2021

# Converged Subset Portfolios: An Extension to Subset Optimization

Coleman Cornell

**Claremont McKenna College**

# Converged Subset Portfolios: An Extension to Subset Optimization

**Submitted to**
**Professor Benjamin Gillen**

by
Coleman Cornell

for
Senior Thesis
Spring 2021
May 2, 2021

**Acknowledgments**

Thank you to the professors who inspired and challenged me. Professor Lincoln, thank you for your continual support and mentorship since the day I arrived at CMC. I am beyond thankful for your belief in me. Professor Garín, you were the perfect academic advisor for me. I cannot thank you enough for how hard you pushed me and how much you forced me to grow. Professor Gillen, there is no way to briefly describe the past two years and what they meant to me. Thank you for this part of my academic journey.

To my close friends at CMC, thank you for the continual support and care over these years. I would like to give special thanks to the Bens, JD, and Vasu for their encouragement and suggestions over this nearly two-year long project. The interest you have shown in this thesis as it developed from a summer research project to a complete paper has been appreciated.

Finally, I'd like to thank my family for their constant love. Blake, thank you for your friendship and advice. I am beyond grateful for the experiences we shared this past year. To my parents, I am truly blessed to have you in my corner. Your support, understanding, and affirmation have been fundamental to who I am today. Thank you.

## Abstract

The limited span of useful data, coupled with increasingly expansive asset universes, cripples the traditional mean-variance problem. When optimizing in these environments, the pronounced effect of estimation error yields extremely unstable portfolios when evaluated out-of-sample. As a proposed solution to the "curse of dimensionality," Gillen (2016) presents subset optimization as a technique to reduce the impact of estimation error. Instead of optimizing jointly over the entire asset universe, subset optimization naïvely aggregates over many "subset portfolios" that each optimize over a much smaller random sample of assets. Given the inefficiencies when using naïve aggregation, converged subset optimization is presented as an extension to subset optimization. Simulation and backtest experiments illustrate the potential for outperformance when implementing this method of convergence.

# Contents


**1 Introduction** **4**

1.1 Literature Review . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5

**2 The Algorithm: Converged Subset Portfolios** **8**

2.1 The Mean-Variance Problem . . . . . . . . . . . . . . . . . . . . . . . . . . 9

2.2 Presentation of the Algorithm . . . . . . . . . . . . . . . . . . . . . . . . . 10

2.3 Proof of the Algorithm's Convergence . . . . . . . . . . . . . . . . . . . . . 13

**3 Simulation Tests of Converged Subset Portfolio Performance** **17**

3.1 Simulation Asset Universe . . . . . . . . . . . . . . . . . . . . . . . . . . . 18

3.2 Effective Efficient Frontiers . . . . . . . . . . . . . . . . . . . . . . . . . . 19

**4 Backtest Performance of Converged Subset Portfolios** **23**

4.1 Backtest Asset Universe . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23

4.2 Backtest Performance Under Mean-Variance Preferences . . . . . . . . . . . 25

4.3 Minimum-Variance Performance . . . . . . . . . . . . . . . . . . . . . . . . 27

**5 Conclusion** **28**

**References** **30**

# 1 Introduction

Since its introduction, the classic mean-variance optimization problem developed in Markowitz (1952) has been a foundational problem in financial economics due to its ability to simply represent the two primary objectives of any investor: maximize returns and minimize volatility. Despite the problem's intuitive nature, the practical implementation of mean-variance optimization has been hampered by the sensitivity of recommended portfolios to estimation error. As explained in Michaud (1989), this sensitivity places mean-variance optimization as an "estimation-error maximizer" and often produces poorly allocated portfolios. This estimation error becomes increasingly severe in universes with large numbers of securities due to the commonly referenced "Curse of Dimensionality." In order to tackle this curse of dimensionality, Gillen (2016) introduces subset optimization which aggregates a large number of small, individually optimized subsets of the asset universe. This method sacrifices some potential utility, but allows investors to optimize over the entire asset universe without suffering from extremely pronounced estimation error. The converged subset optimization strategy presented in this paper is a direct extension of this subset optimization strategy. Rather than naïvely aggregating over the entire space of subsets generated when using subset optimization, converged subset optimization systematically creates constituent portfolios to seed an iterative process. This process continues until a fixed portfolio is eventually found. To my knowledge, this method of convergence has not been considered as an extension to the subset optimization strategy.

The paper's discussion begins in Section 1.1 by reviewing much of the key literature for mean-variance optimization across many assets. This review begins by detailing statistical models that incorporate factors from asset pricing models (Sharpe 1964, Lintner 1965, Fama and French 1992, Black and Litterman 1992, Pastor and Stambaugh 2000) and macroeconomic variables (Schwert 1981, Keim and Stambaugh 1986, Ferson and Harvey 1999) to enhance portfolio allocations. A significant part of the literature also leverages Bayesian techniques to provide better estimated returns (Jorion 1986, Frost and Savarino 1986) and regularize covariance matrices (Ledoit and Wolf 2003, Ledoit and Wolf 2004a, Goto and

Xu 2015. Despite advances in estimating the moments of the return generating process, there was still significant evidence suggesting that instability was largely due to the curse of dimensionality (Kritzman 2006, Kan and Zhou 2007). This gave way to a large literature that proposed alternatives to the optimization phase. This large pool of proposed optimizers included stricter constraints (Jagannathan and Ma 2003, DeMiguel et al. 2009a, Fan et al. 2012), Bayesian techniques in the optimization phase (Garlappi et al. 2007, Golosnoy and Okhrin 2009, Tu and Zhou 2010, Anderson and Cheng 2016), or completely novel approaches to optimization itself using firm-characteristics or turnover conditions (Brandt et al. 2009, Olivares-Nadal and DeMiguel 2018, Hautsch and Voigt 2019). Despite the vast and extremely diverse literature on the mean-variance problem, this paper more closely follows early proposals with model averaging approaches (Breiman 1995, Michaud 1998) and is a direct extension of the subset optimization approach outlined in Gillen (2016).

Section 2 provides a brief overview of the mean-variance problem and formally presents and proves the converged subset optimization algorithm. Much of the key sampling properties of the complete subset portfolios discussed in Gillen (2016) extend to converged subset portfolios, but there is some discussion on how the ex-ante exchangability of complete subset portfolios is not immediately extended to the convergence property. In Section 3, the performance of converged subset portfolios are compared to those of complete subset portfolios through a number of simulation exercises. These exercises provide key insights about the sampling properties and recommended implementation of the complete subset optimization algorithm. In Section 4, converged subset portfolios are evaluated against converged subset portfolios, a naïvely weighted portfolio, and a market index in a dynamic backtest environment, which provides a closer approximation to its expected future performance. Section 5 concludes by discussing potential refinements and plans for future experimentation.

## 1.1 Literature Review

The issue of estimation error with sample moments has led to an extensive literature on alternative methods of estimating expected returns. One popular approach has been the development of factor models for returns. An early example of this was the Capital Asset Pricing Model developed by Sharpe (1964) and Lintner (1965) which characterized the expected return of an asset in terms of how much of its risk is correlated with the market. More recent examples of these factor models have augmented the traditional CAPM with an additional element of investor beliefs (Black and Litterman, 1992), included new factors related to firm fundamentals (Fama and French, 1992), or used factor models to construct priors in a Bayesian inference framework (Pastor 2000; (Pastor and Stambaugh, 2000); Pastor and Stambaugh 2002). A second popular approach has been to use general macroeconomic variables as a predictor for the future price of risky assets. Some of these models have used unexpected inflation as a predictor on stock prices (Schwert, 1981), constructed ex-ante variables which act as proxies for changing risk premia (Keim and Stambaugh, 1986), or used lagged bond spreads, bond yields, and dividend yields in an instrumental variable approach for expected returns (Ferson and Harvey, 1999). A final popular approach to reducing estimation error in expected returns has come through the use of different Bayesian approaches. One such example was the use of Bayes-Stein estimators in Jorion (1986) to account for investor uncertainty. A second example was the Frost and Savarino (1986) implementation of an Empirical Bayes rule which shrunk all expected returns to an identical prior determined by the average returns and covariances of the entire investment universe.

While these alternative methods of estimating expected returns certainly reduced estimation error and produced better performing portfolios, these improvements did not extend to extremely large asset universes. In these large universes, estimation error is very pronounced due to the commonly cited "Curse of Dimensionality" in which there is simply not enough information to have precise sample estimates. Kan and Zhou (2007) summarize this issue well by showing that as the number of assets grows, increasing estimation error causes the mean-variance optimizer to take more and more extreme positions. They characterize this effect in terms of a penalty to the variance of a portfolio that is $O(NT)$. With this formulation, investors believe that the expected volatility of a portfolio grows extremely quickly as additional assets are included in their portfolio. When the amount of assets exceeds the length of the sample period, investors then believe that portfolios have near infinite volatility and they become unable to produce an optimal portfolio. A common approach to handling this issue of dimensionality in estimated covariance matrices has been to apply shrinkage factors (Ledoit and Wolf 2003, 2004a, 2004b, 2012, 2013) or regularization techniques (Goto and Xu, 2015) to the covariance matrix itself. In general, these approaches place stricter restrictions on the construction of the covariance matrix in order to reduce the effect of extreme estimation error.

While the impact of estimation error in mean-variance optimization has led to this deep literature on improved estimators, it fails to fully capture the gap between theory and implementation. As described in Kritzman (2006), the issues of portfolio optimization extend beyond estimation error and begin to enter the realm of modeling error. In essence, the most basic mean-variance optimization is not a powerful enough model to consistently produce good portfolios. This has led to a growing literature on the allocation step itself to create stronger performing optimizers.

One such approach has been the development of Bayesian optimizers that help model the uncertainty investors have in their final portfolios. Goldfarb and Iyengar (2003) and Garlappi et al. (2007) both introduce measures of investor uncertainty to reduce the sensitivity of optimized portfolios to estimation error. Golosnoy and Okhrin (2009) develop an flexible shrinkage technique that groups similar assets and shrinks them to shared priors. Anderson and Cheng (2016) develop a Bayesian averaging approach which carries forward the predictions made in past periods based upon their previous informativeness. Tu and Zhou (2010) create a set of objective-based priors for the portfolio weights themselves to better align estimated portfolios with the economic objectives of the investor. A second set of optimizers sought to constrain the weights of the portfolios themselves. Jagannathan and Ma (2003) imposed non-negativity constraints on their portfolios and found that minimum variance portfolios with the sample covariance matrix were extremely similar to those produced using more complicated covariance estimates. DeMiguel et al. (2009a) and Fan et al. (2012) relaxed these constraints either by implementing constraints on the norm of the weights vector or by limiting gross exposure respectively and found similar results. Some other optimizers take different approaches entirely with some examples modeling portfolio weights on firm characteristics instead of the joint return distribution (Brandt et al., 2009) or by directly modeling turnover costs to allow for more frequent rebalancing (Olivares-Nadal and DeMiguel 2018; Hautsch and Voigt 2019). While several new portfolio optimizers have been introduced, there is still much to improve on when it comes to these novel techniques. In a review of several of these proposed methods, DeMiguel et al. (2009b) show that none consistently outperform the naïve portfolio due to the persistent effects of estimation error.

In the absence of a novel optimizer that can bring consistent performance, the implementation of bootstrap aggregation in portfolio optimization has significant potential. The technique of bootstrap aggregation to develop statistical predictors is well outlined in Breiman (1995). In short, the method develops more stable predictors by generating multiple versions of a single predictor and averaging over them to create an aggregate predictor. An early example of bootstrap aggregation methods in portfolio optimization is proposed in Michaud (1998) in which a resampling approach of historical returns is used to construct portfolios. More recently, the introduction of "Complete Subset Portfolios" in Gillen (2016) presents another implementation of bootstrap aggregation techniques to solve the mean-variance problem. Modeled after the "Complete Subset Regressions" developed in Elliott et al. (2013, 2015), these portfolios create aggregate portfolios by averaging over portfolios optimized over a random sample of the subsets in the asset universe. While the algorithm of Gillen (2016) leads to reduced sensitivity to estimation error, its naïve aggregation rule leaves room for further improvements. In this paper, I present a new aggregation technique that aims to better evaluate the performance of subsets to generate better performing portfolios.

## 2 The Algorithm: Converged Subset Portfolios

In this section, I present the converged subset portfolio strategy developed as an extension to the original subset optimization algorithm presented in Gillen (2016). This section is split into three major components. I first provide a brief overview of the classical mean-variance optimization problem. Next I present two algorithms whose combined use produce converged subset portfolios. Algorithm (1) generates the portfolio spaces for each iterative step, and its optimization process closely follows the original algorithm. Algorithm (2) uses these portfolio spaces to generate converged subset portfolios. Lastly, I present a proof of the convergence mapping properties for Algorithm (2) which demonstrates how my algorithm results in a portfolio of converged weights.

## 2.1 The Mean-Variance Problem

Suppose the investment universe has $N$ securities where asset returns have an unknown expected return vector $\mu$ and unknown covariance matrix $\Sigma$. Investors try to learn about these unknown moments by observing $T$ periods of historical asset returns in the information set $D_T$. Given a risk-aversion parameter $\gamma$ and mean-variance utility, investors attempt to find the solution $w^*$ that maximizes the out-of-sample performance of the portfolio:

$$w^* = \underset{w\in\Delta^{N-1}}{\arg\max}\, w'\mu - \frac{\gamma}{2} w'\Sigma w \tag{1}$$

As shown in 1.1, there is a massive literature that examines many different approaches to solving this problem. The standard "plug-in" method has investors estimating the population means and covariance matrices and optimizing their subjective expected utility conditioned on those beliefs. Given posterior beliefs for the expectations ($\hat{\mu}$) and variance-covariance matrix ($\hat{\Sigma}$) of returns, the SEU-optimizing portfolio weights ($\hat{w}$) are:

$$\hat{w} \equiv \underset{w\in\Delta^{N-1}}{\arg\max}\, w'\hat{\mu} - \frac{\gamma}{2} w'\hat{\Sigma} w \tag{2}$$

With this approach, investors are required to estimate $N$ means and $N(N+1)/2$ variances and covariances. Given the generally limited size of useful samples, these estimates are generally fuzzy and give way to issues of dimensionality.

## 2.2 Presentation of the Algorithm

Fundamentally, converged subset portfolios represent an alternative method to weight the subsets generated in the original subset optimization algorithm. To understand this method of convergence, we must first understand why the equal weighting scheme was adopted in the original algorithm. As stated in Gillen (2016), this naïve weighting scheme for subset optimization was selected for two primary reasons. First, as the amount of subsets drawn approaches the total amount of possible subsets, equally weighting the subsets gives an unbiased estimator for the population complete subset portfolio. Second, as each subset is selected uniformly randomly, then the return series for any given subset is exchangeable with the return series of any other given subset. This exchangeability in turn implies that an equal weighting of subset portfolios is ex-ante optimal when estimating the true complete subset portfolio. With this naïve weighting scheme, complete subset portfolios are essentially analogous to the complete subset regressions outlined in Elliott et al. (2015). However, unlike a complete subset regression in which each subset is considered to be equally as informative about the true model, a complete subset portfolio constructed with only a small portion of all subsets will contain some subsets that are strictly preferable to some other subsets.

This key difference informs the decision to develop a method to evaluate these subset. In order to address the exchangeability property found in a standard complete subset portfolio, the concept of a "Complete Constituent Subset Portfolios" is introduced. Define a "Complete Constituent Subset Portfolio" (CCSP) as a complete subset portfolio where a single asset is included in all of its subsets. With these, define a "Complete Constituent Subset Return Space" as a set of historical returns and weights constructed by the set of CCSPs for each of the $N$ assets in our asset universe. Denote the set of historical returns as $D_{\mathcal{N}}$ and the weights to this return space as $W_{\mathcal{N}}$. This return space handles the exchangeability properties of the original algorithm in a useful way. Within a single CCSP, the individual constituent subset portfolios retain the exchangeability property and inform the decision to equally weight the constituent subset portfolios in the construction of a CCSP. However, this exchangeability principle is not maintained across the different CCSPs

**Algorithm 1** Generate Complete Constituent Subset Portfolio Space using Historical Data $D_N$

**For** $n = 1, 2, \ldots, N$

A Construct Individual Constituent Subset Portfolios
**for** $b = 1, 2, ..., B$

I Select subset of assets $\mathcal{N}_{n,b}$ : Including the $n^{\text{th}}$ asset and $\hat{N} - 1$ randomly selected assets.

II Compute optimal subset portfolio weights. Let $\hat{\mu}_{n,b}$ and $\hat{\Sigma}_{n,b}$ denote means and covariance matrix for securities of $\mathcal{N}_{n,b}$ and solve: $\hat{w}_n^{*,b} = \underset{w \in \Delta^{\hat{N}-1}}{\arg\max}\, w'\hat{\mu}_{n,b} - \frac{\gamma}{2} w'\hat{\Sigma}_{n,b} w$

**Next** $b$

B Equally-Weight Constituent Subset Portfolios for Complete Constituent Subset Portfolio

I Let $\hat{w}_{i,n}^{*,b}$ be asset $i$'s weight in subset portfolio $\hat{w}_n^{*,b}$ (and zero for subsets excluding asset $i$) and define

$$\hat{w}_{i,n}^{*} \equiv \frac{1}{B} \sum_{b=1}^{B} \hat{w}_{i,n}^{*,b}.$$

II Define the vector containing each $\hat{w}_{i,n}^{*}$ by $\hat{w}_n^{*} = \left[\hat{w}_{1,n}^{*}, \hat{w}_{2,n}^{*}, \ldots, \hat{w}_{N,n}^{*}\right]'$.

III Generate $D_n \equiv D_T \hat{w}_n^{*}$, the historical returns for the $n^{\text{th}}$ Complete Constituent Subset Portfolio.

**Next** $n$

C Generate Complete Constituent Subset Return Space

I Let $D_{\mathcal{N}}^{+} = [D_1, D_2, \ldots, D_N]$ denote historical returns for complete constituent subset portfolios.

II Let $\hat{W}_{\mathcal{N}} = [\hat{w}_1^{*}, \hat{w}_2^{*}, \ldots, \hat{w}_N^{*}]'$ denote the matrix containing each $\hat{w}_n^{*}$

III Define the Complete Subset Portfolio Weights $\hat{w}^{*} \equiv \frac{1}{N} \sum_{n=1}^{N} \hat{w}_n^{*}$

---

as the subset portfolio draws are not identically distributed from one portfolio to the next. This means that the naïve equal weighting of these CCSPs is not necessarily optimal and allows for further enhancement.

**Algorithm 2** Generate Converged Subset Portfolios

A *Initialization:* Set $j = 0$, let $D^0_{\mathcal{N}} = [r_1, r_2, \ldots, r_N]$ be historical returns, and $\mathcal{W}^0 = W_{\mathcal{N}}$

B *Apply Algorithm 1.A to I.C.II to Generate Complete Subset Portfolio Space for $D^j_{\mathcal{N}}$ with the additional restriction that each asset weight is non-negative, i.e. $\hat{w}^{*,b}_n \geq 0$*

  I Set $j = j + 1$

  II Let $\mathcal{W}^j = \hat{W}_{\mathcal{N}} * \mathcal{W}^{j-1} = [w^j_1, w^j_2, \ldots, w^j_N]'$ be the cumulative weights of the $j$'th iteration in terms of the original asset space.

  III Set $D^j_{\mathcal{N}} \equiv D_T \mathcal{W}^j$ and $w^{j,*} \equiv \frac{1}{N} \sum_{n=1}^{N} w^j_n$

  IV Define $\delta \equiv \max |w^{j,*} - w^{j-1,*}|$ as the largest change in asset weight for this iteration

C *Evaluate Convergence Criteria:*

  I If $\delta > \theta$, the portfolio has not converged, Repeat Algorithm 2.B

  II If $\delta \leq \theta$, then the portfolio has converged; Define the Converged Subset Portfolio Weights as $w^* = w^*_j$

Given this Complete Constituent Subset Return Space, each Complete Constituent Subset Portfolio can essentially be considered an asset itself. Then, the algorithm to generate these spaces can be used iteratively to weight each of these CCSPs. Intuitively, this extension uses the original subset optimization algorithm to develop these funds, and then iteratively generates funds of these funds. Through this process, the individual subsets are more optimally evaluated when compared to the original algorithm.

For the iterative phase, denote $\mathcal{W}^j$ as the asset weights of the cumulatively calculated Complete Constituent Subset Return Space in terms of the original asset space. Further denote $W_{\mathcal{N}}$ as the weights to the Complete Constituent Subset Return Space that uses the cumulatively calculated space as its asset universe. With this construction, we have the following transformation which reweights the cumulative space in terms of the original asset space:

$$\mathcal{W}^{j+1} = \hat{W}_{\mathcal{N}} * \mathcal{W}^{j}.$$

As we show in the next subsection, this iterative process leads to a converged portfolio such that $||\mathcal{W}^{j+1} - \mathcal{W}^{j}|| \xrightarrow[j\to\infty]{} 0$. Interestingly, this convergence happens when the cumulatively calculate Complete Constituent Subset Return Space regains the exchangeability properties of the original algorithm. This is because the converged portfolio occurs when each row in $\mathcal{W}^{j}$ is equivalent, so each of the funds in this cumulatively calculated return space are identical. Given that each asset is identical, then the returns of any Complete Constituent Subset Portfolio for this space is independent and identically distributed from the returns of any other Complete Constituent Subset Portfolio for the same space.

## 2.3 Proof of the Algorithm's Convergence

The real power of the new algorithm comes with its ability to more optimally weight the individual subsets of a complete subset portfolio, and it does so by taking advantage of the algorithm's ability to be applied recursively upon itself. The following theorem and proof formally state the convergence mapping of the second algorithm.

**Theorem 2.1.** *Suppose $N$ assets' returns have mean $\mu$ and variance $\Sigma$. Let $\hat{W}_{\mathbb{N}}$ represent the matrix containing the weights and $D^{+}_{\mathbb{N}}$ represent the historical returns for a Complete Constituent Subset Return Space generated in Algorithm 1. Then for some convergence threshold $\theta > 0$, there is some $J \in \mathbb{N}$ such that Algorithm 2 with nonnegativity constraints on the constituent subset weights will result in a converge to such that:*

$$\delta \equiv \max_{i=1,\ldots,N} |w^{*}_{j} - w^{*}_{j-1}| < \theta.$$

*Proof.* This proof follows from Brouwer's Fixed Point Theorem and the Nested Interval Property:

**Lemma 2.2** (Brouwer's Fixed Point)**.** *Every Continuous mapping of a closed, bounded, and* *convex subset $K \subset \mathbb{R}^n$ into itself has at least one fixed point $x \in K$.*

**Lemma 2.3** (Nested Interval Property)**.** $\forall n \in \mathbb{N}$, *assume we have a closed interval $I_n = [a_n, b_n]$. Assume also that $I_n \succeq I_{n+1}$. Then the resulting sequence of nested closed intervals*

$$ {}_1 \succeq I_2 \succeq \cdots \succeq I_n \succeq I_{n+1} \succeq \cdots $$

*has a nonempty intersection, i.e.*

$$ \bigcap_{n=1}^{\infty} I_n \neq \emptyset $$

Let $\hat{W}_{\mathcal{N}}$ be the matrix containing the weights of each asset $\hat{w}_n^*$ generated by Algorithm 1 to seed Algorithm 2. Set $\mathcal{W}^0 = \hat{W}_{\mathcal{N}}$. Further, let $D_{\mathcal{N}}^0$ be the historical returns for this Complete Constituent Subset Return Space. When considering the convergence of the portfolio, we focus on Algorithm 2.B.II. With this step, we are presented with the function

$$ \mathcal{W}^j = \hat{W}_{\mathcal{N}} * \mathcal{W}^{j-1}, $$

where

- $\hat{W}_{\mathcal{N}}$ is an $N$x$N$ matrix of weights for the Complete Constituent Subset Portfolio for the current iteration's Historical Return Series

- $\mathcal{W}^{j-1}$ is an $N$x$N$ matrix of cumulatively calculated weights from the prior iteration in terms of the original asset space

- $\mathcal{W}^j$ is an $N$x$N$ matrix of cumulatively calculated weights resulting from the current iteration in terms of the original asset space

Consider $\mathcal{W}^{j-1}$. According to our algorithm, it will have the following construction

$$\mathcal{W}^{j-1} = \begin{bmatrix} w_{11}^{j-1} & w_{12}^{j-1} & \cdots & w_{1N}^{j-1} \\ w_{21}^{j-1} & w_{22}^{j-1} & \cdots & w_{2N}^{j-1} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1}^{j-1} & w_{N2}^{j-1} & \cdots & w_{NN}^{j-1} \end{bmatrix},$$

where $w_{in}^{j-1}$ is the weight of the $i^{\text{th}}$ asset in the $n^{\text{th}}$ Constituent Subset Portfolio in terms of the original asset space. The construction of $\mathcal{W}^j$ is identical, except it is for the following iteration of Algorithm 2. Now instead of looking at the entire matrix, let's instead consider this function in terms of one column only, i.e.

$$\mathcal{W}_i^j = \hat{W}_{\mathcal{N}} * \mathcal{W}_i^{j-1}.$$

With this formulation, we are focusing on how a single asset $i$ is weighted over all $N$ constituent portfolios in the original algorithm. We do this to show how the interval of possible values for an individual asset shrinks in successive iterations.

Now consider the matrix $\hat{W}_{\mathcal{N}}$. Due to the nature of the portfolio optimization problem and the nonnegativity constraints we place on asset weights in the iterative phase, this matrix has two important properties that allow for the convergence of asset weights. First, each row in $\hat{W}_{\mathcal{N}}$ sums to 1. Second, each element in $\hat{W}_{\mathcal{N}}$ is greater or equal to 0. These two properties lead to the following result:

$$a_i^{(j-1)} \leq w_{i,n}^{j-1} \leq b_i^{(j-1)},$$

where

$$a_i^{(j-1)} = \min_{n=1,\ldots,N} (w_{i,n}^{j-1}) < a_i^{(j)}$$

$$b_i^{(j-1)} = \max_{n=1,\ldots,N} (w_{i,n}^{j-1}) > b_i^{(j)}$$

With this result, for every asset $i$ in the $j^{\text{th}}$ iteration, we have the following continuous function:

$$\hat{W}_{\mathcal{N}} : [a_i^{(j-1)}, b_i^{(j-1)}] \to [a_i^{(j)}, b_i^{(j)}].$$

As such, the function that maps asset weights from one iteration to another is a continuous mapping of a convex, closed set into itself. So by Brouwer's Fixed Point Theorem, we have that each asset will have a fixed point. Notably these fixed points are not unique as otherwise the algorithm will converge to the naïve equally weighted portfolio. While this proves that the $\hat{W}_{\mathcal{N}}$ for any given iteration has a fixed point, the basis of Algorithm 2 is to change this matrix with every step.

To prove that the weights converge throughout the iterative process, we will be using the Nested Intervals Property to show that the range of possible values that an arbitrary asset weight may take will shrink in successive iterations. For an arbitrary asset $i$ in the $j^{\text{th}}$ iteration, we can construct an interval $I_i^{(j)} = [a_i^{(j)}, b_i^{(j)}]$ such that $w_{in}^j \in I_i^{(j)}$. As shown above, the formulation of Algorithm 2 has the result that

$$\hat{W}_{\mathcal{N}}(I_i^{(j)}) = I_i^{(j+1)} \subset I_i^{(j)}.$$

An important and needed assumption is that these nested intervals are strict subsets. Due to the nature of the portfolio construction phase itself, this assumption is likely to be true for any application of the algorithm. In order for both the upper and lower bounds of these intervals to be preserved, the complete constituent subset portfolios which correspond to the assets that define those bounds would have to fully invest in their respective bounds. In other words, every single portfolio in a very large sample would need to completely invest in the same common asset. This is already a near impossibility for just one of these bounds, let alone both of them. This assumption is also supported by Brouwer's Fixed Point Theorem as well. For a function satisfying Brouwer's Theorem, we have that for some fixed point $x^* \neq x$, $||f(x) - x^*|| < ||x - x^*||$. In other words, the outputs of the function are closer to the fixed point than the inputs. So if the bounds are not fixed points, then those values will shrink closer to some converged weight. With the conditions of the Nested Interval Property satisfied, we know that no matter how many iterations of Algorithm 2 we perform, the intersection of these intervals will be non-empty, i.e.

$$\bigcap_{j=1}^{\infty} I_i^{(j)} \neq \emptyset.$$

It follows that there will be some $J \in \mathbb{N}$ such that the largest of these asset weight intervals will be smaller than some convergence threshold $\theta > 0$, i.e.

$$\max_{i=1,2,\ldots,N} \left( b_i^{(j)} - a_i^{(j)} \right) < \theta \quad \forall j \geq J$$

This result guarantees that the weights of the assets will converge for the same $\theta$, namely

$$\max_{i=1,\ldots,N} |w_i^{j+1,*} - w_i^{j,*}| < \theta \quad \forall j \geq J$$

where

$$w_i^{j,*} \equiv \frac{1}{N} \sum_{n=1}^{N} w_{i,n}^{j}$$

□

# 3 Simulation Tests of Converged Subset Portfolio Performance

As the usefulness of an optimizer lies in its performance, this section compares the out-of-sample performance of converged subset portfolios to complete subset portfolios in a universe of top US stocks. Across multiple simulation and optimizer parameterizations, I demonstrate that converged subset portfolios have promise as an optimization method for asset allocation.

## 3.1 Simulation Asset Universe

The simulation exercise is conducted on a universe of 500 US stocks drawn from the CRSP database. The selection criteria chooses the 500 stocks in the CRSP database with the largest market capitalization on January 3, 2000[1]. From this selection, we construct a series of weekly returns from January 3, 2000 to December 27, 2019 using the daily returns of these stocks from 2000-2019. These weekly returns are calculated by aggregating the daily returns of a stock from Monday through Friday of every week. For days in which a stock does not have a return, we use the return on CRSP's S&P 500 index. Simulated returns are assumed to be normally distributed with expected returns and covariances calibrated to the sample means and covariances of the training sample.

Annualized summary statistics for this asset universe are shown in Table 1. Note that summary statistics for a market index and naïve portfolio are not presented in this section. This is because the simulation exercise is only used to understand the simulated sampling properties of converged subset portfolios and compare their performance in a static environment to the performance of complete subset portfolios. This secondary comparison to a market index and naïve portfolio will occur in the later backtest exercise.

A potential issue with using weekly returns could come from the heavy-tailed nature of the weekly return distribution. Compared to longer frequencies like monthly or annual returns, weekly return distributions have somewhat fatter tails and thus they don't approximate the normal distribution as well. While this could lead to slight problems when estimating the true optimal portfolio, there are two reasons why the choice of weekly returns is still preferred to monthly returns in this exercise. First, weekly return distributions allow for better estimates of volatility in returns, which should allow for more optimal portfolios. Second, we are comparing the performance of two very similar optimizers in this exercise, so the drawbacks of weekly returns should effect both algorithms in a fairly similar way.

[1]The stocks are chosen according to their market capitalization at the beginning of the sample period to account for survivor bias. Stocks that recently achieve a large capitalization will have higher returns in past periods when compared to assets that dropped out. Because of this, the choice to filter stocks based upon their starting capitalization greatly reduces the effect of the relationship between returns and selection on the sample universe.

Table 1: Simulation Universe Cross-Sectional Return Properties

This table reports the cross-sectional properties of weekly returns for the simulated asset universe of Top 500 US Stocks by market capitalization from 2000 to 2019. Weekly returns are calculated based upon daily returns statistics from the CRSP database. The Mean column reports summary statistics related to the average returns in the universe. The Vol column reports summary statistics related to the volatility of average returns in the universe.

| | Mean | Vol |
|---|---|---|
| Average | 11.37% | 37.71% |
| Std Deviation | 6.45% | 14.53% |
| 1% Quantile | -7.67% | 17.31% |
| 10% Quantile | 3.59% | 21.76% |
| 50% Quantile | 11.92% | 35.11% |
| 90% Quantile | 18.28% | 57.98% |
| 99% Quantile | 26.18% | 76.59% |

## 3.2 Effective Efficient Frontiers

When evaluating the performance of an optimization technique, it is important to test is over both a robust set of specifications and a variety of investor preferences. This motivates presenting the results to these simulation exercises as a series of "converged subset effective efficient frontiers" which plots the out-of-sample expected volatility for a converged subset portfolio at an out-of-sample expected return target.

**Definition 1** (Converged Subset Effective Efficient Frontiers of Size $\hat{N}$)**.** *The Converged Subset Effective Efficient Frontier of Size $\hat{N}$ plots the trade-off between out-of-sample expected return and volatility an investor obtains by implementing the converged subset portfolio while using the estimated sample means, variances, and covariances for securities.*

With this presentation, we are able to quickly visualize the expected performance of portfolios across a wide range of possible return targets. Given that this strategy optimizes for the minimum variance portfolio given a certain level of returns, the optimization problem slightly changes. Given posterior beliefs for the expectations ($\hat{\mu}$) and variance-covariance matrix ($\hat{\Sigma}$) of returns, the SEU-optimizing portfolio weights ($\hat{w}$) for a given return level $\mu^*$

are:

$$
\begin{aligned}
\hat{w}_{\mu^*} \equiv & \underset{w \in \Delta^{N-1}}{\arg\min}\, w' \hat{\Sigma} w \\
\text{s.t.} \quad & w' \hat{\mu} = \mu^* \\
& w' \mathbf{1} = 1
\end{aligned} \tag{3}
$$

Figure 1 shows the results of these simulation exercises across several simulation lengths and subset sizes. Simulations consisted of 260 weeks (5 years), 520 weeks (10 years), or 2600 weeks (50 years) to capture different levels of estimation error. Optimization was done with either 10 or 30 assets per subset to compare the performance of the converged subset optimization algorithm to the standard subset optimization algorithm over a wide range of specifications. Ten equally spaced targets were used ranging from 7% annualized to 25% annualized return. The frontiers presented in the figure are the average across 10 independently and identically distributed simulations which are shared across both the converged and complete subset optimizers. With these simulations, portfolio weights are restricted to be non-negative.

There are a few key findings from this simulation exercise. First and foremost, these simulation provide some evidence of outperformance for converged subset portfolios over the naïvely weighting subset portfolios. Across all simulations, we find that for any given level of out-of-sample expected returns, converged subset portfolios have lower out-of-sample expected volatility. This reduction in estimated volatility ranges from 1% to 4% annualized volatility depending upon the simulation specification and portfolio return target. Given the economic significance of these reductions, these exercises demonstrate some of the potential that converged subset optimization has with the portfolio optimization problem.

The second key finding is that portfolios with subsets of 30 assets have significantly less expected volatility than those with subsets of 10 assets. Depending upon the simulation specification and return target, there appears to be a difference of around 1% expected volatility with converged subset portfolios and a difference of around 2% annualized expected volatility with complete subset portfolios. There appears to be no difference in the out-

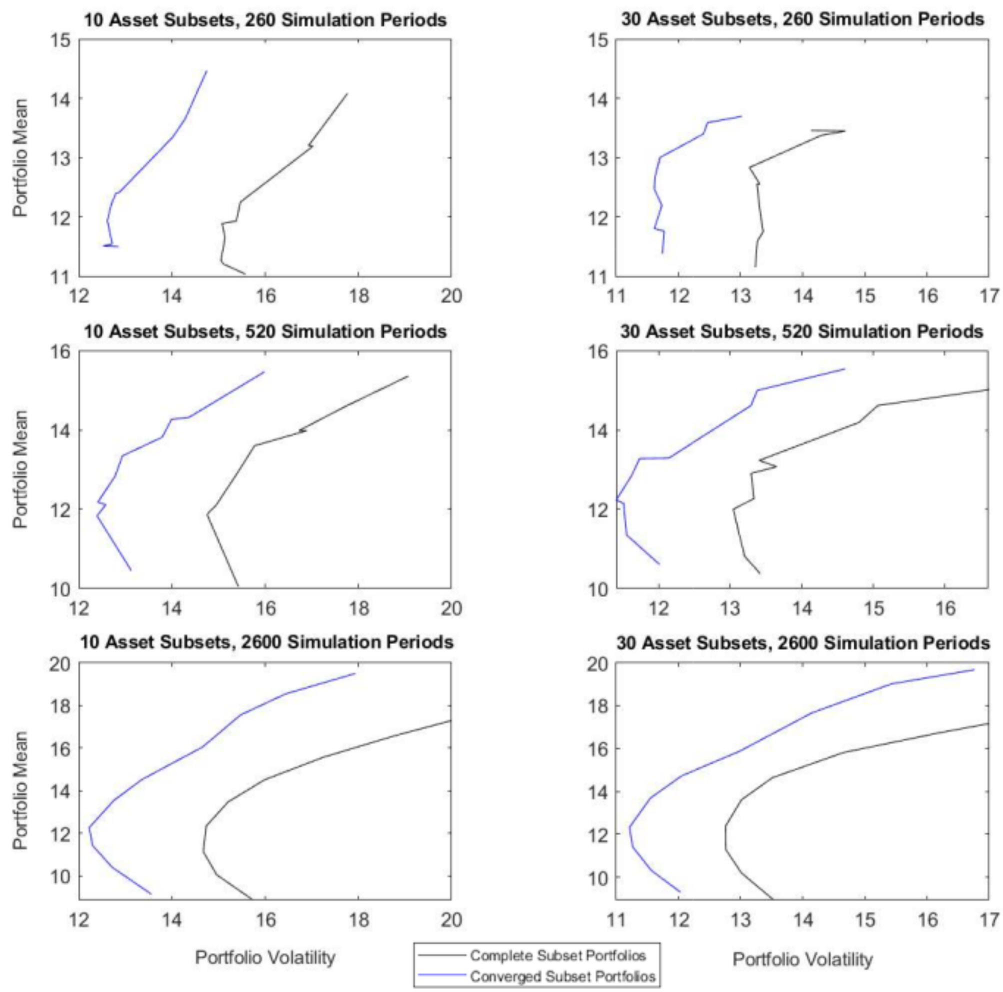


Figure 1: Subset Effective Efficient Frontiers

This figure presents the mean-variance tradeoffs for investors implementing converged subset portfolios and complete subset portfolios using simulated return data calibrated to the sample means and variances of a simulation universe representing 500 Large-Cap US Stocks. For each subset, investors minimize portfolio variance without allowing short sales subject to the expected return target. Converged subset portfolios and complete subset portfolios differ in how they aggregate these subsets. The line for each strategy represents the effective efficient frontiers which plots the mean (y-axis) and volatility (y-axis) of the respective portfolio that minimizes the variance for that level of expected return. Converged subset portfolios dominate complete subset portfolios across multiple subset sizes and simulation lengths, but gains are less pronounced with larger subsets. Portfolio allocations are stable across simulation lengths near the minimum variance portfolio but are extremely unstable at extreme return targets.

of-sample expected returns between portfolios of the different sized subsets. This result follows from Theorem 1 in Gillen (2016) which states that when optimizing with population moments, portfolios with larger subsets will have less expected volatility for any given return target. While the theorem is only analytically true when looking at in-sample results, it is interesting to see it extend to the out-of-sample results as well.

The next key finding is that the gains from convergence appear to be less pronounced in portfolios with subsets of 30 assets compared to the gains exhibited in portfolios with subsets of 10 assets. Intuitively, this result occurs due to the differences in information usage in the original subset optimization algorithm. When including additional assets to a portfolio, investors expand their menu of allocation decisions and allows for better optimization. In essence, this expanded menu allows for investors to use more of their information when determining their optimal portfolio and should lead to better performance. This outcome is shown when comparing the performance of complete subset portfolios constructed with differently sized subsets. Across all three simulation lengths, increasing subset size from 10 assets to 30 assets lead to around 2% less annualized volatility in complete subset portfolios. This additional performance is a direct result of investors becoming able to apply previously unused information. As portfolios with larger subsets have less unused information, then there is less potential gains from applying this method of convergence.

The final key finding is the relative inability to reach extreme return targets when using shorter simulation horizons. Despite each simulation sharing the same set of return targets, the shorter simulation periods (260 and 520 weeks) have much greater difficulty in reaching high expected return targets when compared to the longest simulation period (5200 weeks). In fact, none of the effective efficient frontiers from the 260 period simulations exceed 15% annualized expected return even when trying to reach a target of 25% annualized returns. Although there were few assets in the universe with greater than 20% annualized returns, this problem was much less severe with larger simulations. When looking at these longer simulation, we find that frontiers from the 2600 period simulations were able to reach 20% annualized expected returns. This result suggests that extreme targets suffer most from estimation error and thus require extremely large information sets. This ultimately limits the range of possible return targets when actually implementing the strategy. Interestingly, this same result does not extend to the minimum-variance portfolio. Regardless of the simulation length, the minimum-variance portfolios appeared to be identical. For portfolios with subsets of size 10, the minimum-variance portfolio seemed to always have around 12% annualized return and 12.5% annualized volatility. For portfolios with subsets of size 30, the minimum-variance portfolio appears to have around 12% annualized return and 11.5% annualized volatility. This finding supports a similar result as Jagannathan and Ma (2003), which found that long-only minimum-variance portfolios are generally less prone to estimation error when compared to portfolios with extreme targets.

# 4 Backtest Performance of Converged Subset Portfolios

While the simulation exercises detailed in the previous section demonstrate the potential for the converged subset optimization algorithm, they are extremely limited by the static nature of the simulations. In comparison, backtest exercises add an additional element of dynamic uncertainty, and allows investors to analyze optimization strategies is an environment that is a better approximation to how the market is expected to operate in the future. In this section, the performance of converged subset portfolios is compared to the performance complete subset portfolios, an equally weighted portfolio, and an S&P500 benchmark over a nearly 50 year period. In section 4.2, discussion is focused on the impact of dynamic mispecification of estimated returns has on converged subset optimization. Section 4.3 reveals the potential that converged subset portfolios have as a variance minimizer.

## 4.1 Backtest Asset Universe

The backtest exercise is conducted on a universe of US stocks drawn from the CRSP database from 1963-2019 in which the portfolios are reweighted every year. The training samples for each year include the monthly returns of 500 stocks over the past 10 years. The selection criteria chooses the top 500 stocks by market capitalization at the end of the training period. For example, when evaluating the performance of the algorithm in 1973, our training sample

Table 2: Backtest Evaluation of Benchmark Strategies

This table reports performance statistics from a backtest of several portfolio strategies that act as benchmarks in the backtest exercise. All portfolios constrain weights to sum to one and be nonnegative. The geometric mean, arithmetic mean, and volatility report standard statistics that are annualized from monthly returns. Alpha is the annualized amount the portfolio outperformed the market relative to its risk, and Alpha t-stat measures the significance of that outperformance. Beta-Mkt, Beta-SMB, Beta-HML reflect the portfolio factor loadings from a Fama-French 3 Factor model. Panel A reports results for a complete subset strategy with subsets of size $\hat{N} = 10$ with portfolios reweighted each January according to the sample means and covariances of the past 120 months. Panel B reports the same a complete subset strategy with subsets of size $\hat{N} = 30$. Panel C reports results for the CRSP S&P500 Index and a naïve 1/N portfolio.

| | Panel A: $\hat{N} = 10$ | | Panel B: $\hat{N} = 30$ | | Panel C: Market | |
|---|---|---|---|---|---|---|
| 1973-2019, $T = 120$ | $\gamma = 2$ | $\gamma = 10$ | $\gamma = 2$ | $\gamma = 10$ | S&P 500 | 1/N |
| Geometric Mean | 8.85% | 10.54% | 7.56% | 9.88% | 7.06% | 11.13% |
| Arithmetic Mean | 10.49% | 11.50% | 9.68% | 10.93% | 8.21% | 12.41% |
| Volatility | 18.00% | 13.85% | 20.45% | 13.40% | 15.07% | 15.90% |
| Alpha | -0.53 % | 0.91% | -1.04% | 0.81% | -2.78% | 0.09% |
| Alpha t-stat | -0.774 | 1.463 | -1.106 | 1.175 | -14.195 | 0.182 |
| Beta-Market | 1.073 | 0.880 | 1.136 | 0.882 | 0.993 | 1.022 |
| Beta-SMB | 0.001 | -0.131 | 0.051 | -0.130 | -0.188 | 0.011 |
| Beta-HML | -0.203 | 0.090 | -0.419 | -0.049 | 0.020 | 0.223 |

will consist of the monthly returns from 1963-1972 of the top 500 stocks by market capitalization at the end of 1972. Periods in which a stock does not have a return are replaced by the return on CRSP's S&P 500 index. The sample means and covariances of the training period are assumed to be the true parameters of the return distribution over the evaluation period. When assets are delisted during the evaluation period, investors realize the delisting return and reallocate their holdings to the S&P benchmark. This approximates them allocating these funds to the rest of their portfolio based upon the relative capitalization of the assets.

Table 2 reports summary performance statistics for a number of benchmark strategies to compare against converged subset portfolios. For the complete subset portfolio benchmarks, we see two interesting developments in the backtest environment. First, we find that portfolios implemented by more risk averse investors earned higher average returns than portfolios implemented by more risk neutral investors. Secondly, we find that portfolios generated with subsets of size 10 generally outperform subsets of size 30 and that this outperformance is more pronounced for risk neutral investors. These two results seem to largely stem from the increased estimation error, especially for expected returns, inherent to the backtest environment. Considering backtests using sample means and variances often contain large amounts of this dynamic misspecification, it is understandable that portfolio strategies that are more exposed to estimation error would suffer. Lastly, we note that the 1/N portfolio achieved the highest average geometric return which, similar to DeMiguel et al. (2009b), marks this portfolio as a difficult benchmark to beat.

## 4.2 Backtest Performance Under Mean-Variance Preferences

Table 3 reports summary performance statistics for portfolios generated using the converged subset optimization strategy across a variety of specifications. Compared to both the complete subset portfolios and the 1/N portfolio, the portfolios generated by the converged subset algorithm were underperforming across all specifications. A particularly poor result is that converged subset portfolios were unable to ever achieve greater geometric mean returns or lower annualized volatility compared to their complete subset benchmarks. Further evidence of this relative underperformance is displayed in 4 which compares the performance of converged subset portfolios to each of the benchmarks in rolling monthly and annual windows. While converged subset portfolios were able to outperform complete subset portfolios and the 1/N portfolio in nearly half of all months, their performance was significantly worse in the rolling annual windows.

Interestingly, the relationship between risk-aversion, subset size, and portfolio performance in converged subset portfolios were more pronounced than those found with complete subset portfolios. As the strength of these relationships are largely driven from estimation error, they provide some evidence as to why converged subset portfolios performed so poorly in the backtest experiment. In short, the additional rounds of optimization that the converged subset optimization algorithm requires exposes the strategy to more estimation error. Considering dynamic misspecification is most severe for expected returns, it stands to reason

Table 3: Backtest Evaluation of Long-Only Converged Subset Portfolios

This table reports performance statistics from a backtest of the converged subset optimization strategy implemented by an investor with mean-variance preferences. Portfolios are reweighted every January according to the sample averages and covariances of the previous 120 months. All portfolios constrain weights to sum to one and be nonnegative. The geometric mean, arithmetic mean, and volatility are calculated for each converged subset portfolio and annualized from monthly returns. Alpha is the annualized amount the portfolio outperformed the market relative to its risk, and Alpha t-stat measures the significance of that outperformance. Beta-Mkt, Beta-SMB, Beta-HML reflect the portfolio factor loadings from the Fama-French 3 Factor model. Panel A reports results for a converged subset strategy with subsets of size $\hat{N} = 10$. Panel B reports results for a converged subset strategy with subsets of size $\hat{N} = 30$.

| | Panel A: $\hat{N} = 10$ | | Panel B: $\hat{N} = 30$ | |
|---|---|---|---|---|
| 1973-2019, $T = 120$ | $\gamma = 2$ | $\gamma = 10$ | $\gamma = 2$ | $\gamma = 10$ |
| Geometric Mean | 6.00% | 9.76% | 5.52% | 9.40% |
| Arithmetic Mean | 9.22% | 10.90% | 8.72% | 10.61% |
| Volatility | 25.29% | 15.08% | 25.10% | 15.48% |
| Alpha | -1.24% | 0.80% | -1.78% | 0.712% |
| Alpha t-stat | -0.559 | 0.815 | -0.871 | 0.742 |
| Beta-Market | 1.174 | 0.893 | 1.203 | 0.902 |
| Beta-SMB | -0.014 | -0.192 | -0.003 | -0.150 |
| Beta-HML | -0.523 | -0.034 | -0.569 | -0.125 |

that complete subset portfolios that place great value in expected returns would regularly achieve poor performance.

While the converged subset portfolios performed poorly against complete subset portfolios and the 1/N portfolio, the strategy still performed particularly well against the S&P 500 index. For investors with a $\gamma$ of 10, the converged subset portfolios earned geometric mean returns that were more than 2 percentage points higher than the index, and outperformed in more that 60% of all rolling annual windows. This result shows that despite the strategy not performing well compared to other benchmarks, it still has the potential to produce somewhat strong portfolios even when using the most simple of estimation models.

Table 4: Backtest Performance of Converged Subset Portfolios against Benchmark Strategies

This table reports the relative performance of Converged Subset Portfolios against some benchmark strategies Portfolios are reweighted every January according to the sample averages and covariances of the previous 120 months. All portfolios constrain weights to sum to one and be nonnegative. Complete Monthly and Complete Trail 12 Mnth report the frequency with which the converged subset portfolio monthly return and trailing twelve month returns exceeded that of the complete subset portfolio with the same parameters. Market Monthly and Market Trail 12 Mnth report the same for performance relative to the Market portfolio return. 1/N Monthly and 1/N Trail 12 Mnth report the same for performance relative to the 1/N portfolio. Panel A reports results for a converged subset strategy with subsets of size $\hat{N} = 10$. Panel B reports results for a converged subset strategy with subsets of size $\hat{N} = 30$.

| | Panel A: $\hat{N} = 10$ | | Panel B: $\hat{N} = 30$ | |
|---|---|---|---|---|
| 1973-2019, $T = 120$ | $\gamma = 2$ | $\gamma = 10$ | $\gamma = 2$ | $\gamma = 10$ |
| Complete Monthly | 47.7% | 48.2% | 48.2% | 48.6% |
| Complete Trail 12 Mnth | 47.4% | 42.1% | 45.0% | 38.3% |
| S&P 500 Monthly | 53.5% | 56.7% | 52.0% | 54.3% |
| S&P 500 Trail 12 Mnth | 51.5% | 68.4% | 50.3% | 66.2% |
| 1/N Monthly | 46.6% | 47.0% | 47.2% | 45.2% |
| 1/N Trail 12 Mnth | 42.0% | 41.8% | 39.6% | 39.4% |

## 4.3 Minimum-Variance Performance

The dynamic misspecification of expected returns and the stability of the minimum-variance portfolio in the simulation exercises motivated a final backtest exercise for an investor with minimum-variance preferences. Table 5 reports the annualized volatility for minimum-variance portfolios generated by the converged subset optimization algorithm and the original complete subset optimization algorithm. Unlike the previous exercises, the results of this backtest closely followed earlier results from the simulation exercises. With minimum-variance preferences, converged subset portfolios achieved lower annualized volatilities compared to complete subset portfolios, and the gains from convergence were more pronounced for portfolios produced from subsets of 10 assets than the gains from portfolios produced from subsets of 30 assets. Interestingly, the converged subset portfolio for $\hat{N} = 10$ was able to achieve a lower annualized volatility than the complete subset portfolio for $\hat{N} = 30$.

One explanation for why this backtest environment and the simulation exercises were

Table 5: Backtest Evaluation of Long-Only Minimum Variance Strategies

This table reports annualized portfolio volatility statistics from a backtest of converged subset portfolios and complete subset portfolios implemented by an investor with minimum variance preferences. Portfolios are reweighted each January according to the sample means and returns of the previous 120 months. All portfolios constrain weights to sum to one and be nonnegative. $\hat{N} = 10$ and $\hat{N} = 30$ report annualized volatility for portfolios with subsets of size 10 and subsets of size 30 respectively.

| 1973-2019, $T = 120$ | $\hat{N} = 10$ | $\hat{N} = 30$ |
|---|---|---|
| Converged | 11.78% | 11.48% |
| Complete | 12.76% | 11.96% |

better aligned may come from their more similar objective functions. In the simulation exercise, portfolios were minimum-variance portfolios constrained to a certain in-sample expected return. While the expected returns were important in determining the estimated portfolios, the actual optimization stage was a variance-minimizer. A second explanation for these results could come from the differing degrees of estimation error between the mean-variance exercise and the minimum-variance exercise. The most important difference between the two is the effect of estimation error in expected returns on portfolio weights. Unlike the mean-variance formulation, the absence of expected returns when calculating minimum-variance portfolios eliminates the impact that misspecified expected returns has on calculated portfolios. This is particularly important when considering how poorly estimated expected returns can be. As previously shown in the simulation exercises, expected return estimates from small data sets were subject to large errors even in the static simulation environment. When including the dynamic component of the backtest, this misspecification can cripple converged subset optimization. However, if we remove estimated means from the problem, we unlock the potential of converged subset optimization as a variance-minimizer.

# 5 Conclusion

This paper presents converged subset optimization as an extension to the subset optimization asset allocation algorithm. By leveraging the sampling properties of subset portfolios, this new algorithm seeks to improve upon the naïve weighting scheme of the original algorithm. Across both simulation and backtest experiments, the potential performance gains from these refinements is shown. This early promise and the general flexibility of converged subset optimization warrant further experimentation. Beyond testing less restrictive specifications, the implementation of more sophisticated estimation techniques such as factor models or shrinkage estimates could augment the performance of converged subset portfolios. These possible extensions could be particularly useful considering their potential to reduce the dynamic misspecification of expected returns. Despite these remaining questions, early simulation and backtest experiments demonstrate the power that converged subset optimization could have for the most conservative of investors.

# References


Anderson, Evan W. and Cheng, Ai-Ru (Meg). Robust Bayesian portfolio choices. *Review of Financial Studies*, 29(5):1330–1375, 2016.

Black, Fischer and Litterman, Robert. Global portfolio optimization. *Financial Analysts Journal*, 48:28, 1992.

Brandt, Michael W.; Santa-Clara, Pedro, and Valkanov, Rossen. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies*, 22(9):3411–3447, 2009.

Breiman, Leo. Bagging predictors. *Machine Learning*, 24(2):123–140, 1995. ISSN 1573-0565.

DeMiguel, Victor; Garlappi, Lorenzo; Nogales, Francisco, and Uppal, Raman. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Management Science*, 55:798–812, 2009a.

DeMiguel, Victor; Garlappi, Lorenzo, and Uppal, Raman. Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Review Financial Studies*, 22(5): 1915–1953, 2009b.

Elliott, Graham; Gargano, Antonio, and Timmermann, Allan. Complete subset regressions. *Journal of Econometrics*, 177(2):357–373, 2013. ISSN 0304-4076.

Elliott, Graham; Gargano, Antonio, and Timmermann, Allan. Complete subset regressions with large-dimensional sets of predictors. *Journal of Economic Dynamics and Control*, 54: 86 – 110, 2015. ISSN 0165-1889.

Fama, Eugene F and French, Kenneth R. The cross-section of expected stock returns. *the Journal of Finance*, 47(2):427–465, 1992.

Fan, Jianqing; Zhang, Jingjin, and Yu, Ke. Vast portfolio selection with gross-exposure constraints. *Journal of the American Statistical Association*, 107:592–606, 2012.

Ferson, Wayne E and Harvey, Campbell R. Conditioning variables and the cross section of stock returns. *The Journal of Finance*, 54(4):1325–1360, 1999.

Frost, Peter A and Savarino, James E. An empirical Bayes approach to efficient portfolio selection. *Journal of Financial and Quantitative Analysis*, 21(3):293–305, SEP 1986.

Garlappi, Lorenzo; Wang, Tan, and Uppal, Rama. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *The Review of Financial Studies*, 20(1):41–81, 2007.

Gillen, Benjamin. Subset optimization for asset allocation. *Caltech Social Sciences Working Paper*, 2016.

Goldfarb, Donald and Iyengar, Garud. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.

Golosnoy, Vasyl and Okhrin, Yarema. Flexible shrinkage in portfolio selection. *Journal of Economic Dynamics and Control*, 33(2):317–328, 2009.

Goto, Shingo and Xu, Yan. Improving mean variance optimization through sparse hedging restrictions. *Jounral of Financial and Quantitative Analysis*, 50(6):1415–1441, 2015.

Hautsch, Nikolaus and Voigt, Stefan. Large-scale portfolio allocation under transaction costs and model uncertainty. *Journal of Econometrics*, 212(1), 2019.

Jagannathan, Ravi and Ma, Tongshu. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, 58(4):1651–1683, AUG 2003.

Jorion, Philippe. Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21(3):279–292, SEP 1986.

Kan, Raymond and Zhou, Guofu. Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42(3):621–656, SEP 2007.

Keim, Donald B and Stambaugh, Robert F. Predicting returns in the stock and bond markets. *Journal of Financial Economics*, 17(2):357–390, 1986.

Kritzman, Mark. Are optimizers error maximizers? *The Journal of Portfolio Management*, 32(4):66–69, 2006.

Ledoit, Olivier and Wolf, Michael. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10:603–621, 2003.

Ledoit, Olivier and Wolf, Michael. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365411, FEB 2004a.

Ledoit, Olivier and Wolf, Michael. Honey, I shrunk the sample covariance matrix - Problems in mean-variance optimization. *Journal of Portfolio Management*, 30(4):110–119, SUM 2004b.

Ledoit, Olivier and Wolf, Michael. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *The Annals of Statistics*, 40(2):1024–1060, FEB 2012.

Ledoit, Olivier and Wolf, Michael. Optimal estimation of a large-dimensional covariance matrix under Stein's loss. *Mimeo*, 2013.

Lintner, John. The valuation of risk assets and selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1), 1965.

Markowitz, Harry. Portfolio selection. *The Journal of Finance*, 7(1):77–91, March 1952.

Michaud, Richard. The markowitz optimization enigma: Is 'optimized' optimal. *Financial Analysts Journal*, 45(1):31–42, 1989.

Michaud, Richard. *Efficient Assset Management: A Practial Guide to Stock Portfolio Optimization*. Oxford University Press, 1998.

Olivares-Nadal, Alba V. and DeMiguel, Victor. Technical note - a robust perspective on transaction costs in portfolio optimization. *Operations Research*, 66(3):733–739, 2018.

Pastor, Lubos. Portfolio selection and asset pricing models. *Journal of Finance*, 55(1): 179–223, FEB 2000.

Pastor, Lubos and Stambaugh, Robert. Comparing asset pricing models: an investment perspective. *Journal of Financial Economics*, 56(3):335–381, JUN 2000.

Pastor, Lubos and Stambaugh, Robert. Investing in equity mutual funds. *Journal of Financial Economics*, 63(3):351–380, MAR 2002.

Schwert, G William. The adjustment of stock prices to information about inflation. *The Journal of Finance*, 36(1):15–29, 1981.

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

Tu, Jun and Zhou, Guofu. Incorporating economic objectives into Bayesian priors: Portfolio choice under parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 45: 959–986, 2010.