of all these portfolios.

Where do the resampling parameters (r_{res}, C_{res}) come from?

Resampled portfolio optimization.
Resampled portfolios.

Important facts about resampled portfolios.

- For small S, the choice of ψ_{res} is important. For large S, the density generator can be chosen arbitrarily.
- The resampling parameters r_{res} and C_{res} are estimated from the historical sample $R_1, \ldots, R_S$.
- The estimators which are used to derive r_{res} and C_{res} are used for the estimation of the bootstrapped estimators μ_{res} and Σ_{res} as well.

Continuity properties.

- For fixed S, the resampled portfolio is continuous in the resampling parameters:

$$y^*_{r_k,C_k,S,\psi_{res}} \to y^*_{\bar{r},\bar{C},S,\psi_{res}} \quad \text{for } r_k \to \bar{r},\ C_k \to \bar{C}.$$

- For fixed resampling parameters, it holds independent of ψ_{res}:

$$y^*_{r_{res},C_{res},S,\psi_{res}} \to x^*_{r_{res},C_{res}} \quad \text{for } S \to \infty.$$

Resampled portfolio optimization.
Consistency.

Theorem (Schöttle, Werner – 2006).

Let r_{res} and C_{res} be derived by continuous and consistent estimators for r and C. Then the resampled portfolio $y^*_{r_{res},C_{res}}$ is a consistent estimator for $x^*_{r,C}$, independent of the choice of ψ_{res}.

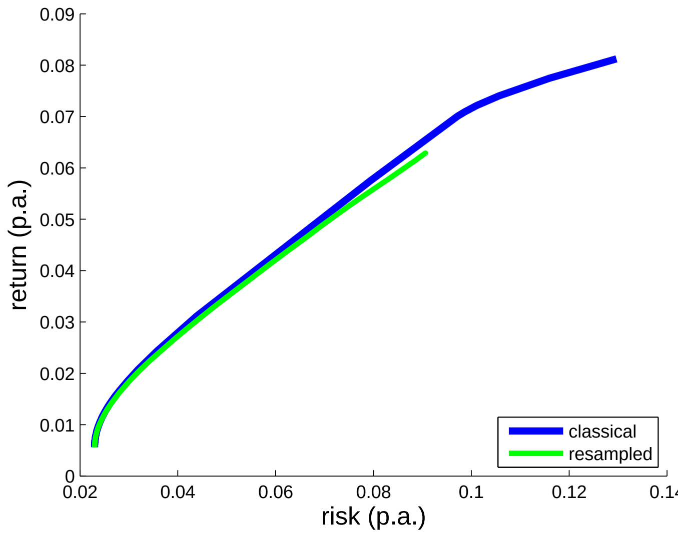
Remarks.

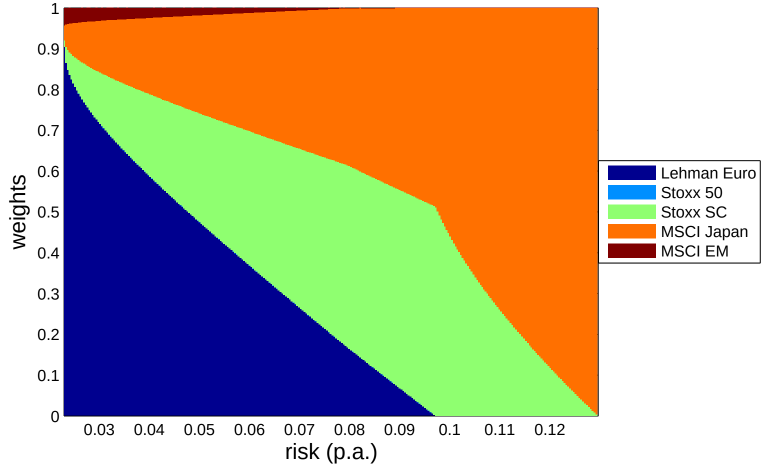
- The key to consistency is again continuity of the solution of (PO) with respect to the uncertain parameters.
- Thus, the result can be easily generalized to the case that X depends (Hausdorff) continuously on r and C.
- Compactness of X is crucial for the above continuity result.
- What about asymptotic normality?

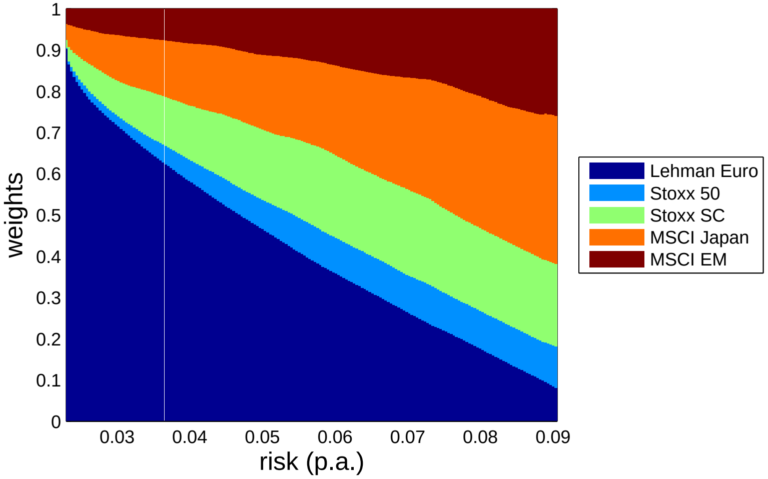
Resampled portfolio optimization.
Costs and benefits of resampling.

Observations.

- The resampled frontier is close to the original frontier.
- The resampled frontier is shorter than the original frontier.
- The resampled portfolio allocations look more reasonable.







Hypo Real Estate

HOLDING