MIT Sloan School of Management

MIT Sloan School Working Paper 5174-16

Advances in Factor Replication

Megan Czasonis, Mark Kritzman, and David Turkington

July 31, 2016

ADVANCES IN FACTOR REPLICATION

Megan Czasonis, Mark Kritzman, and David Turkington

THIS VERSION: AUGUST 12, 2016

Megan Czasonis is a Vice President at State Street Associates in Cambridge, MA
mczasonis@statestreet.com

Mark Kritzman is the CEO of Windham Capital Management in Boston, MA and Senior Lecturer at MIT Sloan School in Cambridge, MA
kritzman@mit.edu

David Turkington is a Senior Vice President at State Street Associates in Cambridge, MA
dturkington@statestreet.com

Abstract

Factor investing has gained widespread acceptance among institutional investors. Some investors believe it is preferable to stratify the investment universe into factors to manage portfolio risk more effectively, while other investors focus on factors because they believe they yield risk premiums. Factors such as economic variables are not directly investable. Investors, therefore, need to identify a combination of securities that tracks the movements in the economic variable. Other factors, however, are directly investable, such as securities with a certain attribute. Often times, though, investors choose to invest in a subset of the factor securities that are inexpensive to trade. In order to identify the best factor-tracking portfolio, investors must estimate covariances from historical observations whose realizations in the future are prone to several types of estimation error. We introduce a non-parametric procedure to account for estimation error, which enables us to incorporate the relative stability of covariances directly into the factor replication process. We show that adjusting for the stability of covariances in this way produces replicating portfolios that are significantly more reliable than portfolios that are either blind to estimation error, formed using Bayesian shrinkage, or by implication, formed by resampling.

We thank Stephen Brown, Sanjiv Das, Paul Koehler, and Harry Markowitz for helpful comments.

ADVANCES IN FACTOR REPLICATION

Institutional investors are increasingly focusing their attention on factor investing. Some investors argue that reconfiguring portfolios into factors allows for more effective risk management, while other investors believe that factors offer risk premiums. Some factors are not directly investable, such as economic variables. Investors must, therefore, identify a combination of securities that tracks the movement of the factor. Many factors, though, are directly investable, such as securities with a particular attribute, but investors often choose to invest in a subset of factor securities that are inexpensive to trade. In each situation, investors must estimate covariances from historical observations, which exposes them to several types of estimation error. We introduce a non-parametric procedure to account for the various types of estimation error, thus enabling us to incorporate the relative stability of covariances directly into the factor replication process.[1] We show that adjusting for estimation error in this way yields replicating portfolios that are significantly more reliable than portfolios that are either blind to estimation error, formed using Bayesian shrinkage or, by implication, formed by resampling.[2]

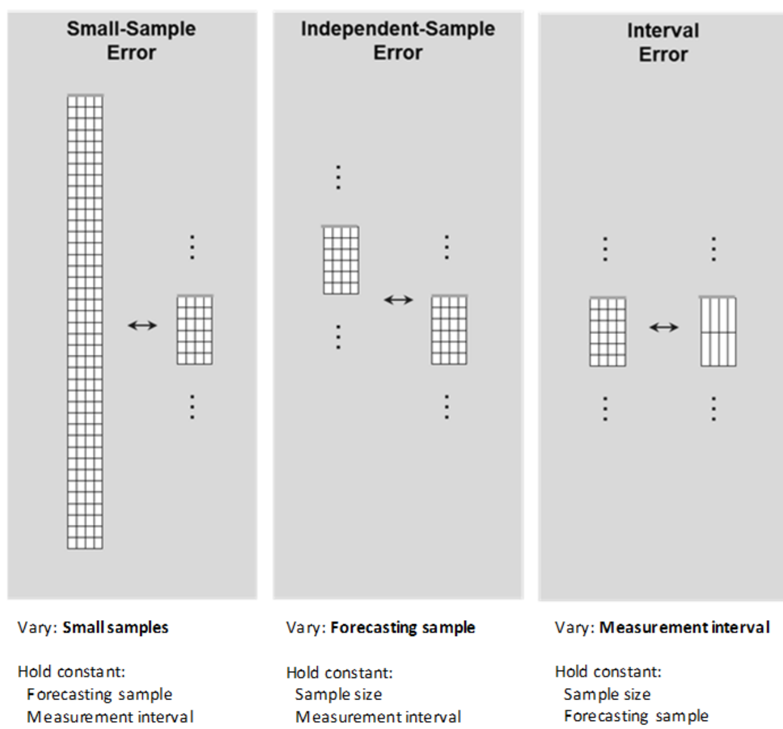
Types of Estimation Error

Investors often rely on long samples of high-frequency[3] historical data to identify factor-replicating portfolios intended to deliver results over a shorter, future period. This reliance on high-frequency historical data exposes investors to three types of estimation error.

First, small-sample error arises when covariances from a long sample are used to forecast covariances of a specific smaller sample. Even though the true values of a long sample are known, the realizations of those values in shorter sub-samples are likely to be meaningfully different. Second, independent-sample error arises when known covariances from one sample are projected onto a separate, independent sample. Third, interval error arises when covariances of high-frequency observations, such as weekly observations, differ from covariances of longer-period observations, such as annual observations. This distortion occurs if autocorrelations or lagged cross correlations of the higher-frequency observations differ from zero.[4]

Exhibit 1 illustrates these three types of estimation error.

Exhibit 1: Sources of Estimation Error



Investors typically attempt to mitigate the effect of estimation error either by Bayesian shrinkage, in which each observation is blended with a prior belief such as the cross-sectional average of all the observations, or by resampling, in which many replicating portfolios are generated from a distribution of means and covariances, and then averaged to arrive at the error-adjusted optimal portfolio.[5] Although Bayesian shrinkage is potentially beneficial, resampling will be ineffective if the weights of the replicating portfolio include unconstrained long and short positions. It will yield the same replicating portfolio as one would derive by ignoring errors.[6] We discuss this issue more thoroughly in Appendix A.

We propose that investors treat estimation error as a distinct component of risk. Unfortunately, it is not possible to do so analytically. We, therefore, introduce a non-parametric procedure for adjusting a sample of returns to reflect the relative stability of the assets' covariances. This approach yields portfolios that are substantially different and substantially more stable than those that either ignore estimation error, rely on Bayesian shrinkage, or are formed by resampling.

A Non-Parametric Procedure for Constructing a Stability-Adjusted Return Sample

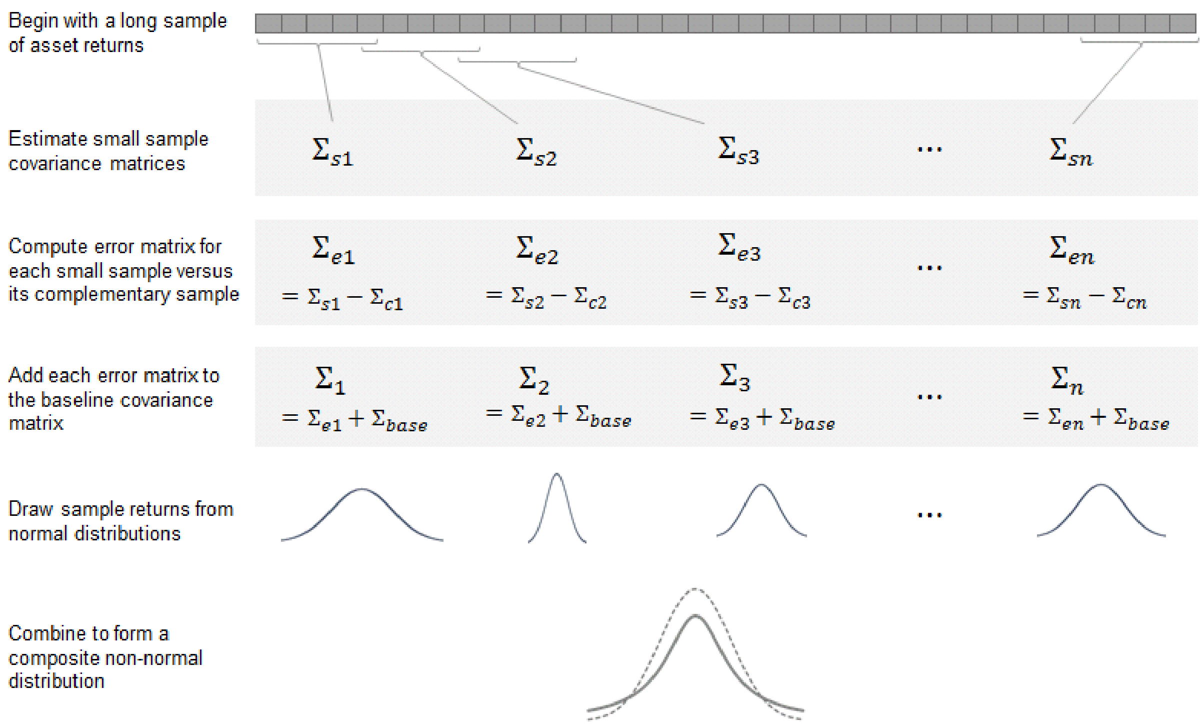
We now describe how to construct a return sample that accounts for small-sample error, independent-sample error, and interval error.

1. We begin by selecting a large sample of returns.

2. We then select all possible overlapping sub-samples of a chosen size and compute their covariance matrices based on observations of the same interval as our replicating horizon.[7]
3. We then subtract the sub-sample covariances from the covariances estimated from the remaining observations in the large sample, which we call the complementary sample.[8] [9]
4. Next, for all sub-samples, we add the errors to the covariances of a base-case sample, which, for example, could be the original sample. It is the sample investors would default to if they were to ignore estimation error. [10]
5. Then, assuming normality of log returns, we generate 1,000 simulated return samples from each error-adjusted covariance matrix. We assume the same mean returns for all sub-samples; only the covariance matrices change with the sub-samples.
6. Finally, we combine these return samples into a new, large sample, which can be thought of as a stability-adjusted return sample.

This process is depicted in Exhibit 2.

Exhibit 2: Constructing a Stability-Adjusted Return Sample



We should note several features of this process. First, the composite errors incorporate all three sources of error. They reflect small-sample error because the sub-samples are smaller than the original sample. They reflect independent-sample error, because each sub-sample is distinct from the remaining observations in the large sample. And they capture interval error, because the sub-sample covariances are estimated from longer-interval observations than the observations used to estimate the large-sample covariances.

We should also note that the resultant return distribution will not be normal even though the distributions of the sub-samples are normally distributed. The combined distribution will have fatter tails than a normal distribution, because the sub-sample distributions have different variances. It is also possible that the combined distribution, though likely symmetric, will be non-elliptical depending upon how the variances and correlations differ across the sub-samples. We provide an intuitive explanation of this possibility in Appendix B.

Finally, we should note that the distribution of the stability-adjusted return sample is not the expected distribution of a large sample of future returns. Rather, it is a distribution of past returns reshaped by historical estimation error to better mirror the distribution of a future small sample of longer-interval returns.

Replicating Factors from Stability-Adjusted Return Samples

As we just described, the process of combining many return distributions, which themselves are normal, results in a composite distribution which is not normal and possibly non-elliptical. This poses a potential challenge to factor replication. The conventional approach for replicating factors is to employ mean-variance optimization to minimize tracking error between movements in the factor portfolio or factor value and the replicating securities.[11] Mean-variance optimization assumes that returns are elliptically distributed, of which the normal distribution is a special case, or that investors have preferences that can be well approximated by mean and variance. Although most power utility functions, such as the log-wealth utility function, can be reasonably approximated by mean and variance, utility functions that have kinks or inflection points cannot. Thus, in most cases we can simply estimate a covariance matrix from the stability-adjusted return sample and replicate the factor using mean-variance optimization.[12]

The alternative approach to replicating factors is called full-scale optimization, but this approach is computationally challenging. It is limited to situations in which the combination of the number of assets, the sample size, and the granularity of the search is computationally manageable. It is a direct, numerical utility maximization process, as described below.

1. We select a particular utility function, which need not be amenable to approximation by mean and variance.
2. We choose a particular combination of factor replicating securities and apply it every period to the returns in the stability-adjusted return sample to compute the utility associated with that factor-replicating portfolio for every period.
3. We sum utility across all periods and record this value.
4. We choose a different set of replicating securities and again compute its total utility across all periods.
5. We proceed as many times as necessary to arrive at the replicating portfolio that yields the highest utility across all periods.

This full-scale approach to optimization accounts for every feature of the data, even beyond kurtosis and skewness. It is thus suitable for special cases in which the distribution is non-elliptical and investor utility cannot be described by mean and variance.

It is important to understand how our approach differs from resampling. Resampling generates many portfolios from a distribution of means and covariances, and averages these portfolios to produce a resampled efficient frontier. In instances in which we apply mean-variance optimization, we form a single replicating portfolio from a single covariance matrix derived from a single, composite distribution of many error-adjusted sub-samples. This is the distribution of the stability-adjusted return sample. In the case of full-scale optimization, we derive a single replicating portfolio from the stability-adjusted return sample.

This distinction between resampling and stability-adjusted replication is critical. By constructing a composite return distribution of many error-adjusted sub-samples, we capture the effect of estimation error on the shape of the return distribution, thus enabling us to build replicating portfolios that are sensitive to the reshaped return distribution arising from estimation error. Resampling ignores the effect of estimation error on the shape of the return distribution, because it is based on an average rather than a union of distributions.

Results

We first measure the efficacy of stability-adjusted factor replication using mean-variance optimization. We allow the portfolio to include both long and short positions without constraint.[13] Because our goal is to minimize tracking error, we set the means equal to zero.[14] We form three replicating portfolios: one derived from covariances computed from the stability-adjusted return sample, one derived from covariances computed from the original return sample, and one derived from covariances computed from the original return sample, but blended equally with a matrix composed of the cross-sectional average asset volatility and average off-diagonal asset correlation.

As our figures of merit,[15] we measure the dispersion in tracking error between the 90th and 10th percentile of all possible sub-samples. We regard replicating portfolios with small differences in tracking error as more stable than portfolios with large differences. We also consider the variability of tracking error through time.

We recognize that these sub-samples come from the same original sample used to form the replicating portfolios. But the sub-samples that we use to measure dispersion in tracking error are each independent of the data used to form the replicating portfolios with the exception of the small fraction of the original sample that overlaps with each sub-sample. Thus, our measure of effectiveness is not a back test, but rather a test of robustness. We implicitly assume that the range of covariances across the sub-samples in our data reasonably captures the likely range of covariances in the future. We see no way around this, because in order to capture small-sample error we need covariances from a large sample with which to contrast the small-sample covariances. If we were to generate strictly out-of-sample, non-overlapping observations, we could only generate very few of them. We therefore believe that many observations, each of which only slightly overlaps with the original sample, is a superior measure of robustness than a much smaller number of purely out-of-sample observations.

One might argue that our test of robustness fails to account for regime shifts, but we do not see it that way. We think of the differences in covariances across sub-samples as evidence of shifting regimes. Dramatic regime shifts, such as the global financial crisis, will significantly affect the relative stability of covariances and thus the optimal replicating weights. If, by contrast, the regime shifts are minor, they will have a smaller impact on the stability-adjusted optimal weights. Covariance matrices that are computed without regard to estimation error or adjusted by Bayesian shrinkage also contain information about regime shifts that occurred in the historical sample, but they only incorporate this information by averaging it into the overall covariance matrix.

We replicate three non-investable factors: industrial production, inflation, and energy using 24 industry groups, as well as three investable factors: size, value and momentum, using a subset of 30 cheaper-to-trade securities.

We also illustrate full-scale optimization by replicating inflation with five asset classes for an investor with a kinked utility function. In this case we measure the dispersion of tracking error as well as the dispersion of downside tracking error as figures of merit. We measure the downside as returns below a kink of 0.0%.

We first present results for mean-variance optimization.

Mean-variance optimization results

Exhibit 3 summarizes the process for replicating the non-investable factors:

Exhibit 3: Non-Investable Factor Replication Process

Factors:
- Industrial production (percentage change in seasonally adjusted Industrial Production Index from St. Louis Fed's FRED website)
- Inflation (percentage change in seasonally adjusted Consumer Price Index from BLS website)
- Energy (percentage change in oil price per barrel from Bloomberg; ticker OILPHIST Index)

Replicating assets:
- 24 US industry groups (MSCI GICS level II Total Return Indices)[16]

Sample:
- Monthly returns, July 1996 through May 2016

Portfolio formation details:
- Small sample window: 5 years
- Horizon for longer interval covariances: 1 year
- Baseline covariance matrix: full sample
- Expected returns: zero

Exhibit 4 shows the tracking error through time of the three replicating portfolios for each factor, along with the dispersion in tracking error between the 10^{th} and 90^{th} percentile sub-samples.

Exhibit 4: Trailing Five-Year Tracking Error and Dispersion across Sub-Samples (Non-Investable Factors)

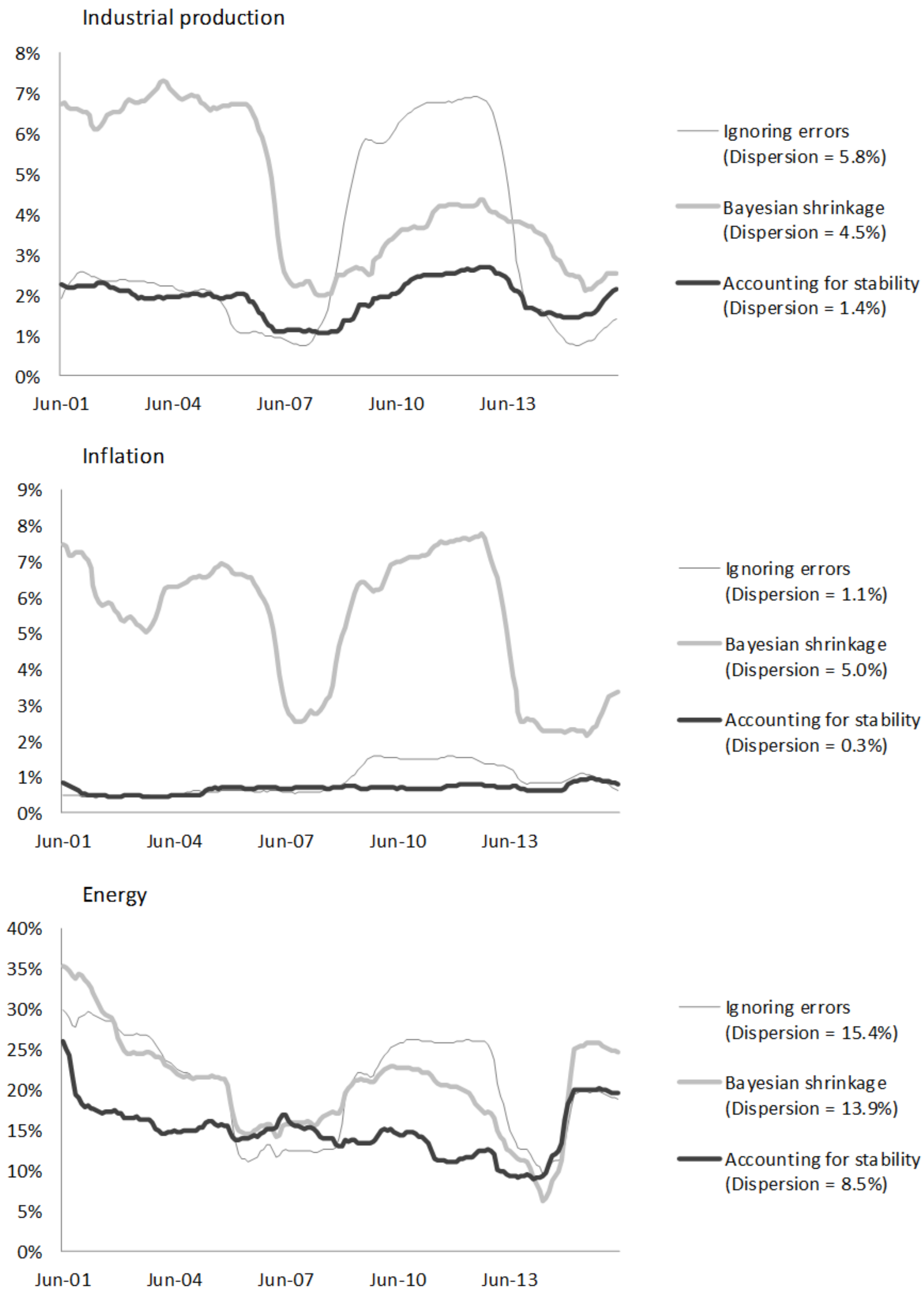


Exhibit 4 starkly reveals that factor-replicating portfolios that explicitly account for the relative stability of covariances are far more stable through time and across a wide range of sub-samples than factor-replicating portfolios that either ignore estimation error, rely on Bayesian shrinkage, or by implication, are formed by resampling.[17] And, in particular, the greatest improvement occurred during the dot com bubble and global financial crisis, when risk was unusually elevated.

Exhibit 5 summarizes the process we use to replicate investable factor universes with a subset of hypothetically cheaper-to-trade securities.[18]

Exhibit 5: Investable Factor Replication Process

Stock universe:

- 515 U.S. stocks (based on MSCI constituents as of December 31, 2015)[19]

Sample:

- Weekly returns, January 2006 through June 2016

Factor portfolio and replicating assets:

For each attribute (size, value, momentum):

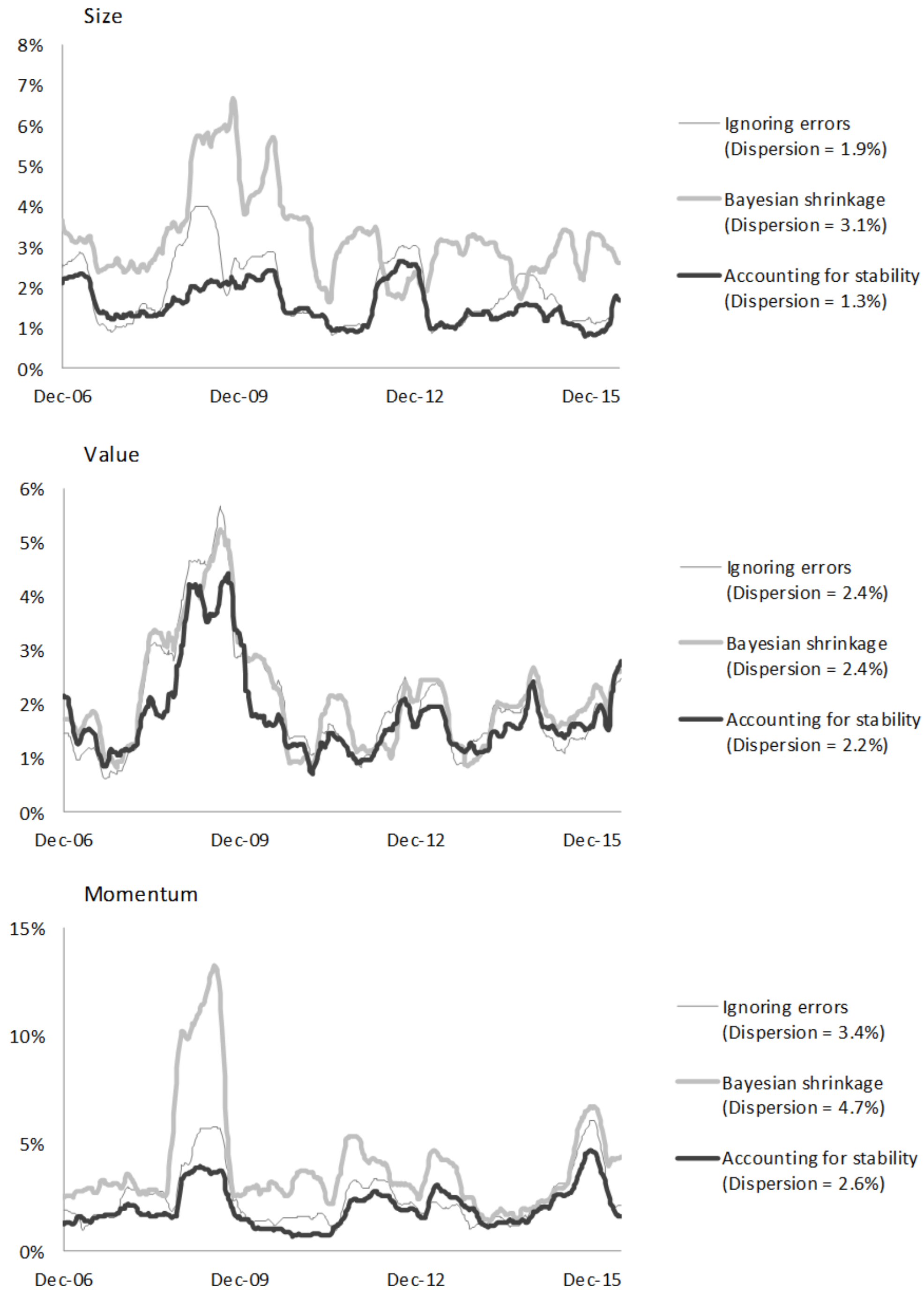
1. Rank stocks on relevant attribute (market cap, price-to-book ratio, or lagged 12-month returns) as of June 24, 2016
2. Isolate top 20% and bottom 20% of stocks (103 stocks per top/bottom bucket)
3. Define benchmark portfolio as the historical return of the top 20% bucket minus the bottom 20% bucket
4. Within each top and bottom bucket, randomly select 15 stocks
5. Define tracking universe as the subset of 30 stocks (“cheaper-to-trade” subset)

Portfolio formation details:

- Small sample window: 1 year (52 weeks)
- Horizon for low frequency covariances: 1 quarter (13 weeks)
- Baseline covariance matrix: full sample
- Expected returns: zero

Exhibit 6 shows the tracking error through time of the three replicating portfolios for each factor, along with the dispersion in tracking error between the 10th and 90th percentile sub-samples.

Exhibit 6: Trailing Five-Year Tracking Error and Dispersion across Sub-Samples (Investable Factors)



The results in Exhibit 6 pertain to just one randomly selected subset of replicating stocks that are assumed to be cheaper to trade. It could be the case that the superiority of accounting for stability is specific to this particular subset of stocks. Exhibit 7 belies this concern. It shows the frequency of trials in which each approach produced the most or least stable results, as well as average instability, which is the average across all 100 trials of the 90th to 10th percentile dispersion in tracking error for each trial. Clearly, accounting for stability yields more stable results far more reliably than ignoring errors, relying on Bayesian shrinkage, or using resampling.

Exhibit 7: Reliability of Portfolio Formation Processes

Among 100 trials, number of times each method was the most or least stable			
	Ignoring errors	Bayesian Shrinkage	Acounting for Stability
Size			
Most stable	6	0	94
Least stable	7	93	0
Average instability	1.73%	2.52%	1.14%
Value			
Most stable	14	12	74
Least stable	39	56	5
Average instability	2.28%	2.47%	1.76%
Momentum			
Most stable	8	0	92
Least stable	5	95	0
Average instability	3.07%	4.42%	2.27%

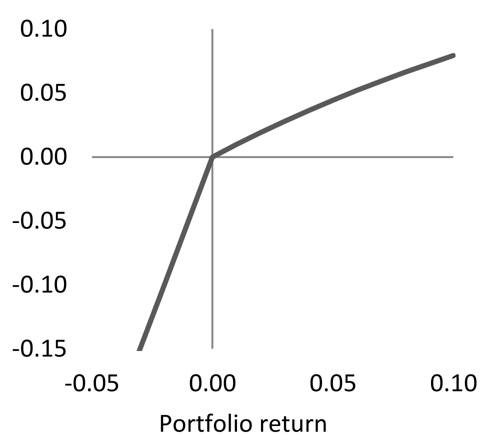
Full-scale optimization

Now we turn to factor replication using full-scale optimization. As we mentioned earlier, full-scale optimization directly maximizes expected utility given a particular sample of returns and a particular utility function. It may be more reliable than mean-variance optimization in cases in which the return distribution is non-elliptical and investors have preferences that cannot be reasonably described by mean and variance. However, it is only feasible if the combination of the number of assets, sample size, and granularity of the search is computationally manageable.

To illustrate the efficacy of full-scale optimization, we assume that the investor has a kinked utility function as shown in Exhibit 8, where *r* is the portfolio return, *k* is the location of the kink in utility, ω is the slope of utility below the kink, and θ is the curvature parameter for power utility above the kink.

Exhibit 8: Kinked Utility Functions

$$U(r) = \begin{cases} \dfrac{1}{1-\theta}\left((1+r)^{1-\theta}-1\right), & if\ r \geq k \\ \dfrac{1}{1-\theta}\left((1+k)^{1-\theta}-1\right)-\omega(k-r), & if\ r < k \end{cases}$$



Investors who face thresholds such as funding requirements have preferences that are better described by kinked utility functions than power utility functions.

We assume a -100% fixed exposure to an "inflation asset" whose returns are the percentage changes in the CPI Index. We allow both long and short positions in the replicating assets, but we force their weights to sum to zero, in order to facilitate tractability. We do not assume the expected returns of the assets equal zero. Because kinked utility implies sensitivity to a particular threshold, we must account for expected return. But our focus is on risk control; thus we set the expected returns equal to each other.

Exhibit 9 summarizes how we deploy full-scale optimization to replicate inflation.

Exhibit 9: Full-Scale Replication Process
(Inflation)

Factor:
- Inflation (percentage change in seasonally adjusted Consumer Price Index from BLS website)

Replicating assets:
- US Stocks (S&P 500 Total Return Index)
- US Treasuries (Barclays US Treasury: Long Total Return Index)
- US Credit (Barclays US Credit Total Return Index)
- Commodities (S&P GSCI Commodity Total Return Index)
- Real Estate (FTSE EPRA/NAREIT United States Total Return Index)

Sample:
- Monthly returns, January 1990 through May 2016

Portfolio formation details:
- Small-sample window: 5 years
- Horizon for longer interval covariances: 1 year
- Constraints: sum of weights must equal zero
- Expected returns: all equal to each other
- Search granularity: 1%
- Utility function: kink at 0%, slope = 5 below kink, power utility with curvature = 5 above kink

Exhibit 10 shows tracking error through time as well as the difference in tracking error between the 10th and 90th percentile sub-samples, while Exhibit 11 shows the same information for downside tracking error below a return of 0%.

Exhibit 10: Trailing Five-Year Tracking Error and Dispersion across Sub-Samples (Inflation)

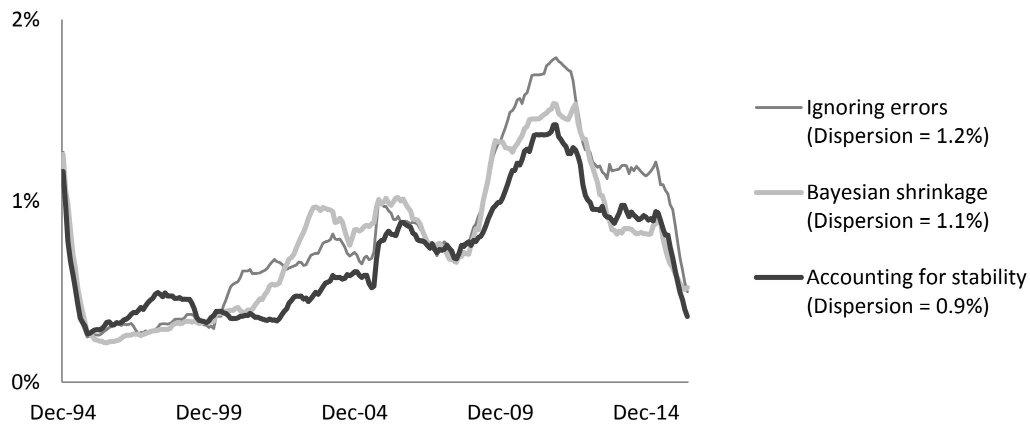
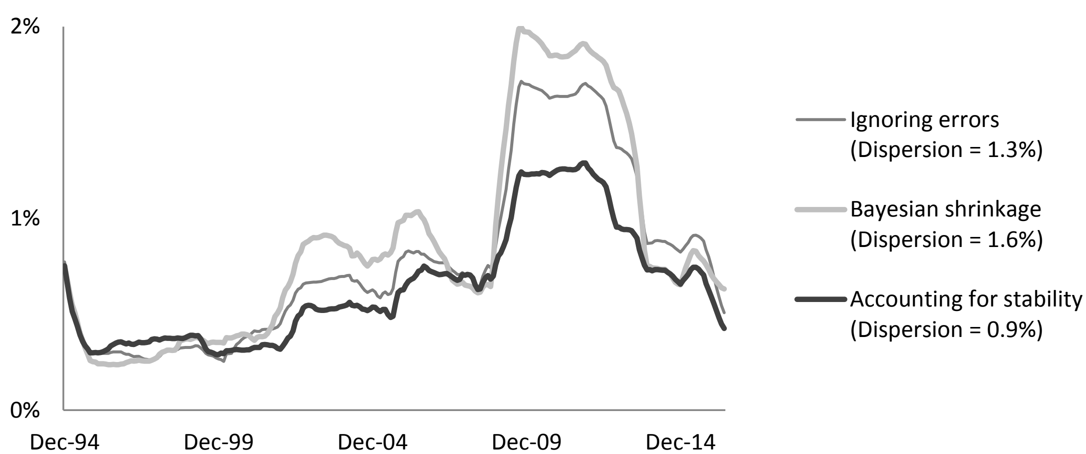


Exhibit 11: Trailing Five-Year Downside Tracking Error and Dispersion across Sub-Samples (Inflation)



Exhibits 10 and 11 offer evidence, at least in the case of inflation, that accounting for stability leads to a better factor replicating portfolio using full-scale optimization just as it did in the case of mean-variance optimization. Moreover, again the greatest improvement occurred during the dot com bubble and the global financial crisis, when risk was elevated – arguably periods when investors would have been most interested in risk control.

Summary

Investors have become increasingly interested in factor investing, but replicating factors is prone to several types of estimation error. Our empirical analysis shows that Bayesian shrinkage does not provide much benefit when applied to the covariance matrix. And we know from first principles that resampling is ineffective if factors are replicated with unconstrained long and short positions. We propose an alternative approach for managing estimation error. We argue that investors should treat estimation error as a distinct component of risk.

Specifically, we introduce a non-parametric procedure for constructing a return sample that explicitly reflects the relative stability of covariances, taking into account small-sample error, independent-sample error, and interval error. We show how to apply mean-variance optimization to replicate factors for investors who have preferences that are well described by mean and variance. However, if investors have utility functions that have kinks or inflection points, we show that mean-variance optimization may not be suitable, because stability-adjusted return distributions may be non-elliptical. We therefore show how to carry out full-scale optimization for special cases in which investors have non-standard utility functions and full-scale optimization is computationally feasible.

We apply mean-variance optimization to replicate non-investable factors with investable assets and to replicate investable factors with subsets of hypothetically cheaper-to-trade securities. And, as an illustration, we also apply full-scale optimization to replicate inflation with a small number of asset classes.

Our analysis reveals that accounting for the relative stability of covariances in the factor replication process yields replicating portfolios that are generically and significantly more stable than portfolios derived either from error-blind replication, that rely on Bayesian shrinkage, or by implication, are formed using resampling.

Appendix A: Equivalence of Resampling and Full-Sample Mean-Variance Optimization

Our objective is to solve for the set of weights that produces the lowest possible standard deviation of excess returns (tracking error) versus a specified target index. This optimal tracking problem can be solved for any set of linear weight constraints, upper bounds, and lower bounds, using quadratic programming. However, it is also interesting to note that in the absence of any constraints on the weights, this problem is equivalent to solving a standard multivariate linear regression. Specifically, our goal is to identify portfolio weights, $\hat{\beta}$, which generate the smallest sum of squared deviations (residuals) versus the target, Y. In matrix notation:

$$\hat{\beta} = \min_{\beta}(Y - X\beta)'(Y - X\beta)$$

where Y is a column vector of target returns, X is a matrix of asset returns with each column representing an asset, andβ is a column vector of asset weights. We assume that all assets have expected returns equal to zero, and that X and Y are shifted accordingly so that their column means are zero. In a portfolio context, assuming zero means is equivalent to assuming that an investor's objective is to minimize tracking error without regard to asset return differences. $\hat{\beta}$ is the Ordinary Least Squares (OLS) solution to the following equation (which does not include an intercept term):

$$Y = X\beta + \epsilon$$

The solution to OLS is well-known and given by:

$$\hat{\beta} = (X'X)^{-1}X'Y$$

The solution can also be expressed in terms of the sample covariance matrix of X, which is given by $\Sigma = X'X/T$, and the column vector of covariances between X and Y, which is given by $Cov(X,Y) = X'Y/T$, remembering that we have assumed X and Y both have zero means for the optimal tracking objective:[20]

$$\hat{\beta} = \Sigma^{-1}Cov(X,Y)$$

The resampling procedure simulates many new samples for X and Y based on their joint normal distribution described by covariance matrix Σ and $Cov(X,Y)$, both of which the resampling procedure assumes are known and are estimated from the full-sample data. Resampling then computes the coefficients (weights) for each of n simulated samples, and averages them to arrive at a final set of portfolio weights.

Under the assumption that X and Y are normally distributed, the weight vector $\hat{\beta}_k$ from a given simulated sample k is normally distributed with the following mean and variance:

$$E[\hat{\beta}_k] = \hat{\beta}$$

$$Var[\hat{\beta}_k] = \sigma^2(X'X)^{-1} = \sigma^2\Sigma^{-1}$$

where σ^2 is the variance of the residuals from the original full-sample regression, and equal to the tracking error squared in a portfolio context. Given that the resampled weights are an average of the weights resulting from each simulated sample, it follows that the resampled weights will converge to the original weights $\hat{\beta}$ when enough simulated samples are used.

$$\hat{\beta}_{resample}(n) = \frac{1}{n}\sum_{k=1}^{n} \hat{\beta}_k$$

$$E\left[\hat{\beta}_{resample}(n)\right] = E\left[\hat{\beta}_k\right] = \hat{\beta}$$

$$Var\left[\hat{\beta}_{resample}(n)\right] = Var\left[\frac{1}{n}\sum_{k=1}^{n} \hat{\beta}_k\right] = \frac{1}{n^2}\sum_{k=1}^{n} \sigma^2\Sigma^{-1} = \frac{\sigma^2\Sigma^{-1}}{n}$$

Appendix B: Elliptical Distributions and Mean-Variance Optimization

We have shown that applying mean-variance optimization to a stability-adjusted return distribution generally yields portfolios with more stable risk than other approaches. However, we should also note that the nature of the stability-adjusted return distribution might impact the decision to use mean-variance optimization in the first place. It can be shown that if an investor's utility function can be approximated by mean and variance or if asset returns are elliptically distributed, of which the normal distribution is a special case, the mean-variance optimal portfolio will be the same as the portfolio that maximizes expected utility. Both of these conditions must be violated to break the guarantee that mean-variance optimality equates to expected utility maximization. There are many plausible utility functions that are not amenable to approximation by just mean and variance. One example is a kinked utility function, in which an investor has strong aversion to losses below a given threshold. It is also plausible that the true return distribution is not elliptical. In particular, even if the stability-adjusted return distribution is symmetrical, it could have features which make it non-elliptical. First, the kurtosis (or more generally the shape of the tail of the distribution) could differ for each asset, depending on the relative stability of its variance over time. Second, the relationships between assets may be nonlinear and impossible to describe using a single covariance matrix.

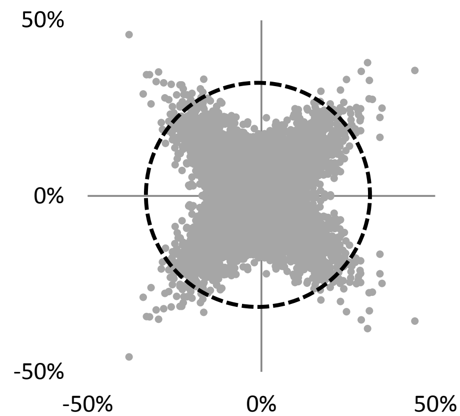
For a vector of asset returns, *x*, to be elliptically distributed, all observations that lie on a given ellipsoid must have the same probability density. In other words, the probability density *f(x)* is proportional to some particular function g(·) of the covariance (Σ) adjusted distance of the observation from the mean (μ):[21]

$$f(x) \propto g[(x-\mu)'\Sigma^{-1}(x-\mu)]$$

Any two observations that are equally unusual in terms of the covariance structure must have equal probability, which means that the shape of the tail must be consistent across assets. For example, a "two-sigma" event for asset A must be just as likely as a two-sigma event for asset B. The stability-adjusted distribution violates this requirement because assets with less stable variance will have fatter tails than those with more stable variance. In other words, the shape of the stability-adjusted distribution depends on the direction of the vector *x*, and not only on the size of the ellipse that contains it.

To see how the co-dependence between assets can make the stability-adjusted distribution non-elliptical, consider the following simple example in which two assets have zero mean, equal variance, and a correlation that is +0.8 half of the time and -0.8 half of the time. The composite distribution will have zero correlation, but that does not mean the assets are not co-dependent. In this case, the iso-probability curves for an elliptical distribution are circles around the origin. Exhibit A1 shows simulated data for the composite distribution, and it is clear that returns are much more likely to occur on some parts of the circle than others, which violates the condition of elliptical distributions even though the distribution is symmetric.

Exhibit A1: Simulated composite return distribution and a sample ellipse



Even if the stability-adjusted sample is not elliptical, it is still possible that mean-variance optimization will offer a very good approximation to the true utility-maximizing portfolio. The quality of approximation is an empirical question.

Notes

The material presented is for informational purposes only. The views expressed in this material are the views of the authors and are subject to change based on market and other conditions and factors; moreover, they do not necessarily represent the official views of MIT, Windham Capital Management, State Street Global Exchange, or State Street Corporation and its affiliates.

References

Britten-Jones, M. 1999. “The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights.” *The Journal of Finance*, Vol. 54, No. 2 (April).

Cocoma, P., M. Czasonis, M. Kritzman and D. Turkington. 2016. “Facts about Factors.” Forthcoming March 2017, *The Journal of Portfolio Management*.

Ingersoll, J. E. 1987. “Theory of Financial Decision Making.” Rowan & Littlefield.

Kritzman, M. and D. Turkington. 2016. “Stability-Adjusted Portfolios.” *The Journal of Portfolio Management*, vol. 42, no. 5, QES Special Issue, pp. 113-122.

Markowitz, M. and K. Blay, *Risk-Return Analysis: The Theory and Practice of Rational Investing* (Volume One) (2014), McGraw Hill, New York

Michaud, R. and R. Michaud. 1998. *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation*. Boston: Harvard Business School Press. 2nd ed. Oxford: Oxford University Press, 2008.

Scherer, B. 2002. “Portfolio Resampling: Review and Critique.” *Financial Analysts Journal*, vol. 58, no. 6, (November / December), pp. 98-109.

---

[1] Although we can approximate the stability of variances analytically, it is not possible to so for covariances. For example, consider two assets which are sometimes 0.8 correlated and sometimes -0.8 correlated. The unconditional correlation may be zero, but a correlation of zero obscures the important fact that the assets sometimes move dramatically in the same direction, and sometimes move dramatically in opposite directions. Therefore, we must use a non-parametric approach to capture instability in the entire covariance matrix.

[2] Let us be clear. We do not explicitly test resampling because we know from first principles that resampling will deliver the same portfolio one would derive by ignoring errors if the replicating portfolio includes both unstrained long and short positions. See Scherer [2002].

[3] We use the term high-frequency to refer to monthly or weekly intervals and not milliseconds.

[4] See Kinlaw, Kritzman, and Turkington [2015] for more detail about this issue.

[5] For a thorough description of resampling, see Michaud and Michaud [1998].

[6] See Scherer [2002] for more detail about this issue as well as empirical validation.

[7] For example, our original sample may comprise weekly returns, but our investment horizon may be one year. Therefore, we would estimate the covariance matrix using annual, overlapping returns. We use log returns to calculate covariance matrices in order to remove the effect of compounding. In particular, we transform each return observation by taking the natural logarithm of one plus the return. The multi-period compounded returns of a normally distributed asset will be highly skewed due to compounding and therefore not normally distributed; however the logarithms of the long-period returns will be normally distributed.

[8] We use overlapping samples to mitigate the distortion that could be caused by choosing a particular start date with independent samples. For example, it could be that a particular period has very high risk and the subsequent period has very low risk. If we were to choose a start date such that we combined half of the first period with half of the subsequent period, we would not capture these extreme episodes of risk.

[9] We remove any strong directional bias from the distribution of errors by subtracting the median error from each individual sub-sample error.

[10] Some of the sub-sample covariance matrices may not be positive semi-definite. This condition becomes more likely as the number of assets grows and the number of independent data points in a given subsample shrinks. We render all covariance matrices invertible by using a standard correction procedure based on principal components decomposition. In particular, the correction computes the eigenvectors and eigenvalues for the original covariance matrix, replaces negative eigenvalues with very small positive numbers, and reconstructs a new covariance matrix. This approach is very similar to using a statistical factor covariance matrix based on a subset of the most important principal components, which is a popular technique for estimating large covariance matrices from limited data samples.

[11] We use the term, mean-variance optimization generically to include instances in which means are non-zero, as well as instances in which means are zero, in which case we are minimizing variance. We also use the term variance to refer to the variance of return differences, which is typically called tracking error.

[12] For an excellent and thorough discussion of the suitability of mean-variance optimization with respect to return distribution and expected utility, see Markowitz and Blay [2013].

[13] To implement an unconstrained set of tracking weights in practice, one would simply need to invest excess proceeds from short sales in the risk free asset, or borrow to fund excess long positions. In either case, the risk free investment or funding cost would have a volatility extremely close to zero, and given that we are concerned with tracking risk in our analysis, these positions would not affect our analysis and can be ignored.

[14] It may be the case that we wish to trade off tracking error with the cost of replication, in which case we would need to include estimates of mean trading costs.

[15] Figure of merit is a term commonly used in operations research. It is a numerical quantity based on one or more characteristics of a system that represents a measure of effectiveness.

[16] The Global Industry Classification Standard (GICS) was developed by and is the exclusive property of MSCI Inc. and Standard & Poor's. GICS is a service mark of MSCI and S&P and has been licensed for use by State Street.

[17] Recall that resampling delivers the same portfolio as one would derive by ignoring estimation error.

[18] We do not actually identify which stocks are cheaper to trade. We randomly select a subset of 30 stocks assuming they are cheaper to trade.

[19] Of the 633 stocks in the MSCI US index as of December 31, 2015, we restrict our universe to the 515 stocks with complete return, market capitalization, and price-to-book data over the period January 2005 through June 2016.

[20] In a similar context which includes non-zero means, Britten-Jones [1999] demonstrates that the maximum Sharpe ratio ("tangency") portfolio corresponding to an unconstrained efficient frontier is equal to the portfolio that results from regression coefficients where the dependent variable Y is a constant vector of ones and the regression model does not include an intercept.

[21] More generally, it is possible for elliptical distributions to have an undefined mean, in which case μ represents the median, or an undefined covariance matrix, in which case Σ represents a dispersion matrix. For a thorough discussion of elliptical distributions and their implications for mean-variance analysis, see Ingersoll [1987].