Portfolio Optimization Using the Atkinson Index*

Nils Billgert†

June 11, 2019

Abstract

Traditional mean-variance optimization of portfolios has received much criticism due to its inability to account for higher order moments and non-quadratic utility. In this thesis, the topic of portfolio optimization is studied using the Atkinson index with CRRA utility. We construct Atkinson-efficient portfolios using computer-generated data in Monte Carlo simulations, as well as using financial asset returns data of assets taken from American stock exchanges. The results show that when normality holds, there are no benefits to the usage of the Atkinson. Under non-normality, however, there are advantages to using the Atkinson instead of the Sharpe ratio. These advantages are great enough for researchers to disregard its greater complexity.

Keywords: atkinson index, performance measure, portfolio optimization

*In partial fulfillment of the Master's of Science in Economics program at the Department of Economics, Lund University. Special thanks go to my supervisor, Asst. Prof. Thomas Fischer, for his patient academic guidance and support.

†Electronic address: nils.billgert@gmail.com

Contents

1

Introduction

Traditional mean-variance optimization of portfolios has received much criticism due to its inability to account for higher order moments and non-quadratic utility (Burr Porter and Gaumnitz, 1972; Jondeau and Rockinger, 2005). When returns are not normally distributed and utility is not quadratic, the mean-variance framework will fail to hold (Baron, 1977, p. 1683; Burr Porter and Gaumnitz, 1972, p. 438). In the context of this framework, it is well known that financial returns suffers from distributional issues, such as non-normality (Jondeau and G. M. Rockinger, 2006, p. 29). Jondeau and Rockinger (2005) are among those who have shown that non-normality can lead to large opportunity costs for risk-averse investors.

There are many measures that have been constructed to deal with the issues of the mean-variance framework, however they are often subject to criticism themselves. Value-at-Risk (VaR) models are quite commonly seen in use among finance professionals, however, along with other drawdown performance measures, it suffers from the weakness that it is a monotonic function of the standard deviation under most return distributions (Eling and Schuhmacher, 2011), which suggests it does not give the investor any substantial amount of extra information. Other performance measures, such as the application of the Gini index to financial returns, have been shown to be easily manipulated.

In a recent paper by Fischer and Lundtofte (2019), the Atkinson Index, primarily known for its use in measuring inequality, is applied to measure risk in financial returns. The Atkinson index covaries weakly with the Sharpe ratio, and it is robust to manipulation (Fischer and Lundtofte, 2019). In this thesis we use this measure to construct an efficient portfolio of assets with non-normal returns for investors with Constant Relative Risk Aversion (CRRA) utility, since as it currently stands, what has been studied is merely its behavior in the context of comparing its value for individual assets. It is interesting to analyze its behavior in the weighting of assets in a portfolio, both in the two-asset simulation case and in the many-asset real-world data case. This is done while carefully auditing the properties of the data used.

In this thesis, we examine efficient portfolios constructed by the Atkinson Index. We do this firstly for the two-asset case using simulated daily financial returns, and secondly for the many-asset case using a set of ten assets selected from the largest stocks by market capitalization in the S&P 500 stock index. We evaluate these efficient portfolios generated by the Atkinson index in relation to benchmark-portfolios, generated as mean-variance efficient portfolios, in order to answer two questions:

1. Are there advantages to using the Atkinson index instead of other measures, specifically the Sharpe ratio?

2. If there are advantages to using the Atkinson index, are the advantages great enough to warrant the more complex calculations required to compute it and the greater complexity of its interpretation?

We start by applying it to simulated data, after which we examine its behavior with real-world financial returns data.

The remainder of this thesis will be structured as follows: Section 2 will consist of a review of previous literature on the topic of alternative measures used in portfolio optimization, and a theoretical background. In section 3, the approach to method will be explained in detail for easy reconstruction of the results, which will be presented and discussed in section 4. Section 5 concludes the thesis.

2

Literature Review & Theory

In a classical paper written by Anthony B. Atkinson (1970), the Atkinson index was proposed as a measure of inequality. Motivating him was the need to address known issues with other measures, such as variance or the Gini index. In particular, his criticism was directed towards the latter, which he explains is not suitable for use under all conceivable circumstances. Below we will outline the issues, Atkinson's proposed solution, and how we can apply it to construct efficient portfolios of financial assets.

2.1 The Mean-Variance Efficient Portfolio

The mean-variance framework was first introduced by economist Harry Markowitz (1952), and has since then been the most commonly used framework to select and weight assets optimally for a portfolio. Previously, it had been proposed that portfolios should be constructed using a mixture of the assets with the highest expected returns, arguing that the law of large numbers applies and that the realized returns of the portfolio should therefore equal the expected value of the portfolio. This is essentially arguing that even market risk can be eliminated through diversification. Markowitz took issue with this assumption, since assets considered for a portfolio are often correlated and that these correlations needs to be taken into account. Even diversification is limited to the elimination of idiosyncratic risk.

A decade later, William F. Sharpe (1964) had developed an extension of this framework that came to be called the Capital Asset Pricing Model (CAPM), introducing the optimization problem that he used to construct mean-variance efficient portfolios. In the CAPM model, we assume that there are k assets with a corresponding excess return vector r that satisfy $r \sim MVN_k(\mu, \Sigma)$, where μ is the expected excess return vector and Σ is the true variance-covariance matrix of the considered assets. With a weights vector w, we minimize the expression (Chen, Hsien, and Lin, 2011).

$$\begin{aligned} \max_{w} \ & \frac{\mu^T w}{\sqrt{w^T \Sigma w}} \\ \text{subject to } & \hat{i}^T w = 1 \\ \text{and } & w_i \geq 0. \end{aligned} \tag{2.1}$$

where $\hat{i}$ is the unit vector. This maximizes the Sharpe ratio (hereafter called the *Sharpe*) of the portfolio given the constraints that the sum of all weights equals one (i.e., that all wealth is invested in the portfolio), and that all asset weights need to be non-negative (short-sale constraints).

In this thesis we use numerical optimization methods, however, the closed-form solution to this maximization problem is useful to understand the parameters involved. This closed-form solution to the problem is given by:

$$w^* = \frac{1}{\hat{i}^T \Sigma^{-1} \mu} \cdot \Sigma^{-1} \mu. \tag{2.2}$$

For a more explicit display of the parameters involved, we consider the optimal weight of asset i, w_i^*, which for the two-asset case is given by:

$$w_i^* = \frac{\mu_i \sigma_j^2 - \mu_j \rho_{i,j} \sigma_i \sigma_j}{\mu_i \sigma_j^2 + \mu_j \sigma_i^2 - (\mu_i + \mu_j) \rho_{i,j} \sigma_i \sigma_j}, \tag{2.3}$$

where μ_i and μ_j are the expected excess returns of assets i and j respectively, σ_i and σ_j are the true standard deviations of assets i and j respectively, and $\rho_{i,j}$ is the coefficient of correlation between assets i and j.

The assumption that the returns follow a multivariate normal (MVN) distribution is problematic since it is almost never satisfied in real-world financial return data, which often exhibit skewness and excess kurtosis. In addition, financial asset returns certainly do not all follow the same distributions. For example, given two assets with the same mean and variance, a negative skewness in only one of them would imply a greater downside risk in this asset than in the other. However, this risk is not recognized by the mean-variance framework and thus not by the CAPM, the performance in the two assets will be considered the same. To maximize expected utility when the assumption of multivariate normality is not satisfied, we have to instead assume a quadratic utility function for mean-variance optimization to be the best method. This assumption is restrictive for a researcher.

2.2 The Atkinson Efficient Portfolio

In this section, we will discuss previous attempts to solve the issues described above within a context of portfolio optimization, as well as the advantages of using our proposed solution.

2.2.1 Some Previously Proposed Solutions

Before we introduce the solution that we use to the issues described in the above section, it is worth noting that previous attempts have been made. Extensions and modifications of the mean-variance framework have been developed to account for skewness (Kraus and Litzenberger, 1976; Konno and Suzuki, 1995) and kurtosis (Lai, Yu and Wang, 2006; Stacy, 2008). However, none of these frameworks consider entire distributions, nor are they proven to be fully compatible with expected utility maximization under various utility functions. Additionally, modifications of the Gini index (hereafter called the *Gini*), mostly known for its use in measuring income inequality (Gini, 1912), have been used for the same purpose (Shalit and Yitzhaki, 1984; Hespeler and Shalit, 2016).

The approach to use the Gini to construct an efficient portfolio, however, has its flaws. Fischer and Lundtofte (2019) show that it is possible to construct an asset using out-of-the-money call and put options such that its Lorenz curve intersects with that of an asset of an asset following a Normal Inverse Gaussian (NIG) distribution. This implies that one cannot make a statement about second order stochastic dominance to rank these assets against each other using the Gini (Cowell, 2000). In finding the Gini-efficient portfolio, not being able to rank the assets would cause issues in deciding on which assets to include in the portfolio and using which weights.

2.2.2 Our Proposed Solution

In order to deal with cases where we have to compare assets that exhibit non-normality, intersecting Lorenz curves, and the appropriate utility function is non-quadratic, Fischer and Lundtofte (2019) derive a measure of financial performance from the Atkinson index (hereafter called the *Atkinson*). Let us first state a simple definition of the Atkinson to illustrate its interpretation in an income inequality context:

$$A = 1 - \frac{y^{CE}}{E[y]}, \tag{2.4}$$

where y is the income of an individual and y^{CE} is the certainty equivalent income for the individual with a given utility function and aversion-to-inequality (or risk-aversion, as we will see below) parameter. Since aversion to inequality ($\nu > 0$) implies that $0 < \frac{y^{CE}}{E[y]} < 1$, the Atkinson is always between zero and one for risk averse individuals. As y^{CE} approaches $E[y]$, the Atkinson approaches 0. This happens when either the distribution of income is approaching equality across the population, or when the aversion to inequality approaches zero. Thus, a lower Atkinson is considered favorable.

The Atkinson has several interpretations. Perhaps the most intuitive way is to think of it as an individual that is faced with a decision under uncertainty: to choose a country in which to live without knowing where on the income distribution he or she will end up. We can then interpret equation 2.4, the general-case Atkinson, as the share of income, in relation to the mean $E[y]$, that the individual is willing to give up to be certain about their place on the income distribution. The certainty equivalent income, y^{CE}, is the amount of income that the individual will get.

Equation 2.4 is a general-case formulation that is easy to interpret. For the special case of constant relative risk aversion (CRRA), the Atkinson is instead written as:

$$A(\nu) = \begin{cases} 1 - \frac{1}{E[y]}(E[y^{1-\nu}])^{\frac{1}{1-\nu}} & \nu > 0,\, \nu \neq 1 \\ 1 - \frac{1}{E[y]}e^{E[\ln y]} & \nu = 1, \end{cases} \tag{2.5}$$

where the parameter ν is the risk aversion of the individual.

Fischer and Lundtofte (2019) apply the Atkinson to fund returns. They do this by considering the Atkinson in terms of a function of future wealth, instead of income. They therefore substitute y with w_0R, where w_0 is the initial wealth, wealth at $t = 0$, and R is the gross return (total return without deduction of fees or other expenses) earned on this initial wealth. The returns can be compounded continuously, $R = e^r$ or discretely $R = 1 + r$, where we in the latter case assume that $r \geq -1$ so that w_0R always stays non-negative, which is required for the Atkinson to be defined. Wealth as a non-negative constant is intuitive as long as we restrict debt to a maximum of w_t, or do not consider it at all, in which case w_t is the amount of cash you have to invest at time t.

Nevertheless, the substitution yields:

$$A = 1 - \frac{w_0R^{CE}}{E[w_0\tilde{R}]} \stackrel{CRRA}{=} 1 - \frac{R^{CE}}{E[\tilde{R}]} = 1 - \frac{R_E^{CE}}{E[\tilde{R}_E]}, \tag{2.6}$$

where $\tilde{R}_E = \frac{\tilde{R}}{R_f}$ is the geometric excess returns. The CRRA case is especially interesting as, for this family of utility functions, the Atkinson is independent of both initial wealth, w_0 and a constant risk-free rate, giving more generalizable results. This result is shown in equation 2.6.

Fischer and Lundtofte (2019) further prove that this measure satisfies n-th order stochastic dominance, even with the presence of intersecting Lorenz curves, unlike previously existing performance measures. This means that the Atkinson takes into account not only the first, second, third or fourth moment, but a potentially infinite amount of moments. This is especially relevant in the case of CRRA or CARA utility functions, where a utility maximization problem using the mean variance framework will fail to assign weights such that expected utility is maximized, unlike in our Atkinson framework. Intuitively, we can understand that there is a negative relationship between performance[1] and the Atkinson. We therefore expect the Atkinson to depend:

- negatively on mean (higher reward). Higher mean is considered higher reward in the sense of higher expected excess returns.
- positively on standard deviation (higher risk). Higher standard deviation is considered higher risk in the sense of higher average fluctuations.
- negatively on skewness (higher reward). Higher skewness is considered higher reward in the sense of more outliers on the positive end of the distribution in relation to outliers on the negative end.
- positively on excess kurtosis (higher risk). Higher kurtosis is considered higher risk in the sense of more frequent outliers on both ends of the distribution, thus fluctuations are more often large when occurring.
- positively on risk aversion. Risk aversion emphasizes any risk in the returns distribution, and downplays any reward.

As for Atkinson-efficient weights in a portfolio, we expect them to depend negatively on the Atkinson. With regards to this measure, we follow the pattern of maximizing the Sharpe, by considering the following optimization problem:

$$\min_{w} \begin{cases} 1 - \frac{1}{E[\tilde{R}]}(E[\tilde{R}^{1-\nu}])^{\frac{1}{1-\nu}} & \nu > 0,\, \nu \neq 1 \\ 1 - \frac{1}{E[\tilde{R}]}e^{E[\ln \tilde{R}]} & \nu = 1 \end{cases}$$

[1]Fischer and Lundtofte (2019) consider the Atkinson a *risk measure*. However, in this thesis it will be considered in terms of performance in a risk-reward sense, as it behaves analogously to the Sharpe ratio (with which it will be compared) with respect to the first two moments. As such, we will call it a *performance measure* in this thesis.

$$\text{subject to } \hat{i}^T w = 1$$

$$\text{and } w_i \geq 0 \tag{2.7}$$

to construct an Atkinson-efficient portfolio. This minimizes the Atkinson of the portfolio given the constraints that the sum of all weights equals one (i.e., that all wealth is invested in the portfolio), and that all asset weights need to be non-negative (short-sale constraints). Considering an entire portfolio of assets, not just individual assets, we have additional class of parameters to consider when calculating the Atkinson: the cross central moments. In accordance with the Sharpe-efficient portfolio, we expect the Atkinson-efficient weights to depend negatively on covariance between assets considered for the portfolio. Moreover, by extension we believe that higher order cross central moments, such as coskewness and cokurtosis, should also affect the Atkinson-optimal weights, positively and negatively respectively. This is due to coskewness having been shown to decrease (Campbell and Akhtar, 2000, pp. 1270-1271) perceived risk and cokurtosis to increase (Fang and T.-Y. Lai, 1997) perceived risk.

3

Method

Since the Atkinson is supposed to account for higher moments, unlike the Sharpe, we want to compare the efficient portfolios that these performance measures yield under both normality and non-normality, in order to find out whether, when and to which extent there are advantages to using the Atkinson. We therefore need to write scripts in Python that construct these portfolios. This is done under various circumstances to investigate and compare the weights they yield under normality, negative skewness and kurtosis.

The construction of these portfolios, especially the Atkinson ones, require numerical optimization methods. To conduct the analysis, we firstly find it useful to perform Monte Carlo simulations where we have control over the properties of the distributions of the data. Secondly, we want to investigate and compare their behaviours under circumstances that would be more likely to occur in reality, and thus perform an optimization on financial assets taken from real-world stock exchanges. To make this possible, we write the optimization scripts in Python. Note that we do not consider a risk-free interest rate, as this would unnecessarily complicate the script without benefit. This could be interpreted as assuming that the risk-free interest rate $r_f = 0$.

3.1 The L-BFGS-B Optimization Algorithm

The limited-bounded Broyden-Fletcher-Goldfarb-Shanno (L-BFGS-B) algorithm (Fletcher, 2000) is an optimization algorithm that finds a local optimum of an objective function. It is a numerical method, and as such it iteratively tests points on the k-dimensional space that the function spans until it finds the optimum. Rather than randomly testing points until it finds the optimum, it does this by analyzing gradients and Hessians in order to assess an appropriate direction in which to move and re-perform the analysis, each time with the aim of further approaching the optimum. This way it tries to minimize the amount of iterations needed.

L-BFGS-B is a version of the BFGS algorithm, which belongs to the Quasi-Newton family of algorithms, a family of algorithms that were created to reduce the computational power needed to perform optimizations using Newton's method (Dennis and Moré, 1977). This is done through approximations of Hessians (and sometimes gradients), rather than performing the exact computations that are are needed in each iteration using Newton's method. This technique is useful in situations where the computational power needed to compute the exact values of these matrices is large, or when they are otherwise unavailable. This particular version of BFGS limits the memory usage by limiting the amount of Hessians that need to be stored, and is extended to perform bound-constrained optimization (Byrd et al., 1995).

We use this algorithm for its easy availability in the SciPy package, and for the performance issues that may be encountered in Monte Carlo simulations using more computationally heavy optimization. While the approximations can cause inaccuracies in the process of finding the optimum, this mainly affects the amount of iteration needed to find it. We do not expect the results to be significantly less accurate than they would have been using the full Newton's method, as Quasi-Newton methods have been shown to be very reliable (Schittkowski, 1981; 1987).

For this algorithm to successfully find an optimum every time it is used, the function to be optimized needs a unique local minimum/maximum, at least under any defined constraints. In order to demonstrate that the functions that we optimize indeed do have unique local minima/maxima, we run an optimization on the same set of $k = 10$ assets 100 times. We use the property of this algorithm that it needs an initial guess to be implemented in Python, by simply providing a new, randomly generated, initial guess each time.

If the function has several local minima/maxima, then the algorithm should find different minima/maxima depending on where the initial guess is placed. While we note that this should not be mistaken for a mathematical proof of whether the functions have unique local minima/maxima, we can use it to draw conclusions about whether this property is likely or not to be present. If the optimization shows different results for multiple initial guesses, then we cannot assume that there is a unique local minimum. If they show the same result, however, then we can assume this. The conclusion should hold for any k number of assets. The initial-guess weights are randomly generated in Python by generating a set of k and normalizing these to sum up to 1. The results are shown in Appendix A.1.

3.2 The Optimization Script

The numerical optimization is performed using Python, which provides numerical optimization algorithms in the SciPy library. Specifically, in minimizing the Atkinson and maximizing the Sharpe under bounds and constraints, the *scipy.optimize.minimize()* method is used. As of SciPy v1.2.1, the method uses the L-BFGS-B optimization algorithm as default optimization algorithm (Jones et al., 2019). The data generation and optimization is looped 1000 times to plot the distributions of the weights and calculate interesting properties to assist us in the analysis.

3.2.1 Data Generation

We need a large data set with specified properties. Therefore, it is convenient to generate this data ourselves, rather than to only use real-world financial returns data. In each run of the script, a different combination of distributions is considered. The script generates a $T \times k$ matrix X of random variables of specified distributions. Due to the lack of a comprehensive data generating package to be used in Python, the data needs to be generated in a couple of different ways, depending on the distributional properties that we want.

To generate data that follow a normal distribution, we use *numpy.random.randn()* from the NumPy library (Oliphant, 2019). For a skew-normal distribution, a family of distributions that generalizes the normal distribution to allow for continuous variation from normality to non-normality (Azzalini, 1985), we use *scipy.stats.skewnorm()* from the SciPy library (Jones et al., 2019) and assign a value of -1 to the skewness parameter. For kurtosis, we define a function ourselves that generates a normally distributed variable Y (again, using *numpy.random.randn()*) and transforms it according to Fleishman's power method (Fleishman, 1978):

$$Z = a + bY + cY^2 + dY^3.$$

With this method it is possible to specify parameters a, b, c, d such that data is generated from a distribution with specific values for the excess kurtosis parameter (Luo, 2011), however, since we are only interested in the sign of this parameter, not the exact value, we are satisfied with specifying $a = b = c = 0$ and $d = 1$. This creates positive excess kurtosis in our data.

The precision with which the skewness and excess kurtosis of the non-normal distributions can be controlled is quite low with these simple specifications, however we get around this limitation by only using the distributions that fall within certain specified intervals of these parameters, throwing away the rest. Means and standard deviations are simpler to control when needed after the data has been generated by addition and multiplication. When investigating behaviors of optimal weights under skewness and kurtosis, we therefore keep means and standard deviations constant over each simulation, in order to isolate the effects of the higher moments.

With this data generating script, a trade-off arises between proximity to the distributional properties we seek and efficiency of the script. Since the data generating process is not very computationally heavy with today's computers, we lean towards prioritizing the first. Therefore we throw away any data that is generated with kurtosis above ± 0.3 when we investigate negative skewness, and any data generated with a skewness greater than ± 0.1 when investigating excess kurtosis. Due to practical limitations, moments above the fourth are not controlled for nor computed.

After we have generated data for the k assets we consider in a simulation, we want to transform it into data that simulate correlated assets. We achieve this by replacing columns $2, 3, ..., k$ of the X-matrix with columns that are calculated to correlate according to a specified $(k-1) \times 1$ ρ_r-vector, in which the correlations with X_1 are specified. The correlated vectors are calculated using the formula:

$$X_{i,correlated} = \rho_i \cdot X_1 + \sqrt{1 - \rho_i^2} \cdot X_{i,uncorrelated}, \;\; i = 2, 3, ..., k. \tag{3.1}$$

3.2.2 Optimization

Having generated an appropriate X matrix, it is then used to calculate a matrix M of expected values to calculate the variance-covariance matrix:

$$\Sigma = \frac{1}{T-1}(X-M)'(X-M). \tag{3.2}$$

The script then specifies a set of parameters that the *scipy.optimize.minimize()* method takes, including the constraints and bounds, along with the function to minimize. We define the bounds of the weights $w_i \in [0,1]$ (short constraints) and then the constraint that $\hat{i}^T w = 1$ (that all wealth is required to be invested in the portfolio). After this, the optimization is ready to be performed. We define one function per performance measure:

1. The calculation of the Atkinson is specified in a function that is given the $k \times 1$ weight matrix, the amount of observations T, the $T \times k$ matrix X, and the risk-aversion parameter ν. The Atkinson is multiplied by a factor of 1 000 000 to deal with step-size issues in the optimization algorithm. Specifically, certain values of the correlation ρ_r and risk aversion ν parameters, the Atkinson takes values that are so small that the algorithm mistakes the gradient for zero at the first iteration and thus believes it has found the minimum. Upscaling the Atkinson parameter solves this issue. Naturally, before any interpretations of the Atkinson are made, we scale it back to have $A \in (0,1)$ after the minimization is performed.

2. As for the mean-variance framework, the optimization algorithm calls a calculation of the portfolio Sharpe (with no risk-less asset: $\frac{E[r_p]}{\sigma_p}$) given the $k \times k$ variance-covariance matrix Σ and the $k \times 1$ weights vector. In order to maximize this function using a minimization algorithm, the function returns the negative of the Sharpe. Similarly to the Atkinson case, we correct this value before any interpretations are made.

This optimizes equations 2.7 and 2.1 respectively.

3.2.3 Monte Carlo Simulations

For each case that we investigate, we loop the data generation and the optimization 1000 times to find the distribution of the weights they yield under each set of circumstances. The circumstances that concern us is not limited to multiple types of distributions, but also include multiple values for the correlation ρ_r and risk-aversion ν parameters. Therefore, we do not only run one Monte Carlo simulation per type of distribution, we run several. In each one we assign different values to the risk aversion and correlation parameters, investigating all combinations that are interesting. Having pre-specified the values we want to investigate for each parameter, the script runs two loops, one within the other. The outer loop changes the correlation parameter ρ_r and the inner one changes the risk-aversion parameter. This is done so that for each correlation coefficient, a Monte Carlo simulation for each interesting value of the risk-aversion parameter is run. Naturally, this is done for all combinations of distributions that we consider.

In our simulations we consider the Atkinson framework with a utility function and risk-aversion parameter, however the mean-variance framework without these. Since our main goal is to investigate the effects of skewness and kurtosis, this will not be an issue. The mean-variance framework does not change with different values for these properties. Hence, different values for the risk aversion parameter would not yield different Sharpe-optimal weights when investigating skewness and kurtosis even if there was a utility function. However, when investigating the measures under normality, we will investigate for which values of risk aversion ν that the two considered measures correlate the most. We will then use this value for our comparisons with real-world data, so that this parameter does not cause any distortions in the optimal weights to occur due to differences in the lower moments that we do not control for.

3.3 Data and Descriptive Statistics

The financial data to be used in the many-asset case covers the period 05/10/18–05/10/19 and was retrieved in prices from their respective exchange, from which the daily percentage returns were calculated. Ten duplicate observations were removed before any calculations were performed. All ten stocks are among the largest, by market capitalization, in the S&P 500 stock price index. Table 3.1 shows the Atkinson, the Sharpe of these stocks, the geometric mean, and the second, third and fourth central sample moments of their respective returns distribution. We see that the average mean of these stocks is 0.04% and the average standard deviation is 1.6%. These values will be used as parameters in our data generation described above. The statistics also indicate that both skewness and excess kurtosis are frequently appearing properties of financial returns data. Table 3.2 shows the correlations between returns of these assets. They also indicate that correlations between financial assets are commonly positive, at least when traded on closely-related exchanges.

	Microsoft	Amazon.Com	Apple	Alphabet A	Facebook Class A
Atkinson	0.001278	0.002404	0.001829	0.001395	0.003600
Sharpe	0.065090	0.029411	0.007869	0.013196	0.002585
Mean	0.001041	0.000641	0.000147	0.000218	0.000060
SD	0.015961	0.021759	0.018636	0.016487	0.023122
Skewness	0.053135	0.015256	-0.431841	-0.262995	-1.795713
Excess Kurtosis	2.475430	2.911917	4.873472	2.865189	19.361587
	Berkshire Hathaway 'B'	Johnson & Johnson	Visa 'A'	Exxon Mobil	Walmart
Atkinson	0.000797	0.000786	0.001046	0.000761	0.000706
Sharpe	0.012101	0.035356	0.056068	-0.021017	0.067690
Mean	0.000152	0.000413	0.000815	-0.000260	0.000833
SD	0.012529	0.011667	0.014501	0.012337	0.012281
Skewness	-0.119880	-2.612088	0.233670	0.078644	1.794183
Excess Kurtosis	4.967074	21.696050	3.408485	1.250523	12.897226

Table 3.1: Distributional Properties

Name/Ticker	MSFT	AMZN	AAPL	GOOGL	FB	BRK.B	JNJ	V	XOM	WMT
Microsoft	1.000000	0.788745	0.676747	0.741200	0.490217	0.521910	0.330011	0.812214	0.449363	0.318276
Amazon.Com	0.788745	1.000000	0.690845	0.776568	0.585452	0.478365	0.287696	0.762234	0.436079	0.284504
Apple	0.676747	0.690845	1.000000	0.646896	0.438606	0.499882	0.295527	0.662839	0.391893	0.220335
Alphabet A	0.741200	0.776568	0.646896	1.000000	0.575098	0.436895	0.263695	0.710142	0.432433	0.184821
Facebook Class A	0.490217	0.585452	0.438606	0.575098	1.000000	0.260226	0.116446	0.437402	0.268256	0.069529
Berkshire Hathaway 'B'	0.521910	0.478365	0.499882	0.436895	0.260226	1.000000	0.381379	0.548149	0.543475	0.379447
Johnson & Johnson	0.330011	0.287696	0.295527	0.263695	0.116446	0.381379	1.000000	0.309685	0.392138	0.307715
Visa 'A'	0.812214	0.762234	0.662839	0.710142	0.437402	0.548149	0.309685	1.000000	0.462881	0.251914
Exxon Mobil	0.449363	0.436079	0.391893	0.432433	0.268256	0.543475	0.392138	0.462881	1.000000	0.343625
Walmart	0.318276	0.284504	0.220335	0.184821	0.069529	0.379447	0.307715	0.251914	0.343625	1.000000

Table 3.2: Correlation Matrix

4

Results

In this section, we present the results from the Monte Carlo simulations as well as from the many-asset case.

4.1 Monte Carlo Simulations

We simulate three cases of daily financial returns data, to use in the construction of Atkinson-efficient and Sharpe-efficient portfolios under short-sale constraints and with all wealth invested:

1. Asset 1 and asset 2 generated from $N \sim (0.0004, 0.016)$.
2. Asset 1 generated from a negatively skewed distribution with mean = 0.0004 and standard deviation = 0.015, and asset 2 generated from $N \sim (0.0004, 0.016)$.
3. Asset 1 generated from a distribution with positive excess kurtosis, mean = 0.0004 and standard deviation = 0.015, asset 2 generated from $N \sim (0.0004, 0.016)$.

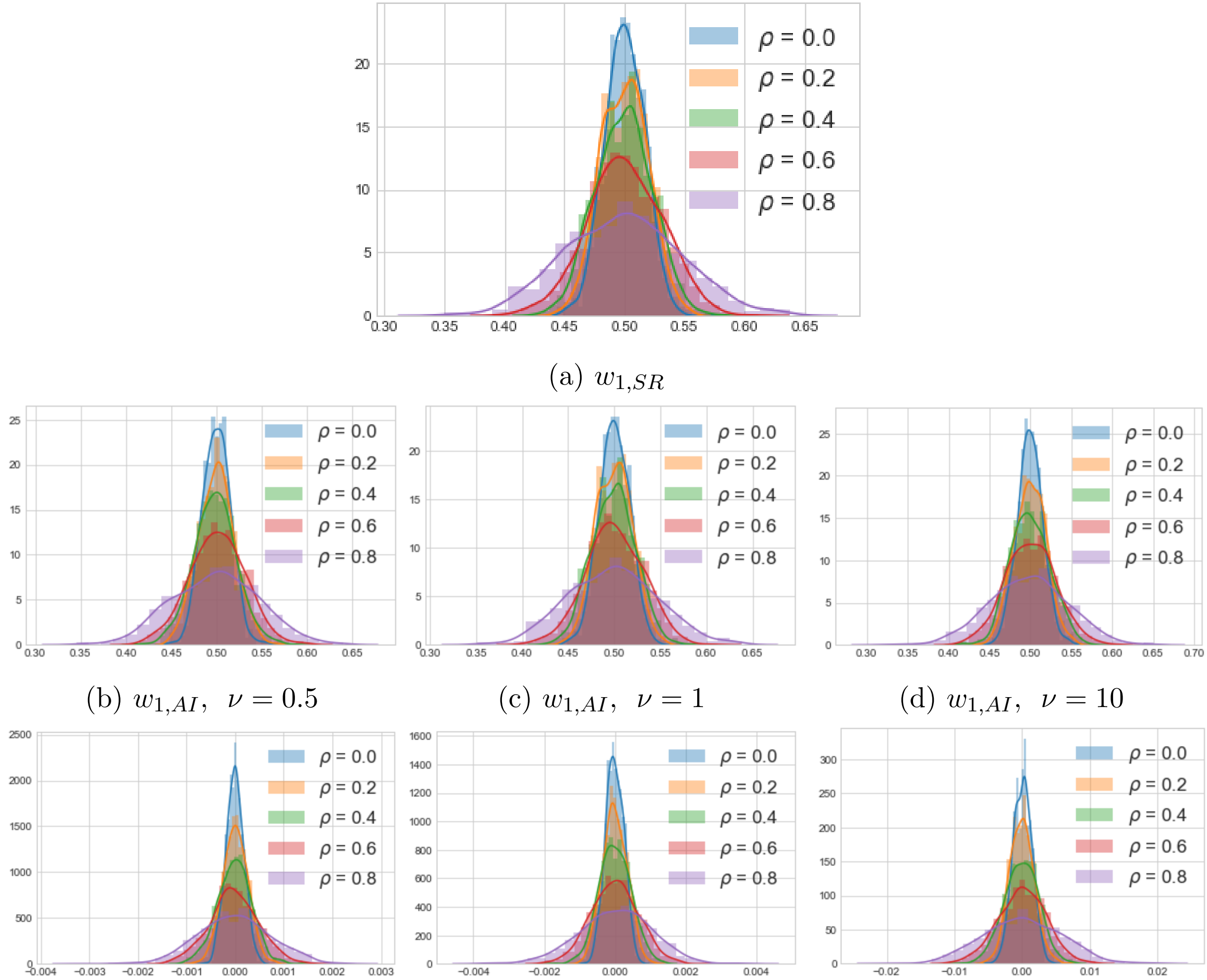
For each case we show a set of six graphs, with the graphs for the first case being supplemented by the coefficient of correlation between the weights yielded by the Sharpe optimization and the Atkinson optimization. Each set of graphs shows only weights of asset 1, from which it is easy to deduce the weight of asset 2 after realizing that their sum needs to equal 1. The first set shows the distribution of the weight assigned to asset 1 by the Sharpe (as a benchmark case) for each value of correlation ρ_r and risk aversion ν, the second set shows the distribution of the weight assigned to asset 1 by the Atkinson for each value of correlation ρ_r and risk aversion ν, and the third set shows the distribution of the difference between the weights assigned to asset 1 by the two risk measures for each value of correlation ρ_r and risk aversion ν. To minimize the amounts of graphs shown in the in this section, we choose to only show the graphs for three values of risk aversion ν: $\nu = 10$, $\nu = 30$ and $\nu = 50$. The graphs corresponding to all values are shown in the Figures section of Appendix A.

After presentation of the results, each case will start with a general discussion, then follows a discussion of the effect of correlation ρ_r, after which the effect of risk aversion ν will be discussed. Each case then ends with a conclusion.

4.1.1 Normality

We begin by running the simulation of a portfolio for $k = 2$ assets that have independently identically distributed (I.I.D.) returns generated from $N \sim (0.0004, 0.016)$. This is done since without higher moments that deviate from those of the normal distribution, we can compare the measures when both measures are able to account for all existing risk in the assets. Only small differences should occur as simulated data will never follow a distribution exactly equal to the normal for a finite set of data points ($T < \infty$). In the case that they both actually do, for example if we let T approach infinity, the weights should be exactly $(0.5, 0.5)$.

We show the weights of asset 1 for $T = 1000$ and $num\_sim = 1000$ (the number of simulations run) in figure 4.1. Figure 4.1a shows the distribution of the Sharpe-efficient weights of asset 1 under normality, which does not depend on risk aversion ν. Figures 4.1b–d show the distributions of the Atkinson-efficient weights for various values of correlation ρ_r and risk aversion ν under normality. Figures 4.1e–g show the distributions of the differences between the Atkinson-efficient and the Sharpe-efficient weights for various values of correlation ρ_r and risk aversion ν under normality. Figure 4.3 shows the mean Atkinson efficient weight under normality as a function of risk aversion ν and its 95% confidence intervals, for various values of correlation ρ_r. Table 4.1 shows the correlation



(a) $w_{1,SR}$

(b) $w_{1,AI}, \ \ \nu = 0.5$ (c) $w_{1,AI}, \ \ \nu = 1$ (d) $w_{1,AI}, \ \ \nu = 10$

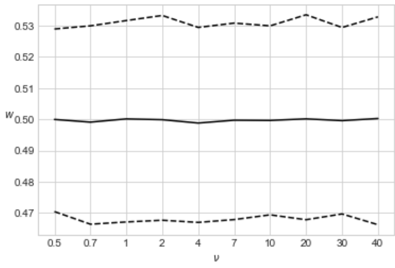
(e) $w_{1,AI} - w_{1,SR}, \ \ \nu = 0.5$ (f) $w_{1,AI} - w_{1,SR}, \ \ \nu = 1$ (g) $w_{1,AI} - w_{1,SR}, \ \ \nu = 10$

Figure 4.1: Efficient Weights under Normality

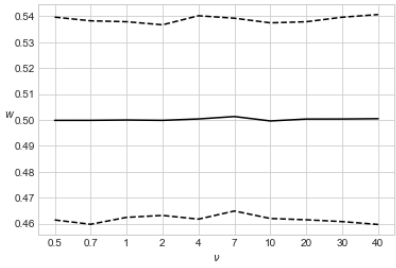
coefficient ρ_w for various values of the risk aversion parameter.

As can be seen in figure 4.1a, figures 4.1b–d and figures 4.3, the weight of asset 1 is distributed with a mean of 0.5 for both risk measures, indicating that both assets are considered approximately the same by both risk measures. The differences between the Atkinson-efficient and the Sharpe-efficient weights are small and on average approximately 0, which we see in figure 4.1e–g. We can see that for different values for correlation ρ_r, the algorithm assigns greater weights to better assets, making the distribution of weights wider.

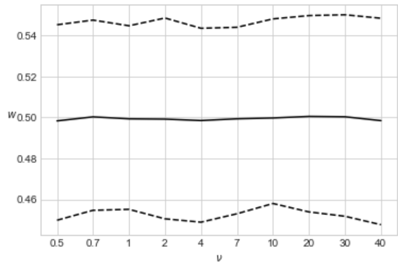
These results are unsurprising. The weight of asset 1 fluctuates around 0.5 for both measures due the two assets being simulated from the exact same true distribution, and



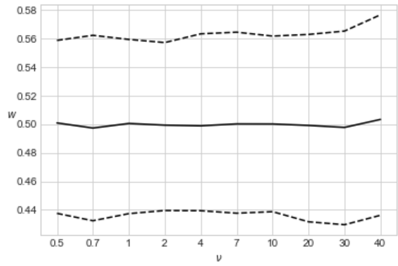
(a) $\rho_r = 0.0$



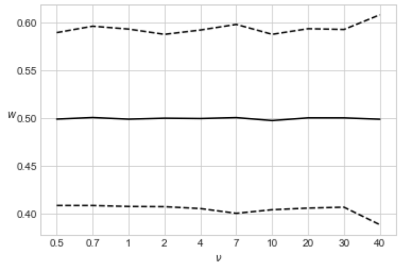
(b) $\rho_r = 0.2$



(c) $\rho_r = 0.4$



(d) $\rho_r = 0.6$



(e) $\rho_r = 0.8$

Weight of asset 1 with 95% confidence interval bounds

Figure 4.3: Mean Atkinson-Efficient Weight under Normality

have only slightly varying sample distributions in each simulation. That the Atkinson-efficient weights often differ from the Sharpe-efficient weights in spite of being calculated on the exact same assets is likely mostly due to the deviation from normal of higher moments, since only the Atkinson takes these into account.

Correlation ρ_r under Normality

Correlation ρ_r has a widening effect on all of these distributions. This is intuitive as it is less meaningful to use correlated assets for differentiation. Thus, any existing differences risk will be receive larger emphasis with more correlation. We can also see this effect in the well-known closed-form solution to Sharpe-efficient weights (see equation: 2.3). That correlation ρ_r widens the distribution of the difference between the efficient weights, as seen in figures 4.1e–g, indicates that the two frameworks consider the correlation differently. This could be due to the calculation of the correlated asset also contributing to increases in the values of higher order cross central moments (coskewness and cokurtosis) in the sample distributions, which only the Atkinson framework should account for. Further, it is clear from figures 4.1b–d and 4.3 that the mean of Atkinson-efficient weights does not depend on correlation ρ_r.

Risk Aversion ν under Normality

ν	$corr(w_{AI}, w_{SR}), \rho_w$
0.5	0.9999
0.7	0.9998
1	0.9998
2	0.9995
4	0.9986
7	0.9963
10	0.9931
20	0.9734
30	0.9464
40	0.9131

Table 4.1: Weight Correlations, ρ_w, under Normality

As risk aversion ν increases, we observe in figures 4.1b–d and 4.3 that the distribution of weights given by the Atkinson is also largely unchanged, however, correlation between the weights given by the Sharpe and by the Atkinson is decreased. We see this decrease of correlation in table 4.1, as well as in figures 4.1e–g where risk aversion ν is seen to have a widening effect on the distribution of the difference between the respective efficient weight.

It is somewhat surprising that the correlation coefficient ρ_w between the weights assigned by the two measures under normality is decreasing with risk aversion ν, indicating that the two risk measures are best comparable for lower values of risk aversion ν. It could be that the combined effect of the increase in risk aversion ν on the higher moments on the Atkinson dominates the effect on the first two moments in such a way that the Atkinson framework assigns weights that differ more from the Sharpe-efficient weights which only account for moment 1 and 2. Nevertheless, we draw the conclusion that the measures are most comparable for lower values of ν and use this result in the many-asset case with assets taken from the real world, where we want to make a fair comparison between the measures applied to real-world data.

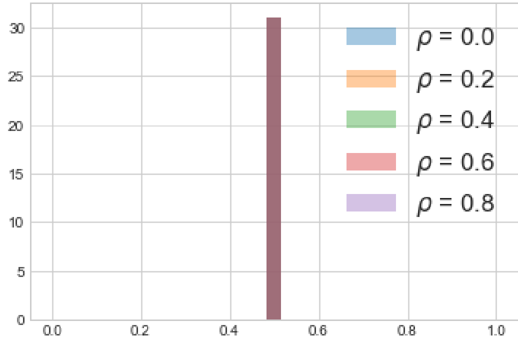
Conclusion under Normality

The conclusion of this section is that we see no great advantages to using the Atkinson instead of the Sharpe when the normality assumption of the Sharpe holds true. Both measures give weights that are approximately equally distributed, especially for lower values of ν. What could be to the advantage of the Atkinson is the additional concern with risk aversion ν, however this also seems to introduce an additional assumed parameter which is often unknown to the researcher. Increasing the value of this parameter also introduces greater distortions of the optimal weights when the higher moments deviate from the normal in the sample – finite-period – data. We consider the Sharpe-efficient portfolios preferable for three main reasons when the normality assumption holds true:

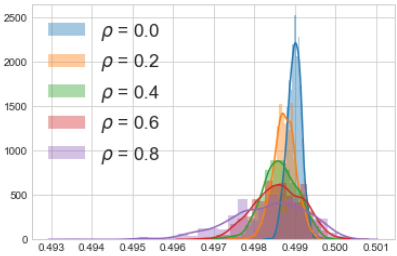
- No distortion of optimal weights due to deviating higher moments when calculated on a finite-period sample.
- Simplicity of calculation and implementation.
- Simplicity of interpretation.

4.1.2 Negative Skewness

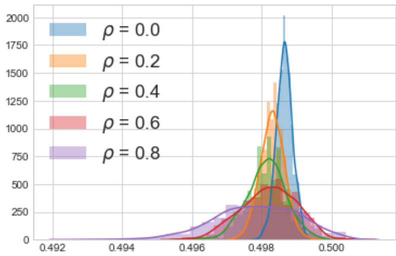
Moving on to more interesting combinations of distributions, we continue by showing the results from simulations of portfolios in which asset 1 is generated from a negatively skewed distribution ($skew_1 \in [-0.836, -0.405]$), and in which asset 2 is generated from $N \sim (0.0004, 0.016)$. The sample distributions of both assets are manipulated post-generation to have a mean of exactly 0.0004 and standard deviation of exactly 0.016. Negative skewness varies within the interval. The absolute value of kurtosis is not allowed to exceed 0.3 for either asset. The results are shown in figure 4.5. Figure 4.5a shows the distribution of the Sharpe-efficient weights of asset 1 under negative skewness, which does not depend on risk aversion ν. Figures 4.5b–d show the distributions of the Atkinson-efficient weights for various values of correlation ρ_r and risk aversion ν under negative skewness. Figures 4.5e–g show the distributions of the differences between the Atkinson-efficient and the Sharpe-efficient weights for various values of correlation ρ_r and risk



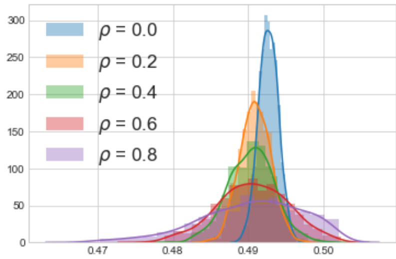
(a) $w_{1,SR}$



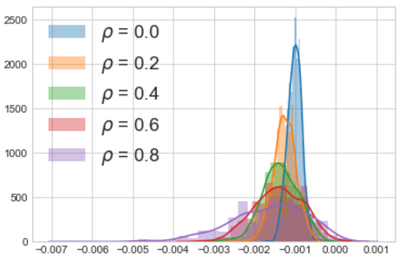
(b) $w_{1,AI}, \quad \nu = 0.5$



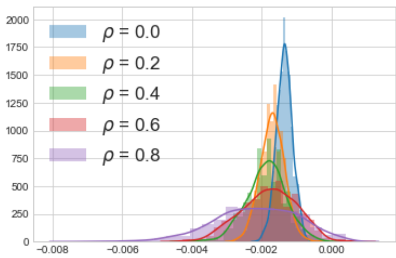
(c) $w_{1,AI}, \quad \nu = 1$



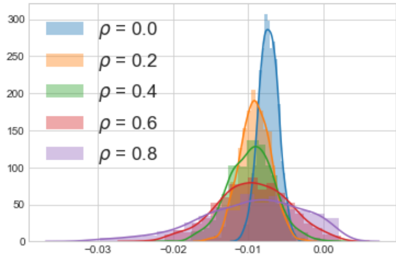
(d) $w_{1,AI}, \quad \nu = 10$



(e) $w_{1,AI} - w_{1,SR}, \quad \nu = 0.5$



(f) $w_{1,AI} - w_{1,SR}, \quad \nu = 1$

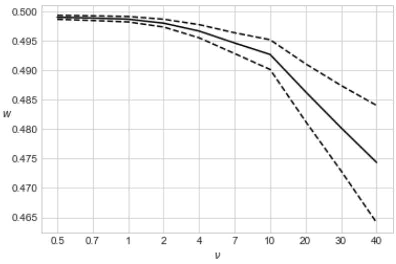


(g) $w_{1,AI} - w_{1,SR}, \quad \nu = 10$

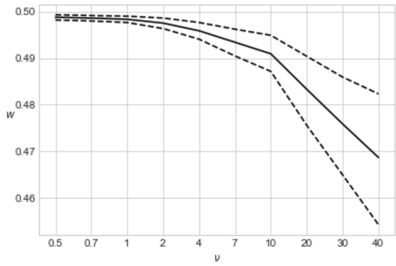
Figure 4.5: Efficient Weights under Negative Skewness

aversion ν under negative skewness. Figure 4.7 shows the mean Atkinson efficient weight under negative skewness as a function of risk aversion ν and its 95% confidence intervals, for various values of correlation ρ_r.

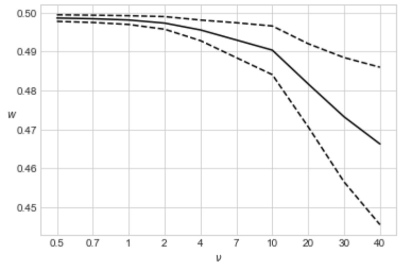
The most obvious observation to make from figure 4.5 is that there is no deviation in Sharpe-efficient weights, seen in figure 4.5a. The risk-reward is considered equal for both assets by the Sharpe since the first two moments of the sample distributions are held constant over the simulations. We then observe in figure 4.5b–d that the Atkinson-efficient weights of asset 1 fluctuate mostly below 0.5 for all values of correlation ρ_r and risk aversion ν. This is because the Atkinson also takes higher moments into account. Naturally, since the Sharpe-efficient weights are always 0.5, this also means that the



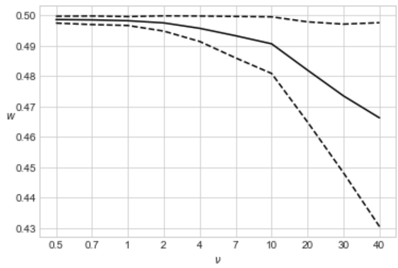
(a) $\rho_r = 0.0$



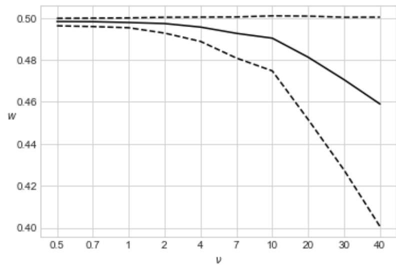
(b) $\rho_r = 0.2$



(c) $\rho_r = 0.4$



(d) $\rho_r = 0.6$



(e) $\rho_r = 0.8$

Weight of asset 1 with 95% confidence interval bounds

Figure 4.7: Mean Atkinson-Efficient Weight under Negative Skewness

difference between the Atkinson-efficient and the Sharpe-efficient weight is distributed mostly below 0 for all values of correlation ρ_r and risk aversion ν, as is seen in figures 4.5e–g. As expected by our theory, particularly obvious is the effect of negative skewness in this simulation which seems to assign lower weights to asset 1, which has considerable negative skewness. This shows that the risk, as measured by the Atkinson, is higher in asset 1 than in asset 2 when asset 1 has negative skewness. To put it simply: when the outliers are more, as well as more extreme, on the negative end of the distribution, the risk as considered by the Atkinson is increased.

Correlation ρ_r under Negative Skewness

Similarly to the case with two normally distributed assets, the weights assigned to asset 1 by the Atkinson fluctuate more for higher correlation ρ_r. For higher correlation ρ_r, we also observe some observations in which the Atkinson assigns a weight to asset 1 greater than 0.5. This unexpected result most likely occurs because of the correlation calculation, which causes the difference in negative skewness between the two assets to lessen to the extent that the combined risk in the higher moments of the distribution is sometimes greater than the risk in the negative skewness.

Risk Aversion ν under Negative Skewness

In this case, risk aversion ν plays a larger role than in the previous, as it is seen to widen the distribution on both sides of the mean, especially to the left as we can see on the lower bound of the confidence interval in figure 4.7. In this graph we can clearly also see that it simultaneously moves the mean leftwards. We see in figures 4.5b–d that for very low values of risk aversion ν, the weight fluctuations are very small, indicating that the Atkinson does not consider the risk of asset 1 to be much greater than that of asset 2 for these values. For higher values, such as for $\nu = 10$, the negative skewness of asset 1 begins to receive attention. We see also in figures 4.5b–d that for this value of risk aversion ν, asset 1 receives Atkinson-efficient weights further to the right of 0.5 than for lower values.

It is interesting that the effect of risk aversion ν is much larger in this case than under normality. It could be due to an exponential effect of risk aversion on the risk attributed, such that the larger magnitudes of the skewness of the sample distributions are considered dis-proportionally heavier under higher values of risk aversion ν. This effect would not be present to the same extent in the previous case since the sample distributions do not vary as much, not even in the first two moments.

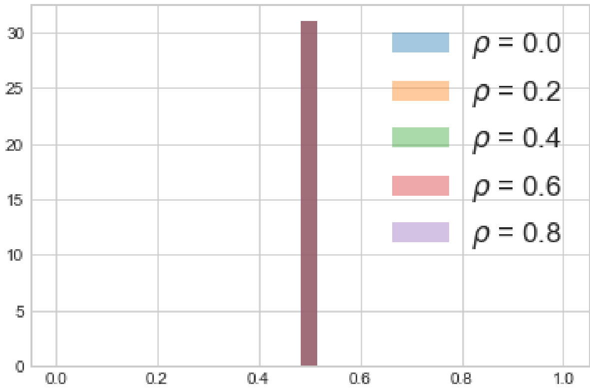
It could be questioned whether the consideration of the negative skewness is large enough for any of these values of risk aversion ν. For very low values of risk aversion ν, the skewness barely makes a difference in the weight yielded, and even for larger values, such as risk aversion $\nu = 10$, negative skewness does not yield weights much lower than 0.47 even in the most extreme cases. Here, we need to keep in mind that out of all sample distributions we generate, even the most negatively skewed has a skewness of merely -0.836, not a very extreme value. If one still considers this an issue, one could solve it by increasing risk aversion ν further, however then the risk in the lower moments may receive somewhat odd consideration in comparison to the standard mean-variance framework, as is indicated by the increasing difference in efficient weights in figure 4.1, and by the lower efficient-weights correlation ρ_w for larger values of ν seen in table 4.1. Either way, the issue of not having the skewness considered at all, as is the case in the mean-variance framework, is definitely even worse.

Conclusion under Negative Skewness

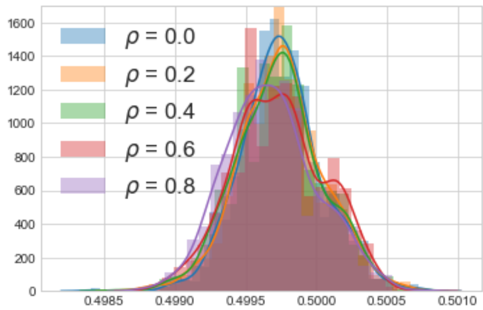
Considering both sides of these results we conclude that, under skewness, the Atkinson is preferable to the Sharpe. The results make it obvious that under skewness, the impact on the risk measured by the Atkinson is to the advantage of the investor whose portfolio is being optimized, especially when being careful with the magnitude of the risk-aversion parameter. Without this consideration for skewness, a rational investor will experience unexpectedly large drawdowns when this skewness is negative, and will not gain as much from the additional expansions when this skewness is positive. Therefore, as long as there is enough skewness, the advantages of the Atkinson outweigh the disadvantages of its interpretation being less intuitive and its computation more complex. Furthermore, considering that computers nowadays handle these computations quite easily, the skewness that is needed for it being warranted should be considered quite low for most researchers.

4.1.3 Positive Excess Kurtosis

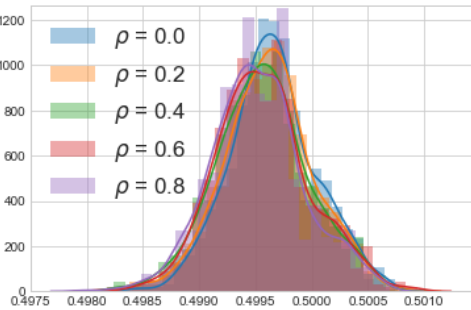
In our final simulation, we generate data to examine the effects of kurtosis on the portfolio weights. Asset 1 in this portfolio is generated from a distribution with positive excess kurtosis ($kurt_1 \in [5.21, 101]$), and asset 2 is generated from a normal distribution, $N \sim (0.0004, 0.016)$. As in the case with negative skewness, the generated data is manipulated to have a mean of exactly 0.0004 and standard deviation of exactly 0.016. Skewness is limited to ± 0.1. The results are shown in figures 4.9 and 4.11. Figure 4.9a shows the distribution of the Sharpe-efficient weights of asset 1 under positive excess kurtosis, which does not depend on risk aversion ν. Figures 4.9b–d show the distributions of the Atkinson-efficient weights for various values of correlation ρ_r and risk aversion ν under positive excess kurtosis. Figures 4.9e–g show the distributions of the differences between



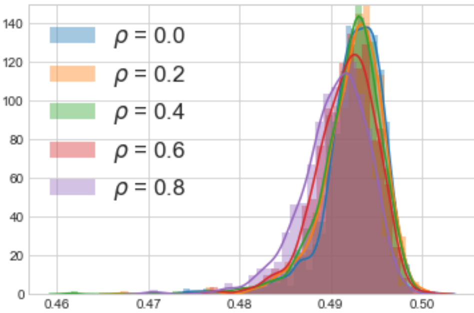
(a) $w_{1,SR}$



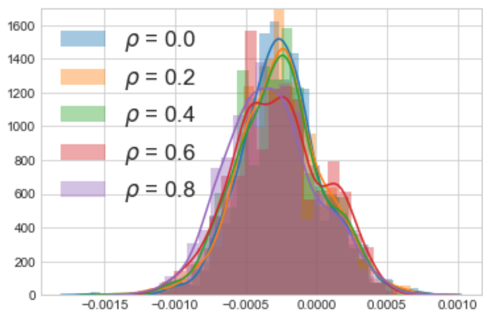
(b) $w_{1,AI}, \;\; \nu = 0.5$



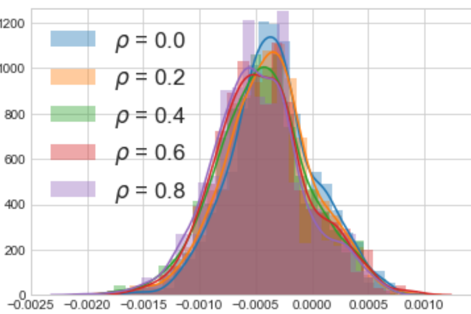
(c) $w_{1,AI}, \;\; \nu = 1$



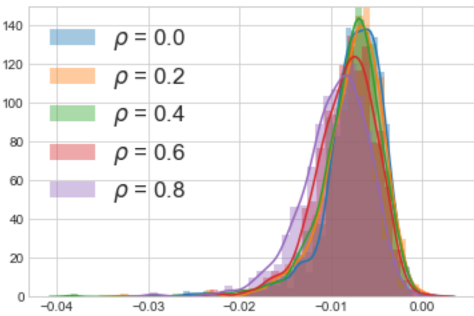
(d) $w_{1,AI}, \;\; \nu = 10$



(e) $w_{1,AI} - w_{1,SR}, \;\; \nu = 0.5$



(f) $w_{1,AI} - w_{1,SR}, \;\; \nu = 1$

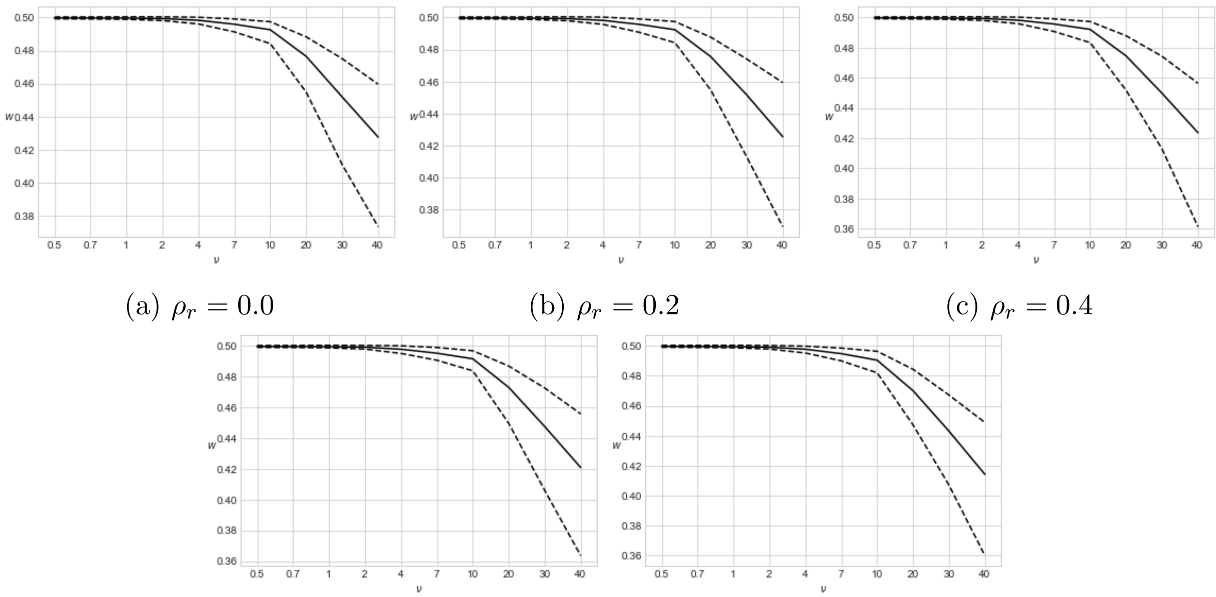


(g) $w_{1,AI} - w_{1,SR}, \;\; \nu = 10$

Figure 4.9: Efficient Weights under Positive Excess Kurtosis

the Atkinson-efficient and the Sharpe-efficient weights for various values of correlation ρ_r and risk aversion ν under positive excess kurtosis. Figure 4.11 shows the mean Atkinson efficient weight under positive excess kurtosis as a function of risk aversion ν and its 95% confidence intervals, for various values of correlation ρ_r.

Again, we see in 4.9a that the Sharpe-efficient weights do not change when the mean and variance are held constant. However, looking at 4.9b–d, we see that the Atkinson distributes efficient weights completely below 0.5 for asset 1, which has positive excess kurtosis. Naturally, since the Sharpe-efficient weights are always 0.5, this also means that



(a) $\rho_r = 0.0$ (b) $\rho_r = 0.2$ (c) $\rho_r = 0.4$

(d) $\rho_r = 0.6$ (e) $\rho_r = 0.8$

Weight of asset 1 with 95% confidence interval bounds

Figure 4.11: Mean Atkinson-Efficient Weight under Positive Excess Kurtosis

the difference between the Atkinson-efficient and the Sharpe-efficient weight is distributed completely below 0 for all values of correlation ρ_r and risk aversion ν, as is seen in figures 4.9e–g.

Correlation ρ_r under Positive Excess Kurtosis

Unlike in the case of skewness, correlation ρ_r does not show any particular effect on the distribution of the Atkinson-efficient weights, as can be seen in figures 4.9b–d. This is surprising and unexpected from the theory as we, again, can have a look at equation 2.3 and recognize that correlation ρ_r should, as is true for the cases of normality and negative skewness, widen the distributions of weights due to the decreased benefit of diversification. The reason that it does not could be due to the correlation coefficient simultaneously working in another way as well: it decreases the difference in the kurtosis between the assets, thus decreasing the difference in risk. It seems that this effect cancels the effect of decreased benefit of diversification in the case of positive excess kurtosis. This effect also exists in the case of negative skewness, where it is the negative skewness that increasingly disappears with higher correlation ρ_r, however, the effect seems to be dominated by the effect of decreased benefit of diversification, making the distribution of weights wider.

Risk Aversion ν under Positive Excess Kurtosis

From figures 4.9b–d and 4.11 we see that risk aversion ν still shows to have an effect on the distribution of the Atkinson-efficient weight of asset 1, as its mean decreases and its distribution widens. Especially, it widens on the negative end. This is similar to the case of negative skewness. This widening effect is expected due to its effect of emphasizing any risk. For risk aversion ν there is no simultaneous, counteracting effect, as there is with correlation ρ as discussed above.

Another remarkable thing about these results is that although excess kurtosis of asset 1 reaches rather extreme values, as high as 101, the weight of this asset does not reach values much lower than 0.46. This reintroduces the question of whether the risk in these higher moments is not considered enough for these values of ν. By the same reasoning as before, the solution to increase the value of risk aversion ν raises other issues and researchers should therefore be careful to increase it too much. However, that the risk is considered at all is still considered positive in our view.

Conclusion under Positive Excess Kurtosis

The result that the widening effect on the efficient-weight distribution of correlation is canceled by its effect on the difference in kurtosis indicates that the Atkinson considers positive excess kurtosis a lesser risk than negative skewness. Contrarily, directly comparing the wideness of the distributions of figures 4.5b–d and figures 4.9b–d we see that the distributions are generally slightly wider for positive excess kurtosis than they are for negative skewness. This would indicate the opposite, that positive excess kurtosis is considered a greater risk than negative skewness. We need to keep in mind that the values of excess kurtosis could generally be considered more extreme than the values of negative skewness in the sample-distributions, so therefore the results of these two sections are not entirely comparable. Therefore, we need to be humble with the uncertainty of any direct comparisons between them.

We conclude that the Atkinson is preferable to the Sharpe when constructing efficient portfolios under excess kurtosis. Similarly to the case of negative skewnesss, a value of risk aversion $\nu = 10$ seems to yield a good compromise between reasonable consideration for the lower moments and enough consideration for the kurtosis.

4.2 The Many-Asset Case

	Microsoft	Amazon.Com	Apple	Alphabet A	Facebook Class A
Atk.-opt weight	0.000000	0.000000	0.000000	0.017822	0.056408
Sha.-opt weight	0.293318	0.000000	0.000000	0.000000	0.000000
	Berkshire Hathaway 'B'	Johnson & Johnson	Visa 'A'	Exxon Mobil	Walmart
Atk.-opt weight	0.105358	0.284029	0.060750	0.157143	0.318490
Sha.-opt weight	0.000000	0.087576	0.108313	0.000000	0.510793

Table 4.2: Efficient Weights in the Many-Asset Case

Table 4.2 shows the Atkinson-efficient and the Sharpe-efficient weights. In this section, we also consider tables 3.1 and 3.2. The construction of these portfolios are made under short-sale constraints and with all wealth invested, as in the previous section. At first glance, it is interesting that many assets in the Atkinson-efficient portfolio have considerable kurtosis, even the assets given the highest Atkinson-efficient weights. This seemingly contradicts the theory and the results of the previous section that kurtosis should be a positive component to the Atkinson (meaning worse performance) and negative to component to the weight given in an Atkinson-efficient portfolio. Similarly, some assets with large Atkinson-efficient weights have considerable negative skewness, also seen to be interpreted as worse for Atkinson-measured performance in the previous section. Below we will see why this may be. To see the effects of correlation on the magnitude of the weights, we will mostly examine the correlations between assets with positive weights, rather than between all assets, as these are the most relevant for the weight of the respective asset.

We can see that the Atkinson and the Sharpe both assign greatest weight to Walmart (WMT). Looking at its Atkinson, $A_{WMT} = 0.000706$, we can see that it carries good performance in spite of its high kurtosis, in comparison to the other assets. Its Sharpe (which, of course, also affects the Atkinson) is high, $S_{WMT} = 0.0677$, and its skewness is strongly positive. Its correlation with the rest of the assets in the portfolio is relatively low as well, contributing to its role in minimizing the Atkinson of the entire portfolio. The reason for its Atkinson-efficient weight being lower than its Sharpe-efficient weight may be due to the high excess kurtosis, $kurt_{WMT} = 12.9$, although its skewness is considerably positive, $skew_{WMT} = 1.79$. Since the higher moments are what causes the difference between the optimal weights for each asset, this indicates that the negative effect of this magnitude of positive excess kurtosis dominates the positive effect of this magnitude of positive skewness. This statement, however, relies on the assumption that the magnitudes, or rather, the effect on the Atkinson, of the even higher moments (from the fifth and up) is negligible. This we do not know.

The Atkinson-efficient weight of Johnson & Johnson (JNJ), $w^A_{JNJ} = 28.4\%$, is interesting. It receives quite a large weight in spite of its considerable negative skewness, $skew_{JNJ} = -2.61$, and highly positive excess kurtosis, $kurt_{JNJ} = 21.7$. This could be due to its correlations with the rest of the assets being low and its Sharpe quite high, $S_{JNJ} = 0.0354$. However, despite these properties of its distribution, the large weight is still surprising since these factors should also be taken into account in the Sharpe-efficient weight, which is lower. This could unexpectedly indicate that these values of the third and the fourth moments are considered positive for the performance by the Atkinson. Similarly, Facebook Class A (FB) also receives a positive Atkinson-efficient weight, $w^A_{FB} = 5.64\%$, despite it having the highest Atkinson of all considered assets, $A_{FB} = 0.0036$, while receiving no weight in the Sharpe-efficient portfolio.

Considering their distributional properties, the weights of these two assets are unexpected by our theory and the results in the previous section, and their respective explanation seems to differ. By the value of the Atkinson for Johnson & Johnson (JNJ), $A_{JNJ} = 0.000786$, which is quite low in comparison to the other assets' Atkinsons, we should conclude that it is the Sharpe and moments five and up, that still gives it its considerable, positive weight. As for the Atkinson-efficient weight of Facebook Class A (FB), its high Atkinson seems in agreement with our theory and our results from the previous section, that its first four moments are considered poor for performance by the Atkinson. The higher moments do not seem to compensate in this case. It suggests that its only redeeming quality is its low correlation with the other assets in the portfolio, giving it a positive Atkinson-efficient weight. One could interpret this as correlation being more heavily disfavored by the Atkinson framework than by the mean-variance framework.

However, comparing figures 4.1a to figures 4.1b–d, we see that the optimal-weight distributions widen similarly with increasing correlation ρ_r. The distribution of weights given by the Atkinson does not widen more than the distribution given by the Sharpe. Thus, it does not seem that the above explanation is entirely accurate. A more accurate explanation could be that the correlation also transfers to the higher moments, in the form of *coskewness*, *cokurtosis*, etc, which are not major properties in the simulation under normality due to the entire lack of higher moment deviations from the normal. Although the higher order cross central moments are not computed and further investigated in this thesis, we can speculate that these are likely to be accounted for by the Atkinson framework, while they are not by the mean-variance framework. This would indirectly disfavor correlation more in the Atkinson framework than in the mean-variance framework.

This interpretation that the correlation transfers over into the higher moments is supported by comparisons of the weights of other highly correlated assets for the respective case. There seems to be a pattern that higher-correlation ρ_r assets are given a lower Atkinson-efficient weight than Sharpe-efficient weight. The heavy weight of Microsoft (MSFT) in the Sharpe-efficient portfolio, $w^S_{MSFT} = 29.3\%$, is one such asset. It is heavily correlated with the positively Sharpe-efficient weighted, $w^S_V = 10.8\%$, Visa 'A' (V) asset, $\rho_{MSFT,V} = 0.812$. We see that while given this considerable weight in the Sharpe-efficient portfolio, it is not given any weight at all in the Atkinson-efficient portfolio despite its mid-level Atkinson. There are no assets in the Atkinson-efficient portfolio given such high weights with correlations that parallel this. The highest correlation between two assets in the Atkinson efficient portfolio is that between Alphabet A (GOOGL) and Visa (V), $\rho_{GOOGL,V} = 0.71$, however their combined weight is no more than $w^A_{GOOGL} + w^A_V = 6\%$

and the Atkinsons of both are mid-level among the considered assets.

5

Conclusion

In this thesis, the topic of portfolio optimization has been studied using the Atkinson as a performance measure with CRRA utility. We have used simulated data in Monte Carlo simulations, as well as daily financial asset returns data of assets taken from American stock exchanges. In the analysis, we used the distributional properties of these assets to study their weights in Atkinson-efficient portfolios, and compared them to Sharpe-efficient weights. We asked the questions whether there are any advantages to using the Atkinson index instead of a standard performance measure, the Sharpe ratio, and whether any advantages are great enough to warrant the more complex calculations required to compute it and the greater complexity of its interpretation.

Aiming to hold all else equal in the Monte Carlo simulations, the Atkinson framework is shown to give lower efficient weights to negatively skewed distributions and to distributions that exhibit positive excess kurtosis. In the many-asset case, it becomes apparent that the considerations of this framework differ remarkably from the mean-variance framework, even using financial returns data taken from real stock exchanges. Seemingly, this is to the advantage of a rational investor maximizing expected utility. However, the magnitude of the risk-aversion parameter needs careful consideration to neither distort weights unnecessarily nor fail to consider the risk in the higher moments adequately. It should also be noted that under normality, there are no benefits to this framework, rather it may distort the weights due to sample-distribution deviations from true normality.

All things considered, the results suggest that there are advantages to using the Atkinson index when the concerned asset distributions exhibit non-normal properties. The advantages are deemed great enough that complexity of interpretation should not daunt the researcher whenever the higher moments of distributions are considerable, nor should the computational complexity have adverse effects on today's computers great enough to deter him.

Further research is needed on the magnitude of the higher moments and how they relate to the Atkinson-efficient weights, not only the sign of these moments. Additionally, it would be useful to not only investigate the higher order central moments, but also the higher order cross central moments, such as coskewness and cokurtosis. Such properties could explain some surprising results, especially in the many-asset case. Lastly, the Atkinson is not the only performance measure that considers higher moments and as such, it is necessary to compare it to portfolios constructed using other such performance measures, in order to truly determine its usefulness.

Bibliography

Atkinson, A. B. (1970). "On the Measurement of Inequality". In: *Journal of Economic Theory* 2, pp. 244–263.

Azzalini, A. (1985). "A Class of Distributions Which Includes the Normal Ones". In: *Scandinavian Journal of Statistics* 12 (2), pp. 171–178.

Baron, D. (1977). "On the Utility Theoretic Foundations of Mean-Variance Analysis". In: *The Journal of Finance* 32 (5), pp. 1683–1697.

Burr Porter, R. and Gaumnitz, J. E. (1972). "Stochastic Dominance vs. Mean Variance Portfolio Analysis: An Empirical Evaluation". In: *The American Economic Review* 62 (3), pp. 438–446.

Byrd, R. H. et al. (1995). "A Limited Memory Algorithm for Bound Constrained Optimization". In: *SIAM Journal on Scientific Computing* 16 (5), pp. 1190–1208.

Campbell, H. R. and Akhtar, S. (2000). "Conditional Skewness in Asset Pricing Tests". In: *Journal of Finance* 55 (3), pp. 1263–1296.

Chen, H.-H., Hsien, T. T., and Lin, D. K. (2011). "Optimal Mean-Variance Portfolio Selection using Cauchy-Schwarz Maximization". In: *Applied Economics* 45 (21), pp. 2795–2801.

Cowell, F. A. (2000). *Measurement of Inequality.* 1st ed. Amsterdam: Elsevier. Chap. 2, pp. 87–166.

Dennis, J. E. J. and Moré, J. J. (1977). "Quasi-Newton Methods, Motivation and Theory". In: *SIAM Review* 19 (1), pp. 46–89.

Eling, M. and Schuhmacher, F. (2011). "Sufficient conditions for expected utility to imply drawdown-based performance rankings". In: *Journal of Banking and Finance* 35 (9), pp. 2311–2318.

Fang, H. and Lai, T.-Y. (1997). "Co-Kurtosis and Capital Asset Pricing". In: *The Financial Review* 32 (2), pp. 293–307.

Fischer, T. and Lundtofte, F. (2019). *Unequal Returns: Using the Atkinson Index to Measure Financial Risk*. Tech. rep. Lund University, Department of Economics.

Fleishman, A. (1978). "A method for simulating non-normal distributions". In: *Psychometrika* 43 (4), pp. 521–532.

Fletcher, R. (2000). *Practical Methods of Optimization*. 2nd ed. New York: John Wiley & Sons.

Gini, C. (1912). *Variabilità e Mutabilità*. Bologna: Unknown.

Hespeler, F. and Shalit, H. (2016). "Mean-Extended Gini Portfolios: A 3D Efficient Frontier". In:

Jondeau, E. and Rockinger, G. M. (2005). "Conditional Asset Allocation Under Non-Normality: How Costly is the Mean-Variance Criterion". In: *EFA 2005 Moscow Meetings Paper*, pp. 29–55.

— (2006). "Optimal Portfolio Allocation under Higher Moments". In: *European Financial Management* 12 (1), pp. 29–55.

Jondeau, E. and Rockinger, M. (2005). "Conditional Asset Allocation Under Non-Normality: How Costly is the Mean-Variance Criterion?" In: EFA 2005 Moscow Meetings Paper. Moscow, Russia: EFA.

Jones, E., Oliphant, T., Peterson, P., et al. (2019). *SciPy: Open source scientific tools for Python*. [Online; accessed 2019-05-16]. URL: `http://www.scipy.org/`.

Konno, H. and Suzuki, K.-i. (1995). "A Mean-Variance-Skewness Portfolio Optimization Model". In: *Journal of the Operations Research Society of Japan* 38 (2), pp. 173–187.

Kraus, A. and Litzenberger, R. H. (1976). "Skewness Preference and the Valuation of Risk Assets". In: *The Journal of Finance* 31 (4), pp. 1085–1100.

Lai, K.-K., Yu, L., and Wang, S. (2006). “Mean-Variance-Skewness-Kurtosis-Based Portfolio Optimization”. In: First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS’06). Hanzhou, Zhejiang, China: IEEE, pp. 292–297.

Luo, H. (2011). *Generation of Non-normal Data – A Study of Fleishman’s Power Method*. Tech. rep. Uppsala University, Department of Statistics.

Markowitz, H. (1952). “Portfolio Selection”. In: *The Journal of Finance* 7 (1), pp. 77–91.

Oliphant, T. (2019). *NumPy: A guide to NumPy*. USA: Trelgol Publishing. [Online; 2019-05-16]. URL: `http://www.numpy.org/`.

Schittkowski, K. (1981). “More Test Examples for Nonlinear Programming Codes, part 1: Convergence analysis, part 2: An efficient implementation with linear least squares subproblems”. In: *Numerische Mathematik* 38 (1), pp. 83–127.

— (1987). *More Test Examples for Nonlinear Programming Codes*. 1st ed. New York: Springer-Verlag Berlin Heidelberg.

Shalit, H. and Yitzhaki, S. (1984). “Mean-Gini, Portfolio Theory, and the Pricing of Risky Assets”. In: *The Journal of Finance* 39 (5), pp. 1449–1468.

Sharpe, W. F. (1964). “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk”. In: *The Journal of Finance* 19 (3), pp. 425–442.

Stacy, J. (2008). “Multi-Dimensional Risk and Mean-Kurtosis Portfolio Optimization”. In: *Journal of Financial Management and Analysis* 21 (2).

A

Appendix

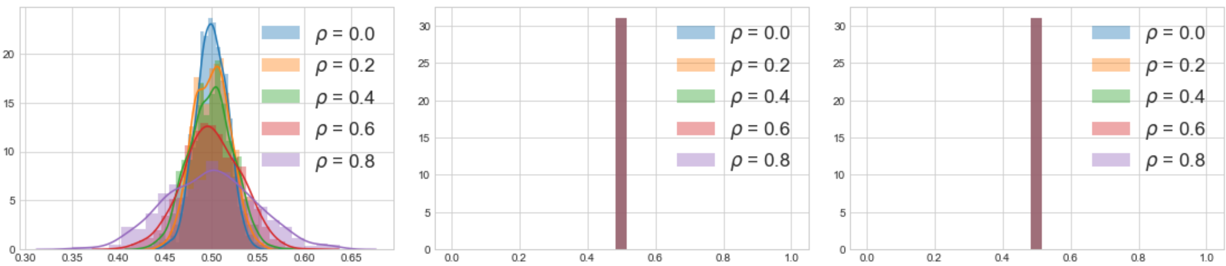
A.1 Function Shapes

Atkinson	min	0.0	0.0	0.0	0.0	0.139	0.382	0.147	0.0	0.18	0.147
	max	$2.32e-12$	$7.85e-12$	$7.65e-12$	$1.84e-13$	0.14	0.384	0.15	$5.838e-12$	0.183	0.149
Sharpe	min	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.659	0.337
	max	$4.98e-06$	$3.96e-06$	$2.22e-06$	$1.44e-06$	$4.4e-05$	$2.81e-06$	$1.12e-04$	$4.82e-06$	0.659	0.341

Table A.1: Min & Max Weights through 100 Optimizations

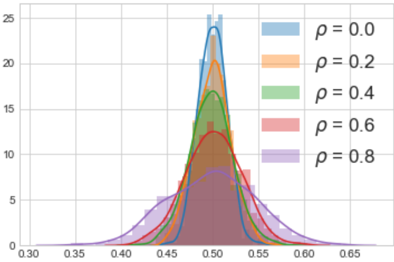
From the result that the minimum and maximum weights are essentially the same for all initial guesses, we draw the conclusion that both the Atkinson and the Sharpe ratio, at least when evaluated within the domain defined by the constraints, are convex with respect to the weights vector w. In other words, we assume that there is a unique local minimum in both functions under these constraints, implying that the L-BFGS-B algorithm is sufficient for our analysis.

Figures

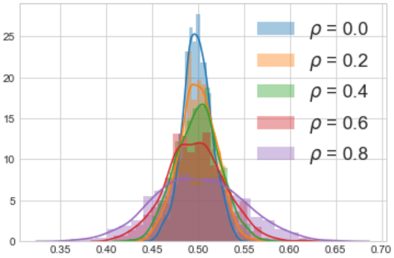


(a) $w_{1,SR}$ (b) $w_{1,SR}$ (c) $w_{1,SR}$

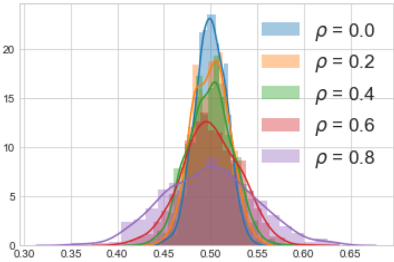
Figure A.1: Sharpe-Efficient Weights



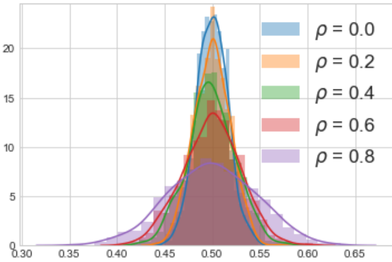
(a) $w_{1,AI}, \;\; \nu = 0.5$



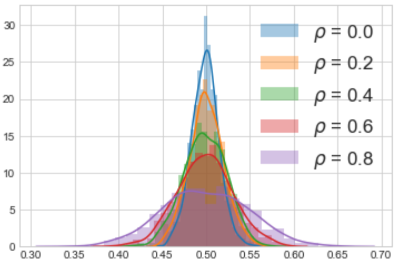
(b) $w_{1,AI}, \;\; \nu = 0.7$



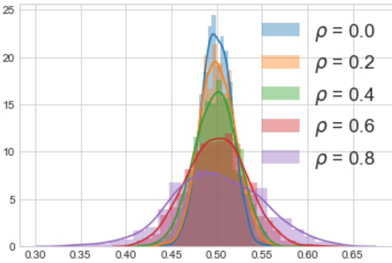
(c) $w_{1,AI}, \;\; \nu = 1$



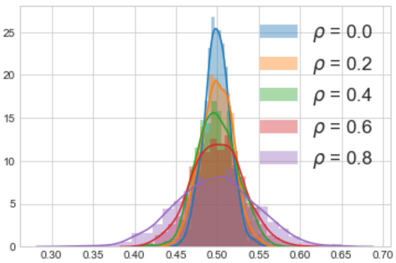
(d) $w_{1,AI}, \;\; \nu = 2$



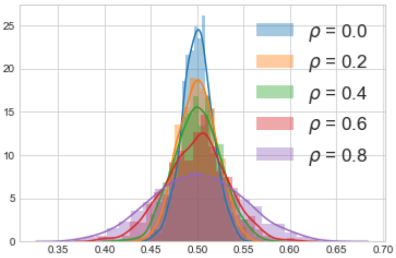
(e) $w_{1,AI}, \;\; \nu = 4$



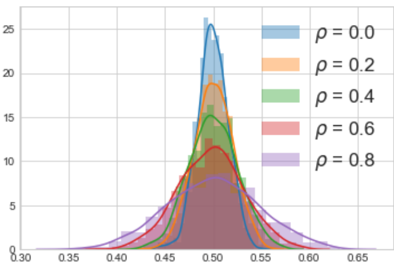
(f) $w_{1,AI}, \;\; \nu = 7$



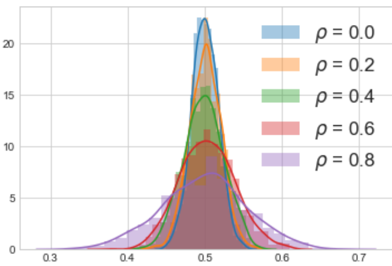
(g) $w_{1,AI}, \;\; \nu = 10$



(h) $w_{1,AI}, \;\; \nu = 20$

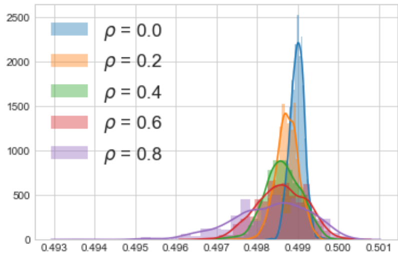


(i) $w_{1,AI}, \;\; \nu = 30$

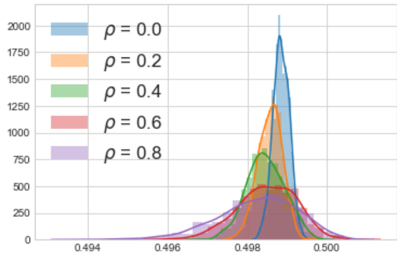


(j) $w_{1,AI}, \;\; \nu = 40$

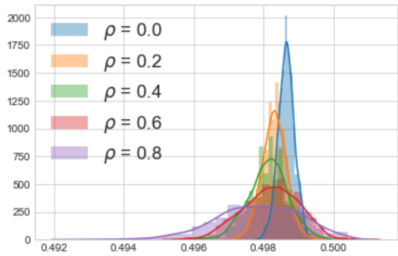
Figure A.3: Atkinson-Efficient Weights under Normality



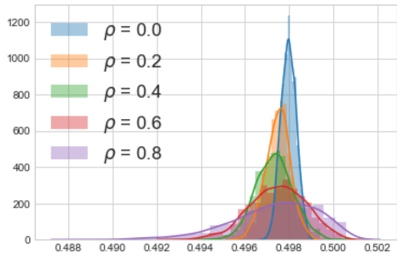
(a) $w_{1,AI}$, $\nu = 0.5$



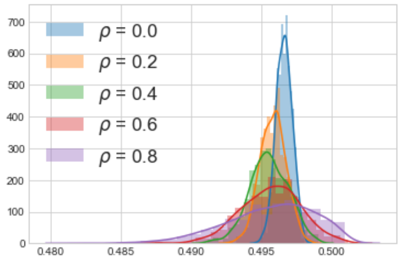
(b) $w_{1,AI}$, $\nu = 0.7$



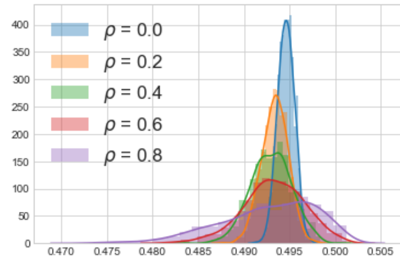
(c) $w_{1,AI}$, $\nu = 1$



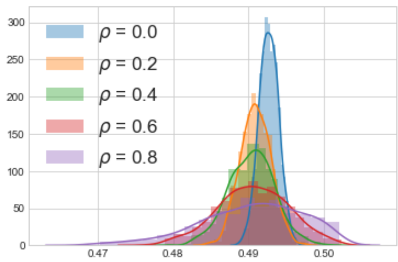
(d) $w_{1,AI}$, $\nu = 2$



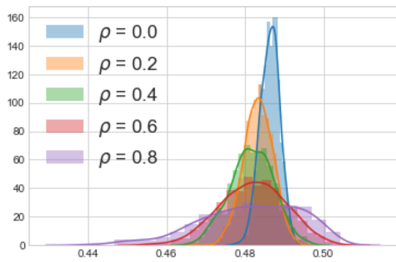
(e) $w_{1,AI}$, $\nu = 4$



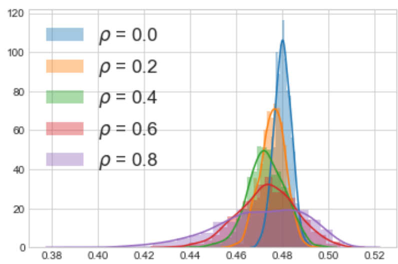
(f) $w_{1,AI}$, $\nu = 7$



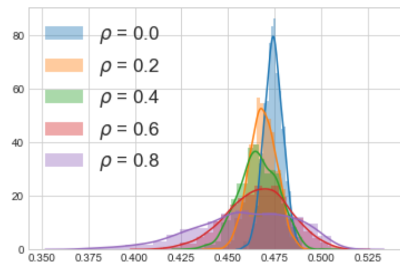
(g) $w_{1,AI}$, $\nu = 10$



(h) $w_{1,AI}$, $\nu = 20$

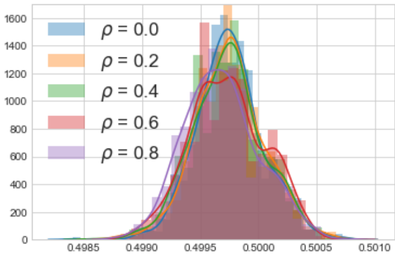


(i) $w_{1,AI}$, $\nu = 30$

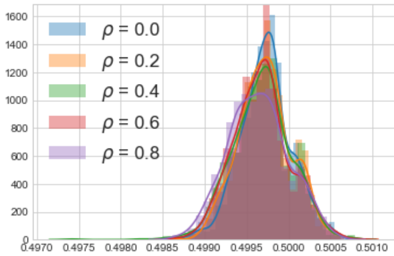


(j) $w_{1,AI}$, $\nu = 40$

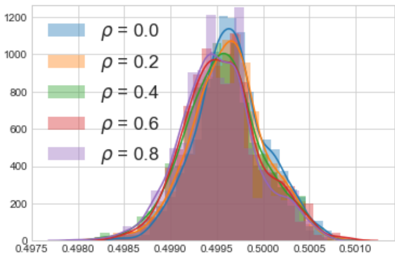
Figure A.5: Atkinson-Efficient Weights under Negative Skewness



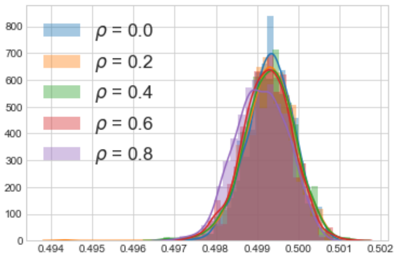
(a) $w_{1,AI}$, $\nu = 0.5$



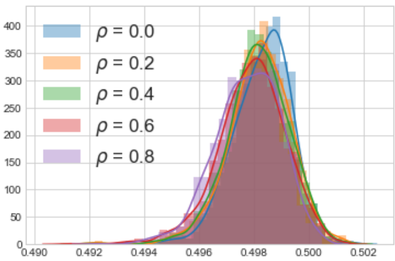
(b) $w_{1,AI}$, $\nu = 0.7$



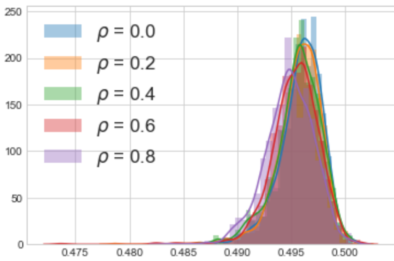
(c) $w_{1,AI}$, $\nu = 1$



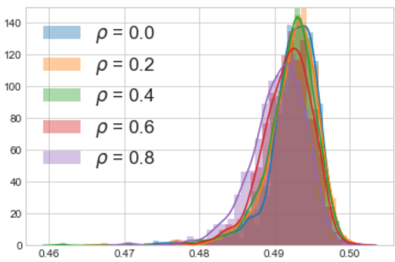
(d) $w_{1,AI}$, $\nu = 2$



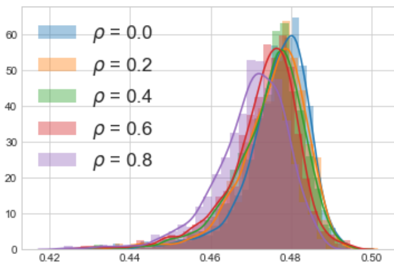
(e) $w_{1,AI}$, $\nu = 4$



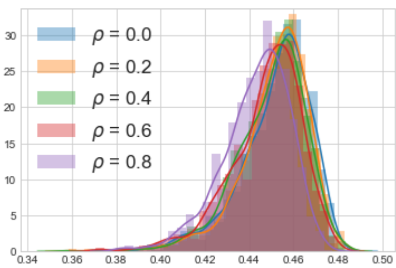
(f) $w_{1,AI}$, $\nu = 7$



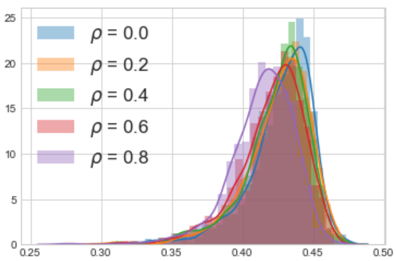
(g) $w_{1,AI}$, $\nu = 10$



(h) $w_{1,AI}$, $\nu = 20$

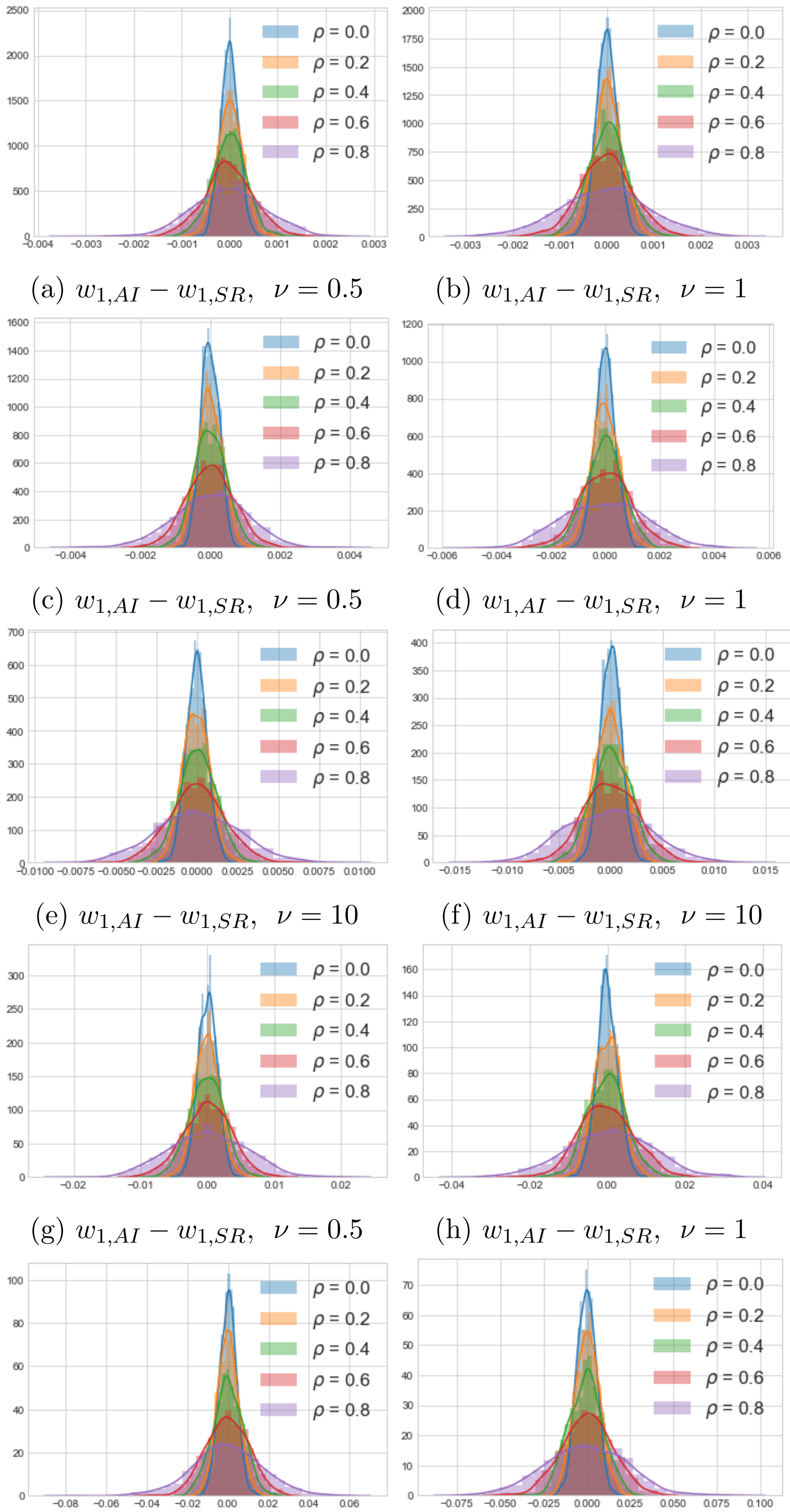


(i) $w_{1,AI}$, $\nu = 30$



(j) $w_{1,AI}$, $\nu = 40$

Figure A.7: Atkinson-Efficient Weights under Positive Excess Kurtosis



(a) $w_{1,AI} - w_{1,SR}, \;\; \nu = 0.5$ (b) $w_{1,AI} - w_{1,SR}, \;\; \nu = 1$

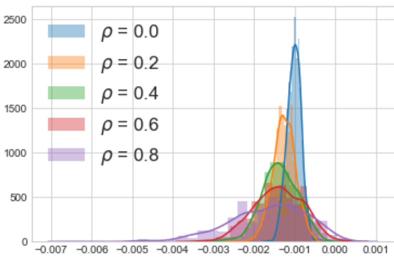
(c) $w_{1,AI} - w_{1,SR}, \;\; \nu = 0.5$ (d) $w_{1,AI} - w_{1,SR}, \;\; \nu = 1$

(e) $w_{1,AI} - w_{1,SR}, \;\; \nu = 10$ (f) $w_{1,AI} - w_{1,SR}, \;\; \nu = 10$

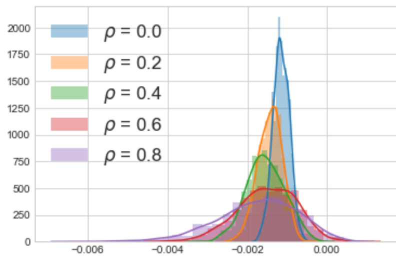
(g) $w_{1,AI} - w_{1,SR}, \;\; \nu = 0.5$ (h) $w_{1,AI} - w_{1,SR}, \;\; \nu = 1$

(i) $w_{1,AI} - w_{1,SR}, \;\; \nu = 0.5$ (j) $w_{1,AI} - w_{1,SR}, \;\; \nu = 1$

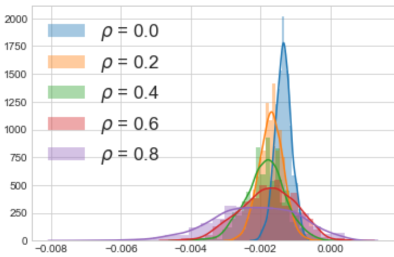
Figure A.9: Efficient Weight Differences under Normality



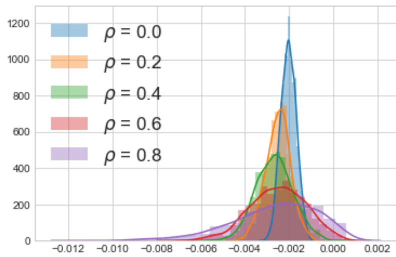
(a) $w_{1,AI} - w_{1,SR}, \ \ \nu = 0.5$



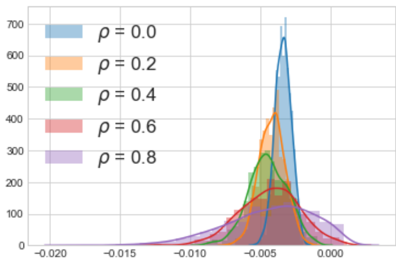
(b) $w_{1,AI} - w_{1,SR}, \ \ \nu = 1$



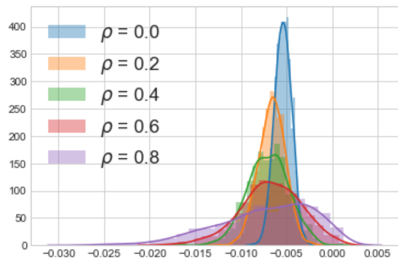
(c) $w_{1,AI} - w_{1,SR}, \ \ \nu = 0.5$



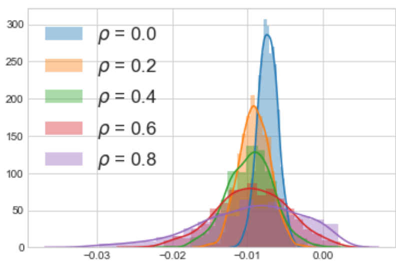
(d) $w_{1,AI} - w_{1,SR}, \ \ \nu = 1$



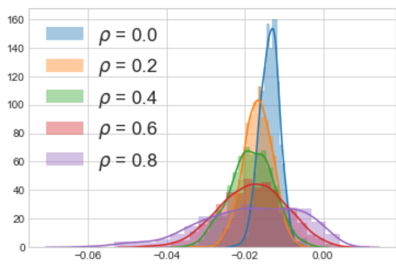
(e) $w_{1,AI} - w_{1,SR}, \ \ \nu = 10$



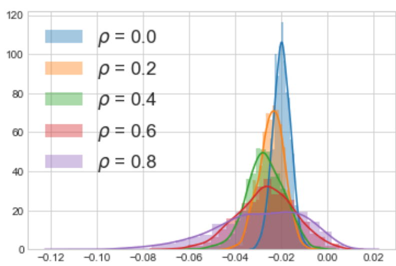
(f) $w_{1,AI} - w_{1,SR}, \ \ \nu = 10$



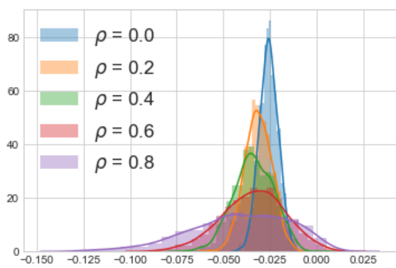
(g) $w_{1,AI} - w_{1,SR}, \ \ \nu = 0.5$



(h) $w_{1,AI} - w_{1,SR}, \ \ \nu = 1$

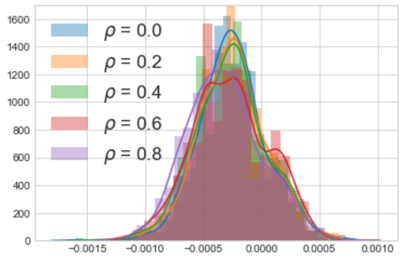


(i) $w_{1,AI} - w_{1,SR}, \ \ \nu = 0.5$

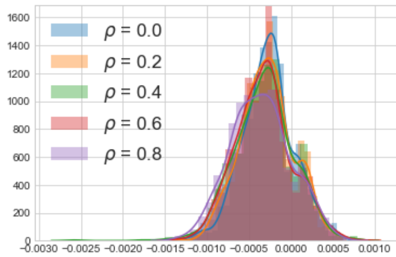


(j) $w_{1,AI} - w_{1,SR}, \ \ \nu = 1$

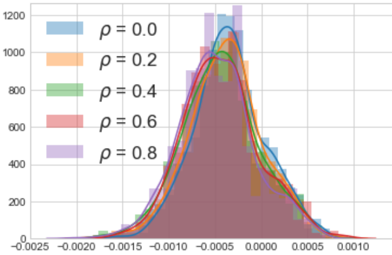
Figure A.11: Efficient Weight Differences under Negative Skewness



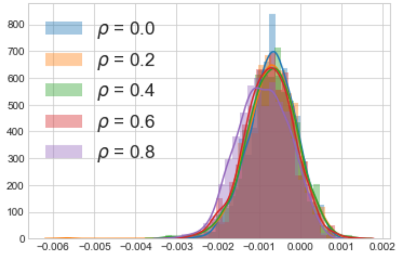
(a) $w_{1,AI} - w_{1,SR}, \;\; \nu = 0.5$



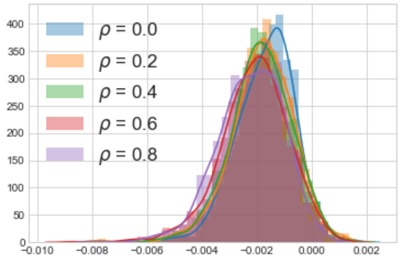
(b) $w_{1,AI} - w_{1,SR}, \;\; \nu = 1$



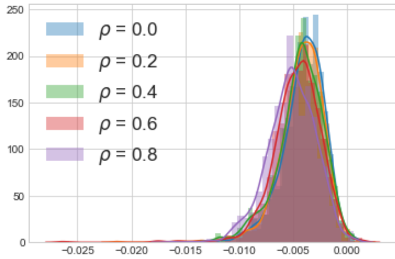
(c) $w_{1,AI} - w_{1,SR}, \;\; \nu = 0.5$



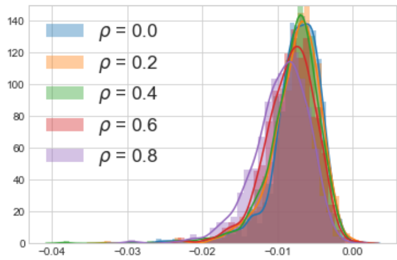
(d) $w_{1,AI} - w_{1,SR}, \;\; \nu = 1$



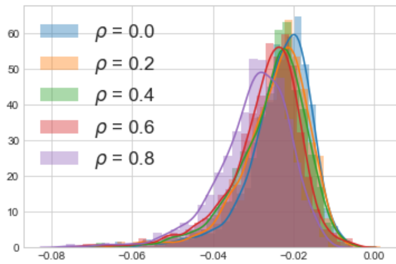
(e) $w_{1,AI} - w_{1,SR}, \;\; \nu = 10$



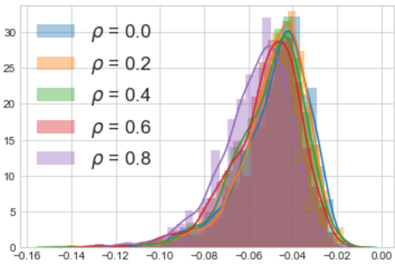
(f) $w_{1,AI} - w_{1,SR}, \;\; \nu = 10$



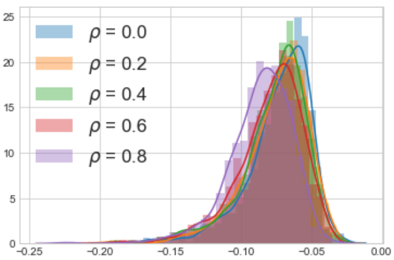
(g) $w_{1,AI} - w_{1,SR}, \;\; \nu = 0.5$



(h) $w_{1,AI} - w_{1,SR}, \;\; \nu = 1$



(i) $w_{1,AI} - w_{1,SR}, \;\; \nu = 0.5$



(j) $w_{1,AI} - w_{1,SR}, \;\; \nu = 1$

Figure A.13: Efficient Weight Differences under Positive Excess Kurtosis