Less is More: Ranking Information, Estimation Errors and Optimal Portfolios

37th Australasian Finance and Banking Conference

Lukas Salcher, Sebastian Stöckl

University of Liechtenstein

Monday 16th December, 2024

Modern Portfolio Theory

- Central idea is Mean-Variance-Optimization:
  - Estimate expected returns and (co-)variances of assets
  - Maximize return per unit of risk (i.e. the Sharpe-ratio)
- Optimized weights lead to poor out-of-sample performance:
  - Optimization is very sensitive to input parameters
  - Higher weights for assets with
    $\Rightarrow$ high expected return and low (co-)variance
    (Michaud, 1989; DeMiguel et al., 2009)
    $\Rightarrow$ larger estimation errors ("estimation error maximiser", Michaud, 1989)
    ▸ Illustration

Handling Parameter Uncertainty in Portfolio Management

- Different techniques have been developed:
  - Portfolio constraints (Frost and Savarino, 1988; MacKinlay and Pástor, 2000)
  - Bayesian-extensions (Jorion, 1985)
  - Shrinkage (Ledoit and Wolf, 2003; 2004; Barroso and Saxena, 2021)
  - Robust portfolio optimization (Garlappi et al., 2007; Kan and Zhou, 2007)
- None of these consistently outperforms the 1/N portfolio (DeMiguel et al., 2009)
- Reasons:
  - Historical forecasts would require huge amounts of data which are usually not available (DeMiguel et al., 2009)
  - But: Even forecasts with modest predictive power (i.e. R^2_{OOS} of 0.5%) improve performance (Allen et al., 2019)

Motivation

Two lessons:

1. Superior OOS performance of 1/N-portfolio suggest dropping all asset-specific information
2. Superior OOS performance of optimal portfolio using forecasts with modest predictive power
3. $\Rightarrow$ Reducing informational content of input parameters improves performance

Motivation

Two lessons:

1. Superior OOS performance of 1/N-portfolio suggest dropping all asset-specific information
2. Superior OOS performance of optimal portfolio using forecasts with modest predictive power
3. $\Rightarrow$ Reducing informational content of input parameters improves performance

Question:

- Is there an ideal level of informational content to address the tradeoff between flawed parameter estimates and not using any information at all?

Motivation

Two lessons:

1. Superior OOS performance of 1/N-portfolio suggest dropping all asset-specific information
2. Superior OOS performance of optimal portfolio using forecasts with modest predictive power
3. $\Rightarrow$ Reducing informational content of input parameters improves performance

Intuition:

- Reducing information (cardinal $\rightarrow$ ordinal $\rightarrow$ nominal) in accordance with the data should reduce estimation errors and thereby improve portfolio performance

Data

- As test assets we use 49 Fama-French industries. We take 100 random draws of 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 and 30 industries. (Potential extension to $5, \ldots, 100$ assets drawn from the S&P 500).
- Aggregate Short Interest (SII) taken from Rapach et al. (2016)
- Time frame: January 1995 to December 2019
- Various techniques to come up with expected returns:
  - Historical mean returns over various horizons
  - CAPM-based forecasts using different predictor variables (e.g.: short interest, variance risk premia, financial uncertainty) CAPM-Forecasts
  - Machine-learning based predictions
  - Forecasts based on simple predictive regressions using aforementioned variables

From Expected Returns to Ranks and Back

- We create our views through ordinal views on the expected asset returns:
  1. Full ranking: $\mu_1 < \mu_2 < \cdots < \mu_5$
  2. Three sub-groups: $\mu_1 = \mu_2 < \mu_3 < \mu_4 = \mu_5$
  3. Two sub-groups: $\mu_1 = \mu_2 = \mu_3 < \mu_4 = \mu_5$
  4. One sub-group: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$
- We implement our views through ▸ Entropy-Pooling:
  - We follow Meucci (2010) given historical scenarios (observations) with prior distribution $\underline{p} := \left(\frac{1}{H}, \ldots, \frac{1}{H}\right)$ we let $\bar{p}$ be defined by

$$\bar{p} := \underset{p \in \mathcal{V}}{\operatorname{argmin}} \, \mathcal{E}\left(p | \underline{p}\right) \tag{1}$$

    where all $\bar{p}_s \geq 0$ and $\sum_{s=t-H+1}^{t} \bar{p}_s = 1$ for all joint historical observations.
  - Posterior distribution will weigh selected historic observations so that all scenarios that do not fulfill the ranking get a smaller weight.

Mean-Variance Framework

- We then use the moments of the posterior distribution as inputs for mean-variance optimization:

$$w_{\mathcal{V}} = \frac{\hat{\mu}_{\mathcal{V}} \hat{\Sigma}_{\mathcal{V}}^{-1}}{\hat{\mu}_{\mathcal{V}} \hat{\Sigma}_{\mathcal{V}}^{-1} 1'}, \text{ where} \tag{2}$$

$$\hat{\mu}_{\mathcal{V}} = \sum_{s=t-H+1}^{t} \bar{p}_s r_{s,\cdot} \text{ and} \tag{3}$$

$$\hat{\Sigma}_{\mathcal{V}} = \sum_{s=t-H+1}^{t} \bar{p}_s \left(r_{s,t} - \hat{\mu}_{s,t}\right) \left(r_{s,t} - \hat{\mu}_{s,t}\right)' \tag{4}$$

- Portfolio constraints:
  - Weights of assets have to lie between -1 and 1
  - Sum of weights has to lie between 0 and 10
  - For robustness: long-only constraint
- Benchmarks: 1/N portfolio and plug-in portfolio (based on exact forecasts)

Performance

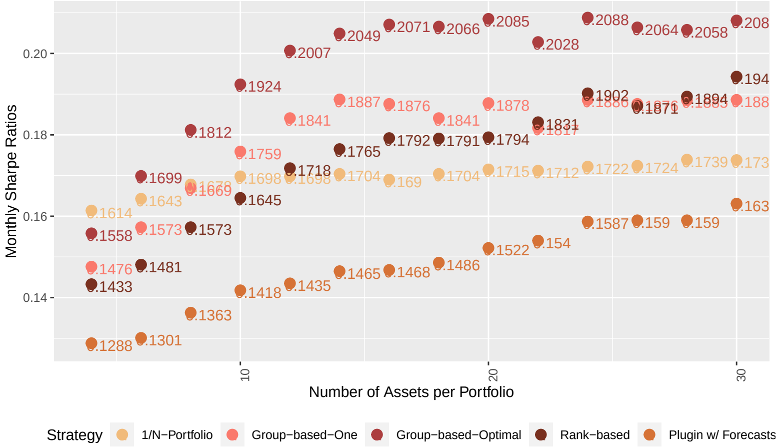

Figure: Monthly Sharpe ratios of rank-based and group-based mean-variance portfolios, plug-in mean-variance portfolios and 1/N portfolios. Portfolios are based on randomly drawn industries and CAPM forecasts.

Out-Performance vs No. of Assets

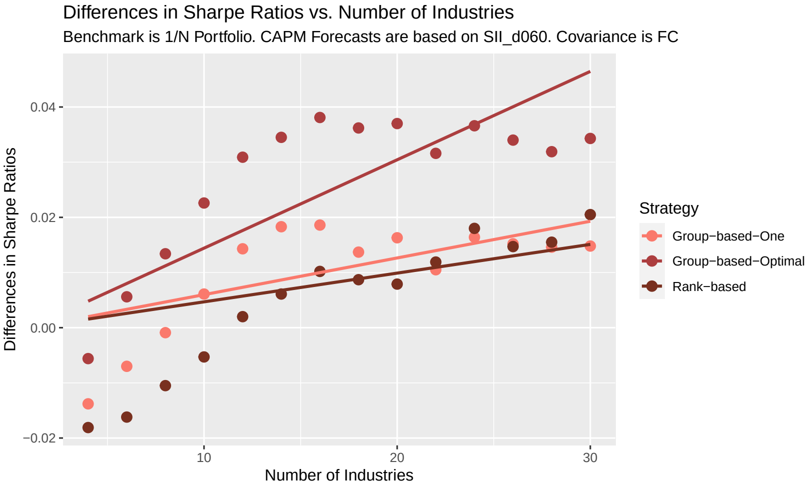

Figure: Benchmark regressions: This figure presents the slope coefficients of plotting Sharpe ratio differences vis-a-vis two bechmark portfolios against the number of assets (100 random draws from FF49 industries, CAPM forecasts).

Out-Performance vs No. of Assets

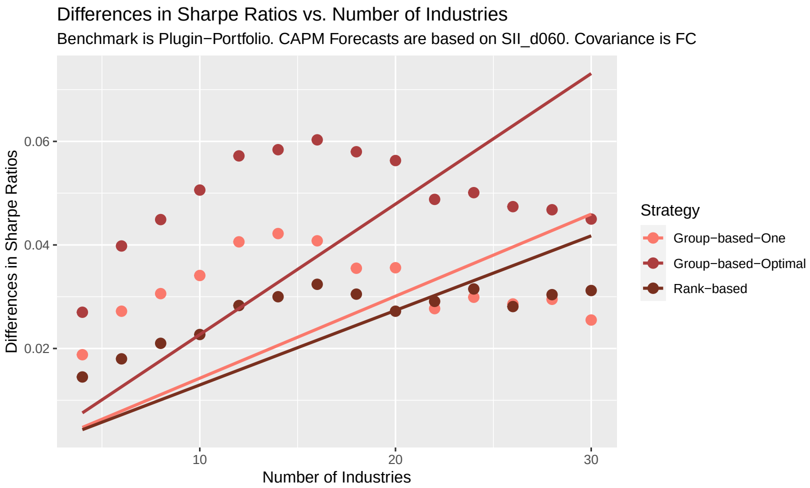

Figure: Benchmark regressions: This figure presents the slope coefficients of plotting Sharpe ratio differences vis-a-vis two bechmark portfolios against the number of assets (100 random draws from FF49 industries, CAPM forecasts).

Possible performance drivers

What are possible performance drivers?

- Are portfolios weights "shrunk" towards equal weights? (cf. Ledoit and Wolf, 2003; 2004; Barroso and Saxena, 2021)?
- Are the estimates based on group-(rank-)based entropy pooling simply better forecasters of future stock returns?

Weight Statistics

Mean absolute deviation against 1/N weights

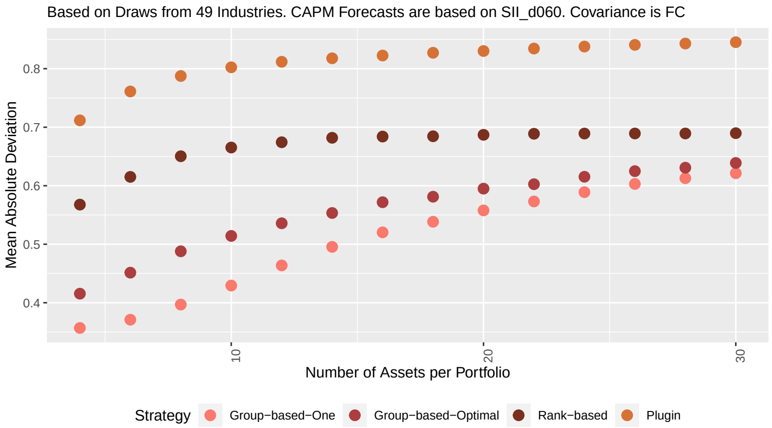

Figure: Mean absolute deviation of optimized industry-sampled portfolio weights relative to equal weights. The number of groups is one and portfolios are based on randomly drawn industries and CAPM forecasts.

Weight Statistics

Mean maximum weight

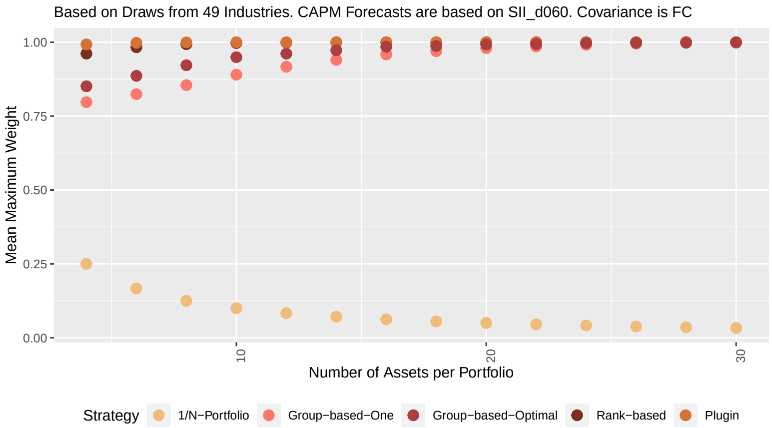

Figure: Mean maximum weight of optimized industry-sampled portfolio weights. The number of groups is one and portfolios are based on randomly drawn industries and CAPM forecasts.

R^2_{OOS} of Strategy Inputs Across All Cross-Sections

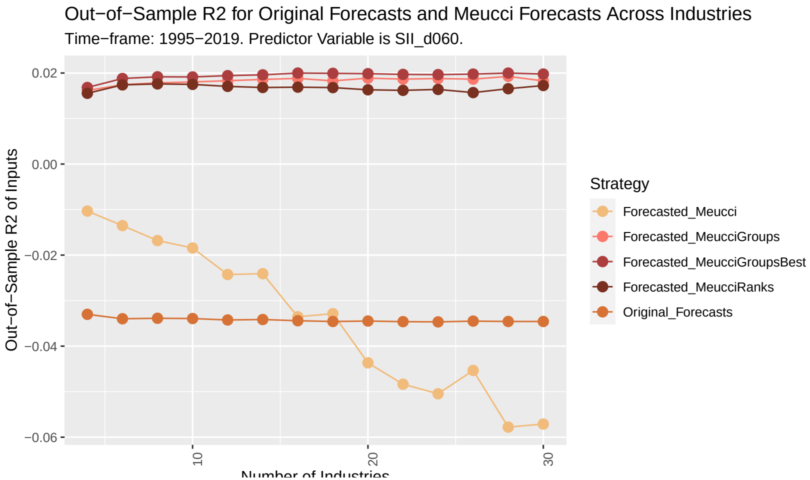

Figure: Aggregate R^2_{OOS} of expected returns used for mean-variance optimization. Benchmark is 60-months rolling average

Cumulative Squared Forecast Error Differences for Cross-Section of 30

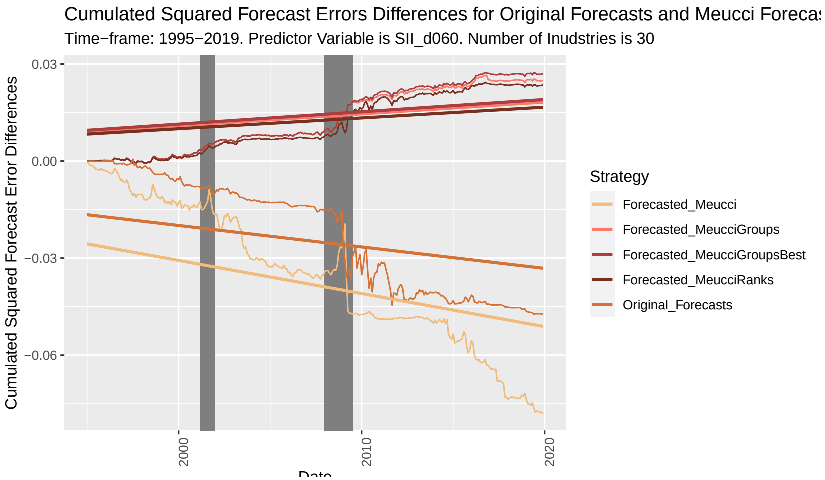

Figure: Cumulative squared forecast error differences between different forecasts and benchmark casts (60-month rolling average).

Realized Standard Deviation of Long-Short Minimum-Variance Portfolios

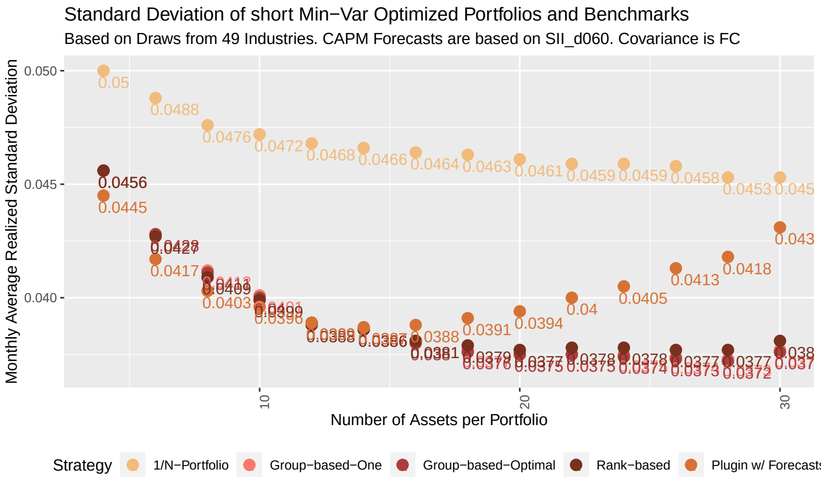

Figure: Realized standard deviation of long-short minimum-variance portfolios.

Robustness

Results are qualitatively similar for...

- Various predictor variables such as:
  - Variance risk premium
  - Financial uncertainty and differences in financial uncertainty
  - Rolling 60-months and 120-months average
  - Machine-learning based forecasts
- Optimized long-only portfolios

Conclusions

- Group-(rank-)based mean-variance optimization increases risk-adjusted performance of optimized portfolios relative to plug-in and 1/N portfolio
- Weights of group-(rank-)based portfolios do not indicate stronger tilt towards 1/N-portfolio then plug-in approach
- Performance gains most likely due to higher accuracy of input parameters (i.e. Group- (rank)based expected returns have positive R^2_{OOS} and covariance estimate results in less realized standard deviation)

Thank you very much for your attention!

Bibliography I

Jorion, P. (1985). International Portfolio Diversification with Estimation Risk. *The Journal of Business* 58 (3), 259–278.

Frost, P. A. and Savarino, J. E. (1988). For better performance: Constrain portfolio weights. *The Journal of Portfolio Management* 15 (1), 29–34.

Michaud, R. O. (1989). The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal* 45 (1), 31–42.

MacKinlay, A. C. and Pástor, Ľ. (2000). Asset Pricing Models: Implications for Expected Returns and Portfolio Selection. *The Review of Financial Studies* 13 (4), 883–916.

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10 (5), 603–621.

Ledoit, O. and Wolf, M. (2004). Honey, I Shrunk the Sample Covariance Matrix. *The Journal of Portfolio Management* 30 (4), 110–119.

Garlappi, L., Uppal, R., and Wang, T. (2007). Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach. *Review of Financial Studies* 20 (1), 41–81.

Kan, R. and Zhou, G. (2007). Optimal Portfolio Choice with Parameter Uncertainty. *The Journal of Financial and Quantitative Analysis* 42 (3), 621–656.

Bibliography II

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies* 22 (5), 1915–1953.

Meucci, A. (2010). Fully Flexible Views: Theory and Practice. *arXiv:1012.2848 [q-fin].*

Allen, D., Lizieri, C., and Satchell, S. (2019). In Defense of Portfolio Optimization: What If We Can Forecast? *Financial Analysts Journal*, 1–19.

Hasler, M. and Martineau, C. (2020). Does the CAPM Predict Returns? *SSRN Electronic Journal.*

Barroso, P. and Saxena, K. (2021). Lest we forget: learn from out-of-sample forecast errors when optimizing portfolios. SSRN Scholarly Paper ID 2771664. Rochester, NY: Social Science Research Network.

Practical Issues With Mean-Variance Optimization

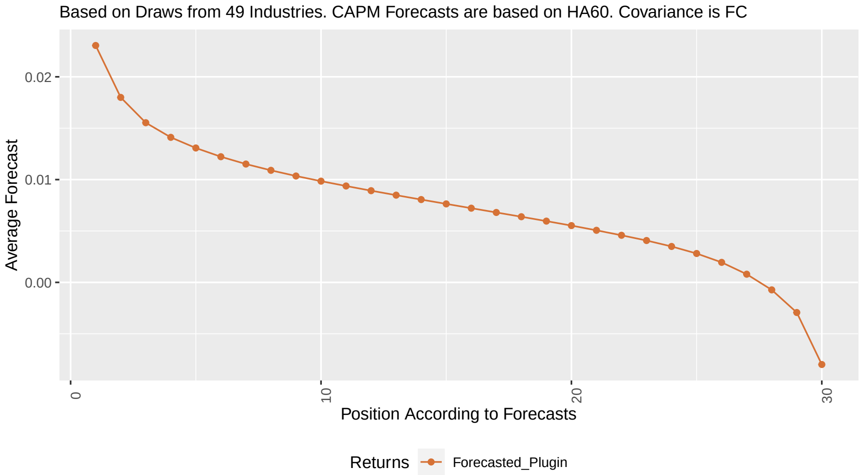

Figure: Monthly expected returns based on 60-month rolling averages and realized returns, optimized portfolio weights and squared forecast errors. Portfolios are based on 100 random draws of 30 assets from 49 FF-industries.

Practical Issues With Mean-Variance Optimization

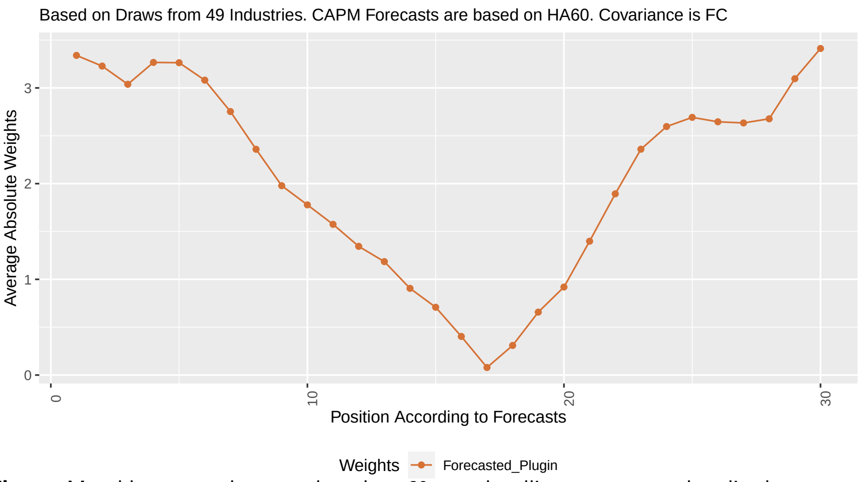

Figure: Monthly expected returns based on 60-month rolling averages and realized returns, optimized portfolio weights and squared forecast errors. Portfolios are based on 100 random draws of 30 assets from 49 FF-industries.

Practical Issues With Mean-Variance Optimization

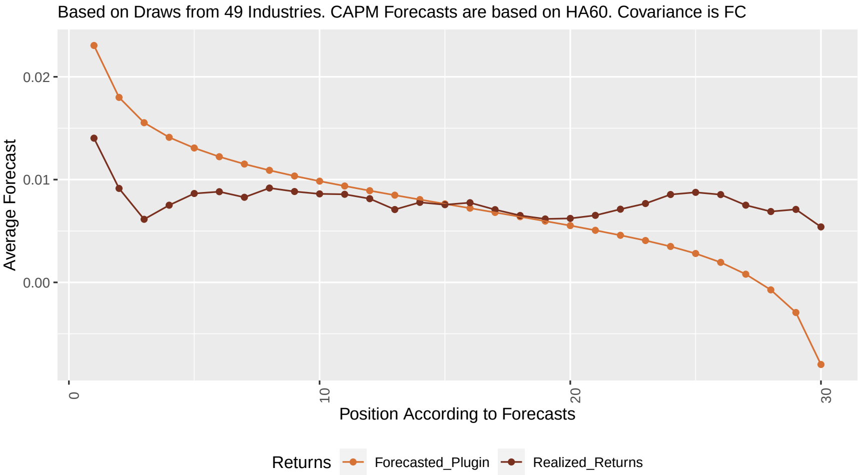

Figure: Monthly expected returns based on 60-month rolling averages and realized returns, optimized portfolio weights and squared forecast errors. Portfolios are based on 100 random draws of 30 assets from 49 FF-industries.

Practical Issues With Mean-Variance Optimization

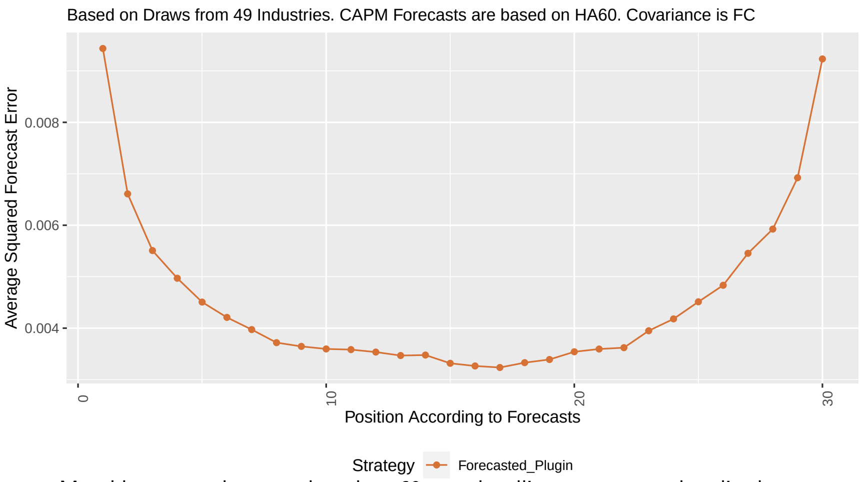

Figure: Monthly expected returns based on 60-month rolling averages and realized returns, optimized portfolio weights and squared forecast errors. Portfolios are based on 100 random draws of 30 assets from 49 FF-industries.

Practical Issues With Mean-Variance Optimization

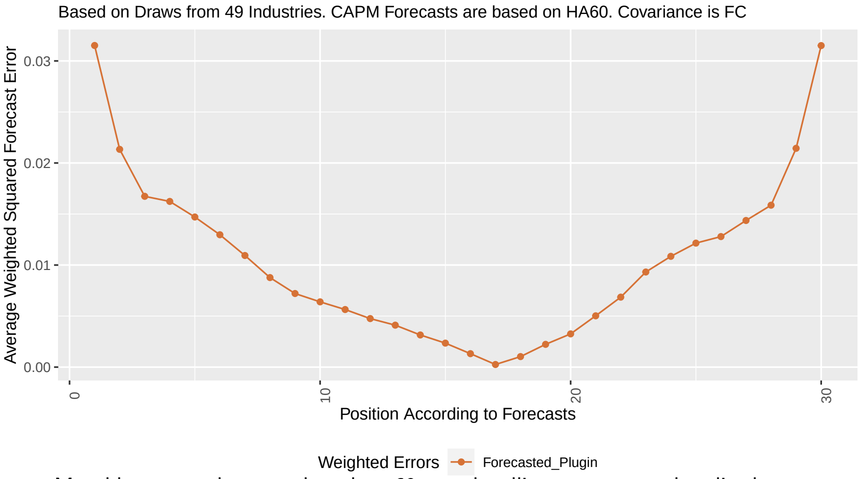

Figure: Monthly expected returns based on 60-month rolling averages and realized returns, optimized portfolio weights and squared forecast errors. Portfolios are based on 100 random draws of 30 assets from 49 FF-industries.

CAPM-Forecasts

- Industry return predictions based on Hasler and Martineau (2020):

$$\mathbb{E}(r_{M,t+1}) = \hat{c}_{1,t} + \hat{c}_{2,t} \cdot SII_t$$

$$\hat{r}_{i,t} \equiv \mathbb{E}_t(r_{i,t+1}) = \hat{\beta}_{i,t} \cdot \mathbb{E}_t(r_{M,t+1}) = \hat{\beta}_{i,t} \cdot (\hat{c}_{1,t} + \hat{c}_{2,t} \cdot SII_t)$$

- where:
  - $\hat{c}_{1,t}$ and $\hat{c}_{2,t}$ based on 60 months rolling window
  - $\hat{\beta}_{i,t}$ based on 24 months rolling window

Back to Data

Entropy-Pooling

- Transform group-(rank-)based views taken from forecasts into mean-variance optimization inputs using entropy-pooling of Meucci (2010):
  - Assume that returns follow a prior distribution $\underline{f}$ and a set of views $\mathcal{V}$
  - Posterior distribution $\bar{f}$ is the one that has the smallest relative entropy w.r.t. $\underline{f}$

$$\bar{f} := \underset{f \in \mathcal{V}}{\operatorname{argmin}}\, \mathcal{E}\left(f | \underline{f}\right) \tag{5}$$

where $\mathcal{E}(f|\underline{f}) := \int f(x) \ln \frac{f(x)}{\underline{f}(x)} dx$

  - For a non-parametric calculation approach we follow Meucci (2010), slightly abusing notation, and, given historical scenarios (observations) with prior distribution $\underline{p} := \left(\frac{1}{H}, \ldots, \frac{1}{H}\right)$ we let $\bar{p}$ be defined by

$$\bar{p} := \underset{p \in \mathcal{V}}{\operatorname{argmin}}\, \mathcal{E}\left(p | \underline{p}\right) \tag{6}$$

where all $\bar{p}_s \geq 0$ and $\sum_{s=t-H+1}^{t} \bar{p}_s = 1$ for all joint historical observations.

Entropy-Pooling

- Assuming, our view is on the ranking of the $\hat{\boldsymbol{\mu}}$:

$$\cdots < \mu_{i1} < \mu_{i2} < \mu_{i3} < \cdots$$

  The posterior distribution will weigh selected historic observations so that all scenarios that do not fulfill the ranking get a smaller weight and the new (weighted) empirical distribution is as close as possible to the prior distribution in terms of relative entropy/Kullback-Leiber-Divergence.
- We then use the moments of the posterior distribution as inputs for mean-variance optimization:

$$\mathrm{w}_{\mathcal{V}} = \frac{\hat{\boldsymbol{\mu}}_{\mathcal{V}} \hat{\Sigma}_{\mathcal{V}}^{-1}}{\hat{\boldsymbol{\mu}}_{\mathcal{V}} \hat{\Sigma}_{\mathcal{V}}^{-1} 1'}, \text{ where} \tag{7}$$

$$\hat{\boldsymbol{\mu}}_{\mathcal{V}} = \sum_{s=t-H+1}^{t} \bar{p}_s \mathrm{r}_{s,\cdot} \text{ and} \tag{8}$$

$$\hat{\Sigma}_{\mathcal{V}} = \sum_{s=t-H+1}^{t} \bar{p}_s \left(\mathrm{r}_{s,t} - \hat{\boldsymbol{\mu}}_{s,t}\right) \left(\mathrm{r}_{s,t} - \hat{\boldsymbol{\mu}}_{s,t}\right)' \tag{9}$$

Performance

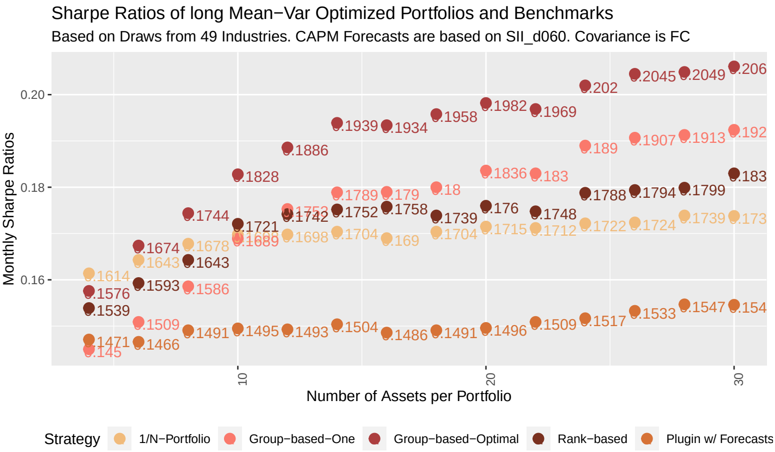

Figure: Monthly Sharpe ratios of rank-based and group-based mean-variance portfolios, plug-in mean-variance portfolios and 1/N portfolios. Portfolios are based on randomly drawn industries and CAPM forecasts.