# The Dispersion Bias

## Correcting a large error in minimum variance portfolios

Alex Shkolnik *(with Lisa Goldberg & Alex Papanicolaou)*
ads2@berkeley.edu

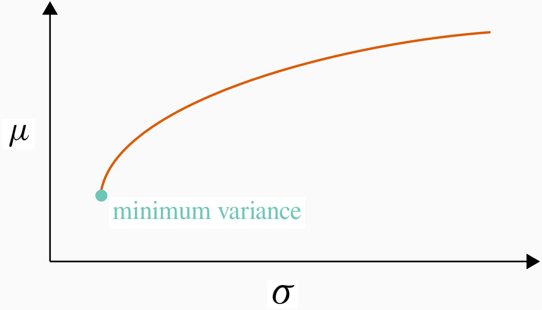
Financial Risks International Forum
March 27, 2018

Center for Risk Management Research
University of California, Berkeley

# Optimized portfolios and the impact of estimation error

# Optimized portfolios

Since Markowitz (1952), quantitative investors have constructed portfolios with mean-variance optimization.




A simple quadratic program given a $\mu$ and a covariance matrix $\mathbf{\Sigma}$.

# The impact of estimation error

In practice, optimization relies on an estimate of the mean and covariance matrix ($\widehat{\boldsymbol{\Sigma}}$ estimates $\boldsymbol{\Sigma}$).

Estimation error leads to two types of errors:

- You get the wrong portfolio: Estimation error distorts portfolio weights so optimized portfolios are never optimal.
- And it's probably risker than you think: A risk-minimizing optimization tends to materially underforecast portfolio risk.

We measure both errors in simulation.

# Measuring the impact of estimation error in simulation

# Measuring errors in weights

(Squared) tracking error of an optimized portfolio $\widehat{w}$ measures its distance from the optimal portfolio $w_*$:

$$\mathscr{T}^2_{\widehat{w}} = (\widehat{w} - w_*)^\top \mathbf{\Sigma} \, (\widehat{w} - w_*)$$

Tracking error is the width of the distribution of return differences between $w$ and $\widehat{w}$.

Ideally, tracking error should be as close to 0.

# Measuring errors in risk forecasts

Variance forecast ratio measures the error in the risk forecast as:

$$\mathscr{R}_{\widehat{w}} = \frac{\widehat{w}^\top \widehat{\mathbf{\Sigma}} \widehat{w}}{\widehat{w}^\top \mathbf{\Sigma} \widehat{w}}$$

This is a ratio of the estimated portfolio risk over the actual risk of the estimated minimum variance portfolio $\widehat{w}$.

Ideally, the variance forecast ratio should be as close to 1.

# Error metrics in simulation

In simulation (given a model for $\boldsymbol{\Sigma}$),

- generate security returns and compute $w_*$ using $\boldsymbol{\Sigma}$,
  ($\boldsymbol{\Sigma}$ is accessible in simulation)
- estimate $\boldsymbol{\Sigma}$ by $\widehat{\boldsymbol{\Sigma}}$ from observed returns and compute $\widehat{w}$,
- measure the error metrics $\mathscr{T}^2_{\widehat{w}}$ and $\mathscr{R}_{\widehat{w}}$.

# Minimum variance

# Why minimum variance?

**Theory**

Error amplification: Highly sensitive to estimation error.

Error isolation: Impervious to errors in expected return.

Insight into a general problem: Informs our understanding of how estimation error distorts portfolios and points to a remedy.

**Practice**

Large investments: For example, the Shares Edge MSCI Min Vol USA ETF had net assets of roughly $14 billion on Sept. 8, 2017.

# True and optimized minimum variance portfolios

The true minimum variance portfolio $w_*$ is the solution to:

$$\min_{x \in \mathbb{R}^N} x^\top \boldsymbol{\Sigma} x$$
$$x^\top 1_{\mathrm{N}} = 1 \,.$$

In practice, we construct an estimated minimum variance portfolio, $\widehat{w}$, that solves the same problem with $\widehat{\boldsymbol{\Sigma}}$ replacing $\boldsymbol{\Sigma}$.

# Factor model & PCA

# One-factor security returns model

The return generating process for $N$ securities is specified by

$$R = \phi \beta + \epsilon$$

where $\phi$ is the return to a market factor, $\beta$ is the $N$-vector of factor exposures, $\epsilon$ is the $N$-vector of diversifiable specific returns.

*The $(\phi, \epsilon)$ are latent variables. We observe $T$ i.i.d. returns to $N$ securities, i.e., $R_1, R_2, \ldots, R_T$.*

# One-factor model covariance matrix

When the $\phi$ and $\epsilon$ are uncorrelated (as we assume), the security covariance matrix can be expressed as

$$\mathbf{\Sigma} = \sigma^2 \beta\beta^\top + \mathbf{\Delta},$$

where $\sigma^2$ is the variance of the market factor and the diagonal entries of $\mathbf{\Delta}$ are specific variances, $\delta^2$.

*Assumption 1.* $\sigma^2/N \to \mu_\infty \in (0, \infty)$ *and* $\mathbf{\Delta} = \delta^2 \mathbf{I}$.

*Assumption 2.* $\{R_i\}_{i=1}^T$ *are i.i.d. with* $R_1 \sim \mathcal{N}(0, \mathbf{\Sigma})$.

*Assumption 3.* $\beta$ *always has some dispersion.*

## Estimation error in factor models

In practice, we have only the estimates $\hat{\sigma}$, $\hat{\beta}$ and $\hat{\delta}$.

$$\widehat{\mathbf{\Sigma}} = \hat{\sigma}^2 \hat{\beta} \hat{\beta}^\top + \widehat{\mathbf{\Delta}}$$

We measure the errors in estimated parameters, of course.

*But our focus is how errors in parameter estimates affect portfolio metrics: (squared) tracking error and variance forecast ratio.*

PCA: compute the sample covariance matrix $\mathbf{S}$ and set:

- $\hat{\beta}$ – first eigenvector of $\mathbf{S}$,
- $\hat{\sigma}^2$ – largest eigenvalue of $\mathbf{S}$,
- $\hat{\delta}^2$ – OLS regression of returns on the estimated factor.

PCA approximates true factors well for $N$ large and $\widehat{\mathbf{\Sigma}} = \mathbf{\Sigma}$.

Sample eigenvectors behave differently for $N$ large and $T$ fixed *(i.e., in the $N \uparrow \infty$ and $T$ fixed asymptotic regime)*.

Current techniques adjust only the eigenvalue $\hat{\sigma}^2$ (typically biased upward). No (direct) corrections of $\hat{\beta}$ are available!

# The dispersion bias

# PCA bias (recent results on sample eigenvectors)

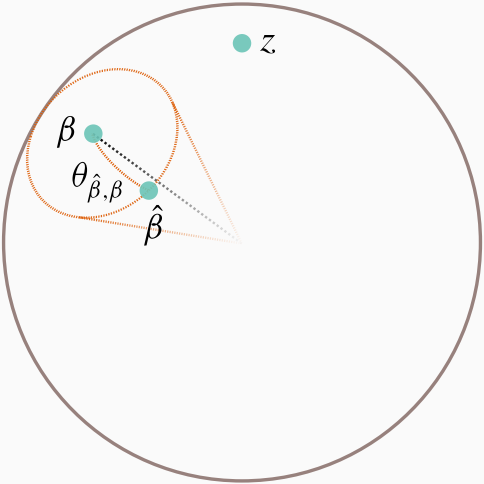
Let $\theta_{\beta,\hat{\beta}}$ be the angle between $\beta$ and its PCA-estimate $\hat{\beta}$.

Recent results (e.g. Shen, Shen, Zhu & Marron (2016)) under our assumptions state that ($N \uparrow \infty$)

$$\cos\theta_{\hat{\beta},\beta} \;\to\; \psi_T^{-1} \tag{1}$$

almost surely for a (non-degenerate) random variable $\psi_T > 1$.

# No reference frame to detect eigenvector bias



# PCA bias characterization

Let $z = 1_N / \sqrt{N}$ (a vector on the unit $N$-sphere).

This is the unique (up to negation) *dispersionless* unit vector.

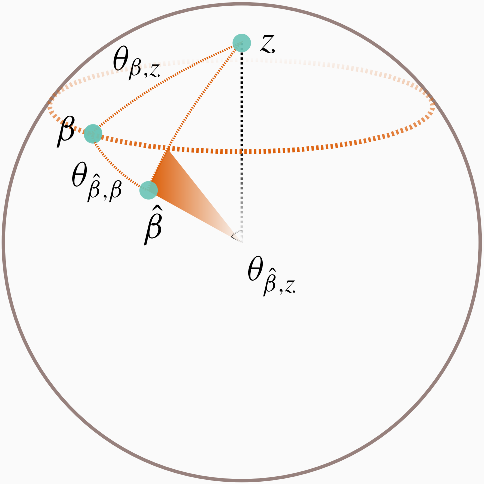
**Theorem**

*Let $\hat{\beta}$ be a PCA-estimate of $\beta$. Then,*

$$\cos \theta_{\beta,z} \overset{a.s.}{\sim} \psi_T \cos \theta_{\hat{\beta},z} \qquad (N \uparrow \infty). \tag{2}$$

In words, $\theta_{\hat{\beta},z}$ is larger than $\theta_{\beta,z}$ with high probability for $N$ large.

# PCA bias illustration



# Error metrics for the minimum variance portfolio

Define $\gamma_{x,y} = x^\top y$ (on unit sphere $\gamma_{x,y} = \cos\theta_{x,y}$) and

$$\mathscr{E}_x = \frac{\gamma_{\beta,z} - \gamma_{\beta,\hat{\beta}}\gamma_{\hat{\beta},z}}{\sin\theta_{\hat{\beta},z}}. \tag{3}$$

*The variable $\mathscr{E}$ drives all the error in our metrics.*

*We prove, $\mathscr{E}_{\hat{\beta}} > 0$ for the PCA-estimate $\hat{\beta}$.*

# Error metrics for the minimum variance portfolio

As $N \uparrow \infty$, (for any $\hat{\beta}$ such that $\inf_N \mathscr{E}^2_{\hat{\beta}} > 0$)

$$\mathscr{T}^2_{\widehat{w}} \sim \frac{\mu^2 \mathscr{E}^2_{\hat{\beta}}}{\sin^2 \theta_{\hat{\beta},z}} \qquad \mathscr{R}_{\widehat{w}} \sim \frac{\hat{\delta}^2 N^{-1}}{\mu^2 \mathscr{E}^2_{\hat{\beta}}}. \tag{4}$$

**Remarkable:** *No dependence on the eigenvalue estimate $\hat{\sigma}^2$!*

(above, $\mu^2 = \sigma^2/N$)

# Bias correction

# Dispersion bias correction

We propose a correction $\hat{\beta}^*$ of the form,

$$\hat{\beta}^* \propto \hat{\beta} + \rho z \qquad \rho \in \mathbb{R}\,.$$

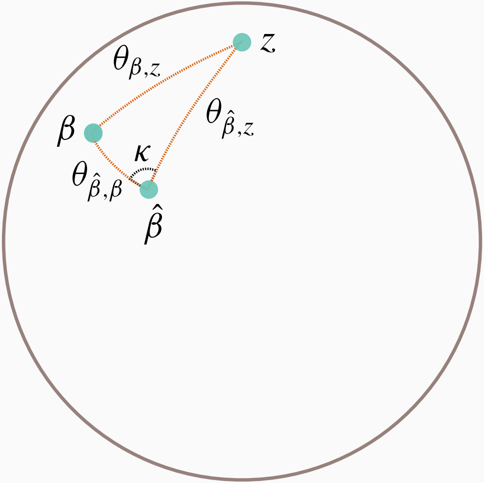
We consider two estimators (i.e., values of $\rho$)

$$\rho_1 = \frac{\gamma_{\beta,z} - \gamma_{\hat{\beta},z}\gamma_{\hat{\beta},\beta}}{\gamma_{\hat{\beta},\beta} - \gamma_{\hat{\beta},z}\gamma_{\beta,z}} \qquad \text{(oracle, } \mathscr{E}_{\hat{\beta}^*} = 0)\,. \tag{5}$$

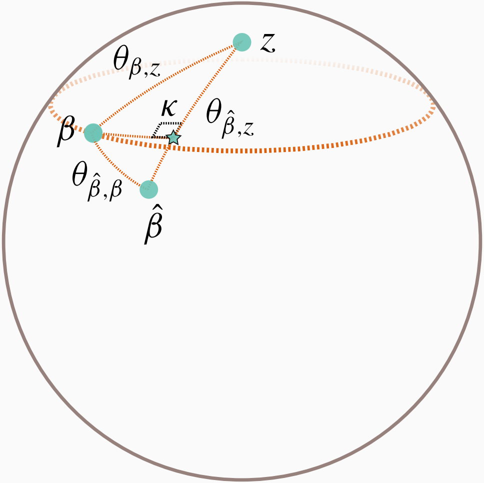
$$\rho_2 = \frac{q\gamma_{\hat{\beta},z}}{1 - (q\gamma_{\hat{\beta},z})^2}(q - q^{-1}) \qquad \text{(data-driven, } \mathscr{E}_{\hat{\beta}^*} \approx 0)\,. \tag{6}$$

where $q$ is computed from observed data only.

# Dispersion bias correction



# Dispersion bias correction



# Main result

**Theorem**

*Under our assumptions, the oracle estimator achieves*

$$\mathscr{R}_{\widehat{w}} \overset{a.s.}{\sim} \hat{\delta}^2/\delta^2 \qquad \mathscr{T}_{\widehat{w}}^2 \overset{a.s.}{\sim} O(N^{-1}) \,. \tag{7}$$

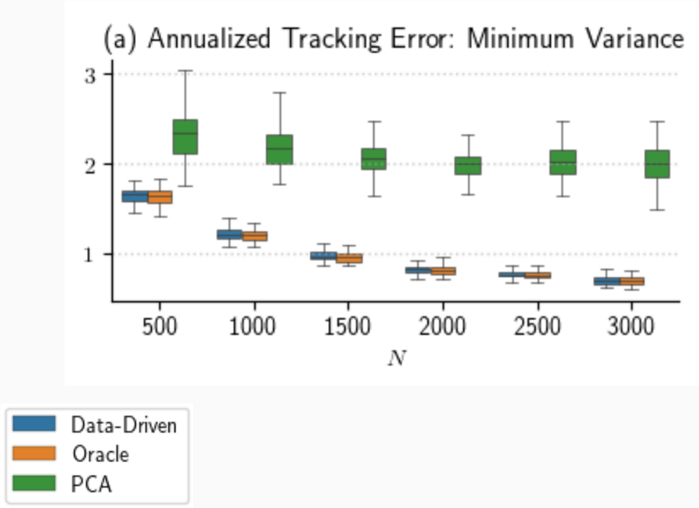
Several complementary results available in paper (online).

# Numerical results

# Calibrating the one-factor model

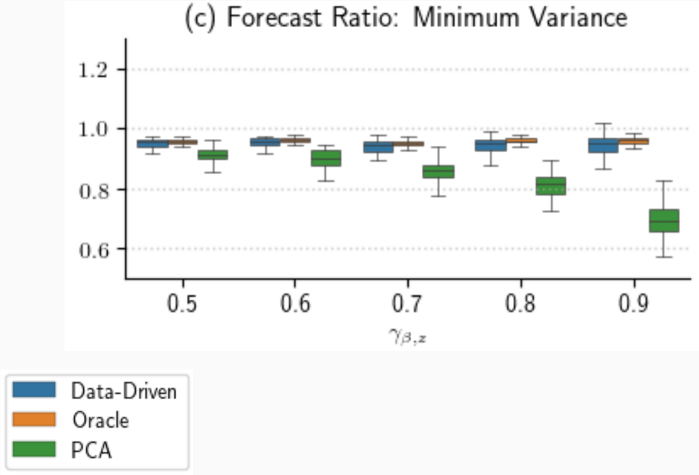
| Parameter | Value | Comment |
|---|---|---|
| $\gamma_{\beta,z}$ | 0.5–1.0 | controls dominant factor dispersion |
| $\sigma^2$ | (roughly) the dominant eigenvalue of $\boldsymbol{\Sigma}$ | annualized factor volatility is 16% |
| $\delta^2$ | specific variances on diagonal of $\boldsymbol{\Delta}$ | annualized specific volatilities drawn uniformly on [10%, 64%] |

# Numerical results ($\gamma_\beta = 0.90$ and varying $N$)



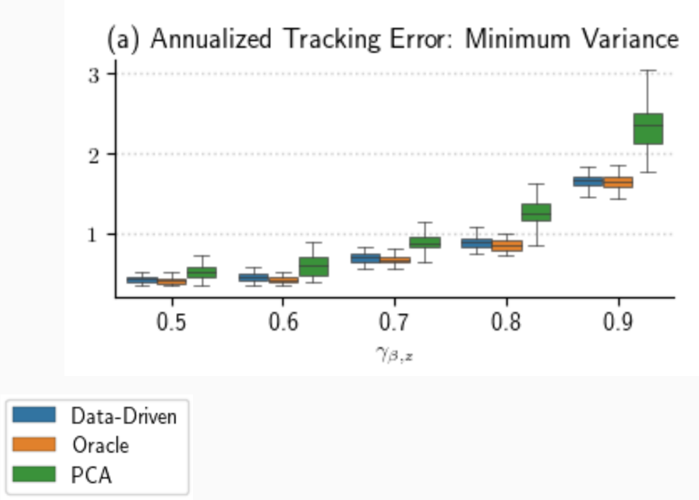

Simulation based on 50 samples

# Numerical results ($\gamma_{\beta,z} = 0.90$ and varying $N$)



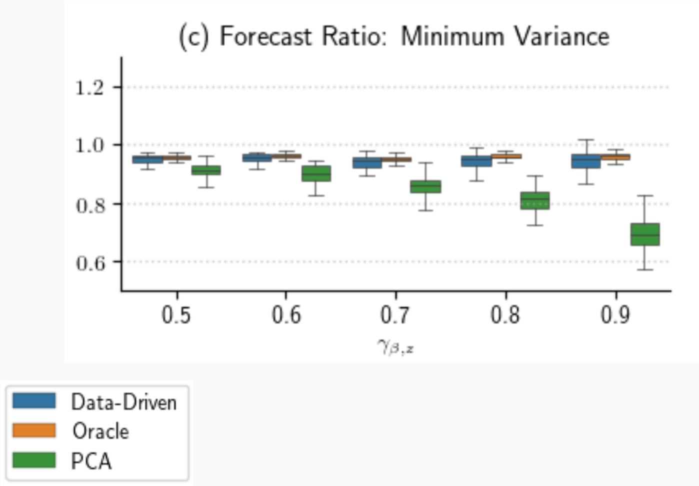

Simulation based on 50 samples

# Numerical results ($N = 500$ and varying $\gamma_{\beta,z}$)




Simulation based on 50 samples

# Numerical results ($N = 500$ and varying $\gamma_{\beta,z}$)




Simulation based on 50 samples

## Summary

We provided a novel characterization of a systematic (dispersion) bias in PCA factors (sample eigenvectors).

– *applicable in many other settings.*

Developed and tested oracle and data-driven corrections to mitigate this dispersion bias (distinct from literature).

Our results can be viewed as an extension and formalization of ideas that have been known by practitioners since the 1970s.

Markowitz, H. (1952), 'Portfolio selection', *The Journal of Finance* **7**(1), 77–91.

Rosenberg, B. (1984), 'Prediction of common stock investment risk', *The Journal of Portfolio Management* **11**(1), 44–53.

Rosenberg, B. (1985), 'Prediction of common stock betas', *The Journal of Portfolio Management* **11**(2), 5–14.

Ross, S. A. (1976), 'The arbitrage theory of capital asset pricing', *Journal of economic theory* **13**(3), 341–360.

Sharpe, W. F. (1964), 'Capital asset prices: A theory of market equilibrium under conditions of risk', *The Journal of Finance* **19**(3), 425–442.

Shen, D., Shen, H., Zhu, H. & Marron, S. (2016), 'The statistics and mathematics of high dimensional low sample size asympotics', *Statistica Sinica* **26**(4), 1747–1770.

Treynor, J. L. (1962), Toward a theory of market value of risky assets. Presented to the MIT Finance Faculty Seminar.

# Some past approaches

Plug-in estimates $\widehat{\boldsymbol{\Sigma}}$:

- sample covariance matrix
  - *underforecasts risk by factor* $(1 - N/T)_+$
- covariance regularization
- low-dimensional approximation (factor model)
- bayes/shrinkages estimates (structured model)

Also, bootstrap resampling and stochastic optimization.

Factor models form our starting point.

# Factor models and equity markets

Beginning with the development of the Capital Asset Pricing Model (CAPM) in (Treynor 1962) and (Sharpe 1964), factor models have been central to the analysis of equity markets.

In a fundamental model, human analysts identify factors. Fundamental models have been widely used by equity portfolio managers since (Rosenberg 1984) and (Rosenberg 1985).

In a statistical model (such as PCA, factor analysis, etc), machines identify factors. An enormous academic literature on PCA models has descended from (Ross 1976).

*PCA is the focus of our analysis.*