

Fischer, Thomas; Lundtofte, Frederik

**Working Paper**

## Unequal Returns: Using the Atkinson Index to Measure Financial Risk

Working Paper, No. 2018:25

**Provided in Cooperation with:**

Department of Economics, School of Economics and Management, Lund University

*Suggested Citation:* Fischer, Thomas; Lundtofte, Frederik (2018) : Unequal Returns: Using the Atkinson Index to Measure Financial Risk, Working Paper, No. 2018:25, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:

<https://hdl.handle.net/10419/260254>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

Working Paper 2018:25

Department of Economics  
School of Economics and Management

# Unequal Returns: Using the Atkinson Index to Measure Financial Risk

Frederik Lundtofte  
Thomas Fischer

October 2018



**LUND**  
UNIVERSITY

# Unequal Returns: Using the Atkinson Index to Measure Financial Risk\*

Thomas Fischer<sup>†</sup>      Frederik Lundtofte<sup>‡</sup>

October, 2018

## Abstract

We apply the Atkinson (1970) inequality index to time series of asset returns to offer a novel measure of financial risk consistent with expected-utility theory. This measure is converted to a certainty-equivalent return serving as a performance measure. We extend the Atkinson index to HARA utility and derive closed-form solutions to our measures for a number of preference-return combinations. Further, we establish relationships between risk aversion and the weights assigned to the cumulants of the return distribution for our performance measure. Using data from hedge funds and asset-pricing anomalies, we find that our performance measure contains additional, economically meaningful information.

*JEL classification: G11*

*Keywords: risk, performance, non-Gaussian distributions, cumulants, hedge funds*

---

\*We would like to thank Hossein Asgharian, Charlotte Christiansen, David Feldman, Björn Hansson, Alexander Herbertsson, Håkan Jankensgård, Eva Liljeblom, Martin Strieborny, Anders Vilhelmsson as well as participants at the CFF-KWC workshop and the 2018 Asian Meeting of the Econometric Society (Seoul) for helpful comments. Nils Dominguez Berndtsson provided excellent research assistance. Financial support from the Thule Foundation (Skandia) is gratefully acknowledged.

<sup>†</sup>Fischer: Department of Economics and Knut Wicksell Centre for Financial Studies, Lund University, thomas.fischer@nek.lu.se

<sup>‡</sup>Lundtofte: Department of Economics and Knut Wicksell Centre for Financial Studies, Lund University, frederik.lundtofte@nek.lu.se

# 1. Introduction

The most commonly-used performance measure for financial assets – well-known to basically any student of Finance – is the so-called Sharpe ratio (Sharpe, 1966, 1994), which is computed as the expected excess return divided by the standard deviation of returns. It is based on the mean-variance framework and is commonly used to rank available investment opportunities. The measure is consistent with maximization of expected utility if and only if the expected utility of the optimal mix of the risky and the risk-free asset is a monotonic function of the Sharpe ratio. This would be the case if, for example, returns are normally distributed or if the utility function is quadratic.<sup>1</sup>

The above-mentioned assumptions – normally distributed returns and quadratic utility – are very strong and certainly unrealistic. Financial returns are commonly found not to follow normal distributions (Cont, 2001) but exhibit negative skewness (downside risk) and excess kurtosis (fat tails). However, the Sharpe ratio does not account for independent higher-order risk and can therefore be manipulated using option-type strategies (Goetzmann et al., 2002, 2007). Secondly, the case of quadratic utility implies a satiation point. Once the input variable of wealth exceeds this threshold value, marginal utility turns negative. It is also quite easy to come up with examples where an asset yields a lower Sharpe ratio, but is preferred according to first-order stochastic dominance (Hodges, 1998) and thus, in such examples, the ranking according to the Sharpe ratio violates any ranking based on expected utility with elementary utility functions that are non-decreasing in wealth.

---

<sup>1</sup>A common misunderstanding is that, since Chamberlain (1983) shows that spherically distributed returns imply mean-variance preferences and Owen and Rabinovitch (1983) show that CAPM can be extended to elliptical distributions, the Sharpe ratio is also a valid performance measure if returns are elliptically or spherically distributed. However, Smetters and Zhang (2013) provide a counterexample in which a spherically distributed asset yields a higher Sharpe ratio than another spherically distributed asset but a lower expected utility. Since Smetters and Zhang (2013) show that it is not valid for spherical distributions and the spherical distribution is a special case of the elliptical one, we can conclude that it is not valid for elliptical distributions either.

Most of the early critics of the Sharpe ratio use formal arguments based on the inconsistency with maximization of expected utility in general or provide intuitive counterexamples in which it fails. Using the standard deviation or variance as a measure of risk has also received a substantial amount of criticism, not least because it fails to distinguish between good (“upside”) risk and bad (“downside”) risk. This calls for alternative, more sophisticated risk and performance measures.

To our knowledge, this paper is the first to apply the Atkinson (1970) index – well-known within the literature of social inequality – to financial returns, thereby obtaining a measure of financial risk that can easily be transformed into a performance measure, which is consistent with maximization of expected utility. As such, it is general in the following sense: If a return distribution  $n$ th-order stochastically dominates another return distribution, then—provided that the elementary utility function is of e.g. CRRA or CARA type—our performance measure will show a higher value.<sup>2</sup> Since Smetters and Zhang (2013) show the impossibility of preference-free higher-order performance measures, our measure is inevitably preference-dependent, but the conclusions can be made more generally valid by varying e.g. the coefficient of relative risk aversion of a CRRA utility function within some reasonable range (say, between one and ten), and checking whether some assets are superior to others for the whole range.<sup>3</sup> Using cumulants, we derive general expressions for our risk and performance measures. We also extend the Atkinson (1970) index to the general class of HARA utility. In addition, we provide analytical solutions for a number of combinations of preferences and return distributions, including the Normal Inverse Gaussian distribution first proposed by Barndorff-Nielsen (1997a). We later apply our performance measure to hedge fund data and show that it has a low rank correlation with existing ones, indicating that it contains additional,

---

<sup>2</sup>Suppose a return distribution  $n$ th order stochastically dominates another return distribution, then—provided that  $(-1)^k u^{(k)}(x) < 0$  for  $k = 1, 2, \dots, n$  and all  $x$ — we show that our performance measure will attain a higher value.

<sup>3</sup>Of course, the conclusions will have to rely on numerics with a reasonable division of the interval of plausible values on the relative risk aversion (a non-integer valued step size is probably preferred).

economically meaningful information. Moreover, we provide an in-depth examination of well-known market anomalies and show that they lose their glamour once considered under a more sophisticated performance measure.

The idea of applying measures of social inequality to financial returns is not new. Yitzhaki (1982) was the first to apply the Gini coefficient (Gini, 1912) to financial returns in order to measure risk and he uses the Gini mean difference to rank assets. He shows that, for non-crossing generalized Lorenz curves, the Gini mean difference is consistent with second-order stochastic dominance. There are, however, many cases in which the generalized Lorenz curves cross. For instance, in the example provided by Goetzmann et al. (2002) to illustrate the manipulability of the Sharpe ratio, we show that the generalized Lorenz curves actually intersect, and that the Gini-based approach suggests that the manipulated is the superior asset, while our performance measure identifies the non-manipulated asset as the superior one for reasonable levels of risk aversion.

Given its generality, one might ask why we do not use a stochastic dominance approach, which seems simple at first sight, especially considering the power of modern computers. Using the results in Eeckhoudt et al. (2009), which in turn build on those of Ekern (1980) and Eeckhoudt and Schlesinger (2006), one could – with the help of a computer – check for the lowest order of stochastic dominance by which a specific asset dominates or is dominated by another one, and this ordering would then be consistent with maximization of expected utility for all elementary utility functions whose derivatives have alternating signs (starting with a positive first-order derivative). The main difficulty with this approach – although it is very general – is that the number of asset pairs that one needs to compare grows very quickly with the number of assets. With the help of an index, we can easily compute its value for each asset and then compare the values of all assets in our investment universe. Moreover, assuming e.g. CRRA or CARA utility, our performance index will be in line with  $n$ th-order stochastic dominance, as explained above.

Goetzmann et al. (2007) and Zakamouline and Koekebakker (2009) propose alternative performance measures that overcome some well-known limitations of the Sharpe ratio. In fact, we show that the measure proposed in Goetzmann et al. (2007) is a nested special case of our approach for Constant Relative Risk Aversion preferences, and we show how their measure can be decomposed into one part related to expected returns and one part related to risk in the form of the Atkinson index. Thus, we both extend their work and show how it is related to our framework. In Zakamouline and Koekebakker (2009), they do not account for higher moments than skewness. Our measures take kurtosis and even higher moments into account. In the bulk of their empirical work, Zakamouline and Koekebakker (2009) assume a parametric distribution, whereas our measures can be easily calculated without specifying a specific parametric distribution. Another contribution relative to Goetzmann et al. (2007) and Zakamouline and Koekebakker (2009) is that we show how our measures are related to the cumulants of the return distribution.

Martin (2013a,b) and Lundtofte and Wilhelmsson (2013) show the advantages of using cumulants in consumption-based asset pricing. Previous attempts to incorporate higher-order moments build on truncating Taylor expansions (Kraus and Litzenberger, 1976), but it is vulnerable to the critique of Brockett and Kahane (1992) showing serious flaws in using this common approach and interpreting e.g. a positive third-order derivative as a preference for skewness. In this paper, we use a similar cumulant-based approach, but instead of applying it to consumption-based asset pricing, we use it to develop general formulae for the Atkinson index and certainty-equivalent returns.

Smetters and Zhang (2013) provide a generalization of the Sharpe ratio for a wider admissible preference-probability space and in the process, they prove an impossibility theorem stating that "any ranking measure that is valid at non-Normal 'higher moments' cannot generically be free from investor preferences" (Smetters and Zhang, 2013, p. 24). Naturally, their impossibility theorem is reflected in our results. However, our goal is not

to generalize the Sharpe ratio but to explore the usefulness of the Atkinson index and rank fund returns directly – not optimal combinations between those returns and the risk free asset. We note that Smetters and Zhang’s (2013) impossibility theorem applies not only to our paper but also to Goetzmann et al. (2007) and Zakamouline and Koekebakker (2009). While Goetzmann et al. (2007) and Zakamouline and Koekebakker (2009)—implicitly or explicitly—assume specific values on preference parameters, we explore the robustness of rankings to variation of preference parameters within a reasonable range.

We employ our measure to rank assets in two empirical settings. We consider both a large number of hedge fund strategies and well-known market anomalies (size, value, and momentum). Our novel measure suggests a substantially different ranking as compared to standard measures such as the Sharpe measures and more elaborate measures based on tail risk. Some of the well-known market anomalies lose their glamour when evaluated under our novel performance measure by loading on higher order risk. For high levels of risk aversion the rank ordering of assets changes substantially, allowing for clear statistical discrimination between the assets in settings with measurement error. These results can be compared to the empirical findings in Eling and Schuhmacher (2007) showing that the ranking of hedge funds is very similar for the various performance measures that they consider.

There is a vast literature proposing alternative performance measures (Caporin et al., 2014) in particular focusing on downside risk (e.g., the popular Value at Risk). As mentioned above, Eling and Schuhmacher (2007) put those measures to the test using empirical evidence from hedge fund returns. They, however, find that the ranking of hedge funds does not change substantially for different performance measures. More formally, in Schuhmacher and Eling (2011) they show that if the distribution of returns satisfies the location-scale (LS) property<sup>4</sup> – which holds for a large number of return distributions – a large number of (drawdown) performance measures increase in a concave

---

<sup>4</sup>That implies a Probability Density Function  $f(y)$  (PDF) with the property  $f(a + by) = bf(y)$ .



manner with the Sharpe ratio. Thus, one might consider the diverse failures of the Sharpe ratio as academic nitpicking, which, yet, have little practical relevance. Our paper challenges this view, in that we find low rank correlations between our proposed performance measures and existing ones, including the Sharpe ratio. While most of the other approaches simply replace variance or standard deviation in the Sharpe ratio by other risk measures, our proposed class of performance measures is firmly grounded in the theoretical foundations of financial economics.

The remainder of the paper is organized as follows. Section 2 contains background information regarding inequality measures; Section 3 presents closed-form solutions for the Atkinson index under CRRA utility; Section 4 extends the results to Hyperbolic Absolute Risk Aversion (HARA) utility and introduces a class of performance measures; in Section 5, we apply our risk and performance measures to hedge funds and market anomalies; finally, Section 6 concludes the paper.

## **2. Background**

Since the theory behind applying inequality measures to financial data is not widely-known among researchers in Finance, this section provides a review of the literature on the financial applications of Lorenz curves and the Gini coefficient and, finally, we provide a general discussion of the Atkinson index and how it can be applied to financial data.

### **2.1. Inequality measures applied to financial time series**

In financial economics the standard deviation has been widely acknowledged as a measure of dispersion and thus of risk. Meanwhile—and seemingly unrelated—there has been a literature in the economics of inequality trying to find a simple measure of the dispersion of economic quantities in a cross-section of individuals that is easy to interpret. The

most popular measure is certainly the so-called Gini coefficient (Gini, 1912). Yitzhaki (1982) was the first to propose its usage for the time series of financial asset returns for the purpose of ranking assets. The Gini coefficient can be defined as

$$G = \frac{E[|x_1 - x_2|]}{2E[x]}, \quad (1)$$

where  $x_1$  and  $x_2$  are independent replicates of the random variable  $x$ . Provided that  $x \geq 0$ , the Gini coefficient is bounded to be between zero and one, i.e.,  $0 \leq G \leq 1$ . High values correspond to high dispersion.

Yitzhaki (1982) proposes the Gini as a measure of financial risk (high values implying high risk) and he introduces the *Gini mean difference* to rank assets. His measure of fund performance is given by

$$\lambda_i = E[x_i] - \Gamma_i,$$

for some fund  $i$  with a mean of  $E[x_i]$ . The measure  $\Gamma_i$  – labeled *Gini's mean difference*<sup>5</sup> – is defined as follows

$$\Gamma_i = 0.5 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x - x'| dF_i(x) dF_i(x') = \int_{-\infty}^{\infty} F_i(x)[1 - F_i(x)] dx,$$

where  $F_i$  is the cumulative distribution function of fund  $i$ 's returns. In fact, this measure is just the Gini coefficient multiplied by its mean  $\Gamma_i = E[x_i]G_i$ . Thus, the performance measure reads

$$\lambda_i = E[x_i] - E[x_i]G_i = E[x_i](1 - G_i). \quad (2)$$

Given the range of the Gini coefficient ( $0 \leq G_i \leq 1$ ) this is an inequality adjusted mean for which  $\lambda_i \leq E[x_i]$ . Yet, this measure is also problematic as it does not generally allow for a ranking of assets that is consistent with second-order stochastic dominance. As

---

<sup>5</sup>Sometimes  $2\Gamma_i$  is referred to as Gini's mean difference (Yitzhaki and Schechtman, 2013, Ch. 2).

explained below, the latter requires a second assumption about non-crossing generalized Lorenz curves (Lorenz, 1905).

An asset  $A$  second-order stochastically dominates an asset  $B$  if their cumulative distribution functions,  $F_A$  and  $F_B$ , respectively, satisfy

$$\int_{-\infty}^{x(q)} (F_B(t) - F_A(t)) dt \geq 0 \quad \forall 0 \leq q \leq 1, \quad (3)$$

with strict inequality for some  $q$ . Formally, this can be checked by employing the concept of the Lorenz curve (Lorenz, 1905), defined as:

$$L(F) = \frac{\int_{-\infty}^{F^{-1}(x)} x f(x) dx}{\int_{-\infty}^{\infty} x f(x) dx} = \frac{\int_0^F F'^{(-1)}(x) dF'}{\int_0^1 F'^{(-1)}(x) dF'}$$

which due to the definition of the arithmetic mean  $E[x] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 F'^{(-1)}(x) dF'$  can be rewritten as:

$$L(F) = \frac{1}{E[x]} \int_{-\infty}^{F^{-1}(x)} x f(x) dx = \frac{1}{E[x]} \int_0^F F'^{(-1)}(x) dF'.$$

Graphically, for  $x > 0$ , the Lorenz curve  $L(F)$  is a concave curve ranging from  $0 \leq F \leq 1$  and  $0 \leq L(F) \leq 1$  which lies below the 45-degree line. If the Lorenz curve  $L(F_A)$  lies strictly above the Lorenz curve  $L(F_B)$ , then we say that  $F_A$  Lorenz dominates  $F_B$ . If we adjust the Lorenz curve by its mean  $E[x_i]$ , we get the so-called generalized Lorenz curve  $GL_i$ :

$$GL_i = E[x_i] L(F_i) = \int_0^F F_i'^{(-1)}(x) dF_i'. \quad (4)$$

As argued in Shorrocks (1983), we can then make a statement about second order stochastic dominance when comparing generalized Lorenz curves. If the generalized Lorenz curve for some asset  $B$  lies strictly above the same curve for asset  $A$ , asset  $B$

second-order stochastic dominates asset  $A$ . In case, the curves intersect, no clear ranking in terms of second order stochastic dominance is possible.

The Gini coefficient can be directly retrieved from the Lorenz curve using

$$G = 1 - 2 \int_0^1 L(F)dF. \quad (5)$$

Graphically, it is the ratio between the area between the Lorenz curve and the 45-degree line as a ratio of the total area under the 45-degree line (which is 0.5). Note that using the Gini coefficient it is (almost always) possible to rank distributions. Nevertheless, it does not have to imply second-order stochastic dominance if generalized Lorenz curves can intersect. This issue is well-known in the economics of inequality (Cowell, 2000).

## 2.2. The Gini coefficient and intersecting Lorenz curves

Let us now illustrate this theory with the help of the example of Goetzmann et al. (2002), also taken as a litmus-test in Zakamouline and Koekebakker (2009). In their paper, Goetzmann et al. (2002) argue that fund managers can gamble the Sharpe ratio by selling out of the money call and put options. Thus, the Sharpe ratio can be increased. Yet, investors are exposed to negative skewness and high kurtosis. For comparison, we adopt the procedure of Zakamouline and Koekebakker (2009) – who also develop an alternative performance measure – and fit a Normal Inverse Gaussian distribution (NIG) both to the stock prices themselves and to the manipulated portfolio. By construction, the returns on the non-modified portfolio follow a normal distribution. The latter is nested within the NIG distribution.

As shown in Figure 1, the manipulated portfolio contains an extreme left tail and thus exposes the investor to a large degree of downside risk. We report the first four moments and the performance measure in Table 1. For the assumed risk-free rate of 5%

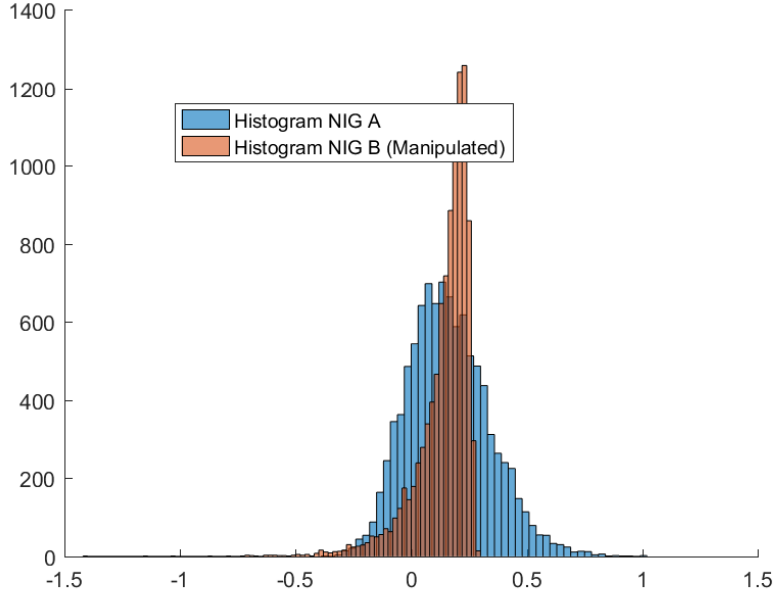


Figure 1: Histograms of returns of standard asset ( $A$ ) and manipulated asset ( $B$ ) for NIG fit.

The figure displays the return distributions for an asset  $A$  and an asset  $B$  having properties as documented in Table 1. Asset  $B$  is a manipulated version of asset  $A$  in order to gamble the Sharpe ratio as suggested in Goetzmann et al. (2002). The results were created by fitting the NIG distribution (for details cf. equation 3) using the four available moments and creating a Monte-Carlo simulation with  $N = 100,000$  observations each.

the Sharpe ratio of the manipulated portfolio is higher, suggesting a favorable investment opportunity.

Figure 2 plots the respective generalized Lorenz curves for arithmetic excess returns  $r_E$ . It becomes apparent that they intersect twice, both at the low end and at the high end. At the low end asset  $A$  is superior, as the manipulated asset exposes the investors to higher downside risk.<sup>6</sup> The same holds true at the right end following from the fact that the non-manipulated asset also displays higher mean returns.<sup>7</sup> Only in the intermediate

---

<sup>6</sup>Moreover, the curve undercuts the x-axis due to negative returns. If we use the exponential transformation ( $R = \exp(r)$ ) the curve is always above the x-axis. Yet, the two curves still intersect twice but the difference is very small. In general, the inequality measure of the transformed value  $R$  is lower than the one featuring negative values.

<sup>7</sup>The right end of the generalized Lorenz curve equals its mean  $GL_i(q = 1) = E[x_i]$  as  $L_i(q = 1) = 1$  and  $GL_i(q) = E[x_i]L_i(q)$ .

<b>Moments</b>	<b>A</b>	<b>B (manipulated)</b>
Mean	0.162	0.139
Std. dev.	0.177	0.12
Skewness	0.456	-2.358
Kurtosis	3.342	12.355
Sharpe ratio	0.631	0.743
Gini coeff.	0.1	0.055
$\lambda$	1.024	1.04
Atkinson index ( $\rho = 3$ )	0.045	0.025
$R_E^{CE} (\rho = 3)$	1.088	1.074

Table 1: Moments, risk and performance measures for the example of Goetzmann et al. (2002).

The manipulated asset  $B$  with the given first four moments produces a higher performance measure than the asset  $A$  if one judges it by the Sharpe ratio respectively the Gini mean difference  $\lambda$  (starting from the Gini coefficient as a measure of risk). This is not the case for certainty equivalent excess return  $R_E^{CE}$  (based on the Atkinson index as a measure of risk) with a risk aversion of  $\rho = 3$ . The risk-free rate is assumed to be 5% in line with Zakamouline and Koekebakker (2009).

area – the one which is targeted by the Sharpe ratio – the manipulated asset  $B$  appears superior. Due to the intersection(s), a ranking – in the sense of second-order stochastic dominance – is generically not possible. Thus, both approaches – the standard Sharpe ratio and the modified approaches of Goetzmann et al. (2007) and Zakamouline and Koekebakker (2009) – deliver rankings that are somewhat unstable in the sense that they are effectively imposing a specific degree of risk aversion and other degrees of risk aversion can potentially alter the rankings. In Table 1, we also see that the Gini mean difference ( $\lambda$ ) ranks the manipulated asset higher. We argue that these issues can be addressed by applying the so-called Atkinson index (Atkinson, 1970) and varying the preference parameters within a plausible range.

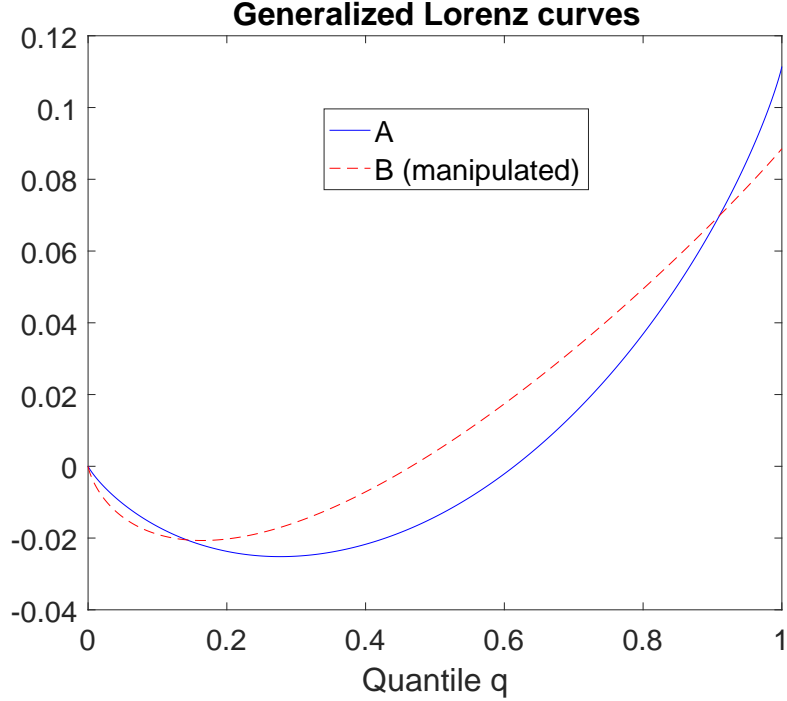


Figure 2: Generalized Lorenz curves for returns of standard asset ( $A$ ) and manipulated asset ( $B$ ).

The generalized Lorenz curves of returns  $r$  from asset  $A$  and the manipulated asset  $B$  (with the moments presented in Table 1) intersect twice meaning that it is not possible to rank them in terms of second-order stochastic dominance.

### 2.3. The Atkinson index

Drawing on standard utility theory with a utility function  $u$ , Atkinson (1970) proposes a dispersion measure and he starts from the certainty equivalent  $x^{CE}$ , defined through

$$u(x^{CE}) = \int_{-\infty}^{\infty} u(x)f(x)dx \leftrightarrow x^{CE} = u^{-1} \left( \int_{-\infty}^{\infty} u(x)f(x)dx \right). \quad (6)$$

He proposes the usage of power utility with a power exponent of  $1 - \rho$  for which  $\rho$  represents the coefficient of relative risk aversion. Thus, the underlying utility function

exhibits Constant Relative Risk Aversion (CRRA). Using this idea, he constructs a measure of inequality – henceforth called the Atkinson index  $A$  – defined as:

$$A = 1 - \frac{x^{CE}}{E[x]}. \quad (7)$$

Note that, for risk averse agents and payoffs with support on the positive part of the real line, we have that the Atkinson index is always between zero and one, where one is the highest degree of inequality, while 0 is the lowest. His initial application was ranking countries not by overall respectively per capita Gross Domestic Product, but by explicitly incorporating income inequality. The index itself is an adjustment for inequality and helps to evaluate income under the *veil of ignorance*.

Solving for the Atkinson index in the case of constant relative risk aversion, we have

$$A(\rho) = \begin{cases} 1 - \frac{1}{E[x]} (E[x^{1-\rho}])^{\frac{1}{1-\rho}} & \rho > 0, \rho \neq 1, \\ 1 - \frac{1}{E[x]} e^{E[\ln x]} & \rho = 1. \end{cases} \quad (8)$$

The index features the free parameter  $\rho$  measuring the degree of risk aversion. For the special case of risk neutrality we would have  $A(\rho = 0) = 0$  and thus  $x^{CE} = E[x]$ . The special case of  $\rho = 1$  (log-utility) implies that the Atkinson index  $A(\rho = 1) = 1 - \frac{GM[x]}{E[x]}$  represents a relationship between arithmetic and geometric mean  $GM[x]$ . In fact, in this case, the certainty equivalent is the geometric mean. In the special case when  $\rho = 2$ , the certainty equivalent equals the harmonic mean.<sup>8</sup>

Now, we want to apply the Atkinson index to fund returns. So far, we assumed some arbitrary random variable  $x$ . The literature in the economics of inequality has focused on the cross-sectional distribution of measures such as income or consumption. In the financial economics setting – which is considered here – utility depends on future wealth. Thus, we can replace the generic value  $x$  with  $w_0 R$  for which  $w_0$  is initial wealth (at time

---

<sup>8</sup>The latter is defined as  $HM(x) = (E[x^{-1}])^{-1}$ .



$t = 0$ ) and  $R$  represents the earned gross return. Below, we also model continuously compounded returns  $r$ , such that even if  $r$  can become negative,  $R = \exp(r)$  stays positive.

To facilitate comparisons with other performance measures, we also define certainty equivalent returns and certainty equivalent geometric excess returns. Note that geometric excess returns (index  $E$ ) are given by the ratio between returns and risk-free rates, while arithmetic excess returns – as e.g. employed in the standard Sharpe ratio – are defined as the difference between returns and risk-free rates. The geometric excess return is (among others) used in the rating of Morningstar (Morningstar, 2016). The certainty equivalent return ( $R^{CE}$ ) is the return on a risk-free investment that yields the same utility as an investment in the risky asset, defined implicitly through

$$u(w_0 R^{CE}) = E[u(w_0 \tilde{R})] \leftrightarrow R^{CE} = \frac{1}{w_0} u^{-1}(E[u(R)]), \quad (9)$$

in line with equation (6). The certainty equivalent geometric excess return is defined as

$$R_E^{CE} \equiv \frac{R^{CE}}{R_f}, \quad (10)$$

and the geometric excess return on a risky asset is defined as

$$\tilde{R}_E \equiv \frac{\tilde{R}}{R_f}. \quad (11)$$

Hence, the Atkinson index equals

$$A = 1 - \frac{w_0 R^{CE}}{E[w_0 \tilde{R}]} = 1 - \frac{R^{CE}}{E[\tilde{R}]} = 1 - \frac{R_E^{CE}}{E[\tilde{R}_E]}. \quad (12)$$

From the above equation, we can see intuitively why the Atkinson index is a risk measure: For a given expected return ( $E[\tilde{R}]$  or  $E[\tilde{R}_E]$ ),  $A$  is decreasing in the risk-adjusted return ( $R^{CE}$  or  $R_E^{CE}$ ) and the risk-adjusted return is lower whenever the risk is higher.

One important reason in favor of using (10) as a performance measure is that it is in line with stochastic dominance. Essentially, there are results connecting expected utility and stochastic dominance, so that —given some assumptions on the utility function— if a return distribution stochastically dominates another return distribution, expected utility and therefore certainty equivalent returns will be higher. The details are put forth in the following proposition.

**Proposition 2.1** *Suppose that we are comparing two return distributions,  $F_A$  and  $F_B$ . Further, suppose that the elementary utility function  $u(w)$  is sufficiently differentiable and satisfies  $(-1)^k u^{(k)}(w) < 0$  for  $k = 1, 2, \dots, n$  and all  $w$ . Then, we have that if  $F_A$   $n$ th-order stochastically dominates  $F_B$ , it must be that  $R_{E,A}^{CE} \geq R_{E,B}^{CE}$ .*

**Proof** The proof is relegated to Appendix A.1.

If we use CRRA or CARA utility, for example, we have that our performance measure is automatically in line (in the above sense) with  $n$ th-order stochastic dominance, regardless of the coefficients of risk aversion. In particular, if the elementary utility function is increasing and  $F_A$  first-order stochastically dominates  $F_B$ , our performance measure will yield a (weakly) higher value for returns following the distribution  $F_A$  (cf. Levin, 2006). Thus, it is not vulnerable to the same type of criticism that Hodges (1998) put forth regarding the Sharpe ratio (see Table 2). In his example, our performance measure will never yield a higher value for the first-order stochastically dominated asset, as long as the chosen utility function is increasing and Table 2 confirms this result in the case when we let the coefficient of relative risk aversion be equal to three. Another implication is that if the elementary utility function is increasing and concave and  $F_A$  second-order

stochastically dominates  $F_B$ , our performance measure will yield a (weakly) higher value for returns following the distribution  $F_A$  (cf. Laffont, 1989, Section 2.5). Further, with CRRA utility, the Atkinson index is independent of both initial wealth  $w_0$  and a constant risk-free rate. As we demonstrate below, these properties do not carry over to all utility functions.

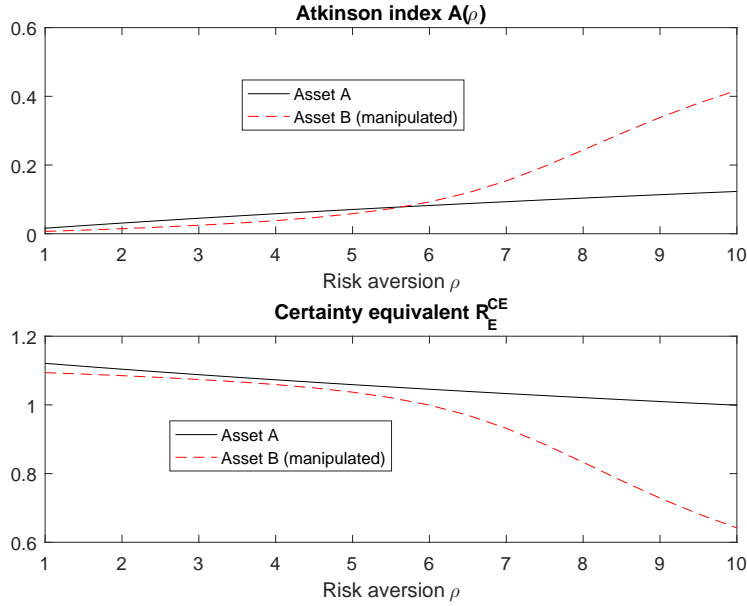


Figure 3: Atkinson index and certainty equivalent excess return for standard asset  $A$  and manipulated asset  $B$  for a range of values on  $\rho$ .

The Atkinson index  $A(\rho)$  (upper panel) as a measure of risk in asset  $A$  (with the moments presented in Table 1) and the manipulated asset  $B$  intersect allowing no general ranking. Yet, the certainty equivalent excess return is always superior for asset  $A$  for a risk aversion in the range  $1 \leq \rho \leq 10$ . The risk-free rate is assumed to be 5% in line with Zakamouline and Koekebakker (2009).

Returning to the example of Goetzmann et al. (2002) with an original distribution  $A$  and a manipulated one,  $B$ , and using a CRRA utility function, we find that, for intersecting generalized Lorenz curves – as depicted in Figure 2 – the value of  $\rho$  becomes important. For low values of  $\rho$ , agents focus on the upside and thus the manipulated distribution  $B$  yields a lower Atkinson index (indicating risk). Only for a sufficiently high  $\rho > \rho^*$  (which in our case is approximately 6, cf. Figure 3), does the Atkinson index indicate that asset  $A$  carries a lower risk. In contrast, the certainty equivalent excess

Probability	Excess Return [%]	
	Asset C	Asset D
0.01	-25	-25
0.04	-15	-15
0.25	-5	-5
0.4	5	5
0.25	15	15
0.04	25	25
0.01	35	45
Mean	0.05	0.051
Std. dev.	0.1	0.103
Skewness	0	0.305
Kurtosis	3.4	4.487
Sharpe ratio	0.5	0.493
Gini coeff.	0.0535	0.0544
$\lambda$	1.0002	1.0007
Atkinson index ( $\rho = 3$ )	0.015	0.0156
$R^{CE}$ ( $\rho = 3$ )	1.041	1.042

Table 2: Moments and performance measures for the example of Hodges (1998). The table presents the probability distribution of an asset  $C$  and another asset  $D$  as suggested by Hodges (1998) and the associated first four moments. Asset  $D$  first-order stochastically dominates asset  $C$  by having a more favorable return in one state. Yet, the Sharpe ratio fails to identify it as the superior asset. Meanwhile, both the Gini mean difference  $\lambda$  (built on the Gini as a risk measure) and the certainty equivalent  $R^{CE}$  (based on the Atkinson index as risk measure) with a risk aversion of  $\rho = 3$  identify asset  $D$  as superior.

return  $R_E^{CE}$  is superior for the unmanipulated asset  $A$  over the whole range of values on  $\rho$  (cf. lower panel of Figure 3). In fact, the manipulated asset  $B$  loses in relative attractiveness for high values of  $\rho$ .

The example depicted in Figure 1 compares returns following a normal distribution – being the basic working hypothesis in Finance and also the underlying assumption behind the Sharpe ratio – with a distribution obeying a NIG distribution. In the following, we derive closed-form measures for these two important parametric distributions. Moreover, we present a non-parametric approach using cumulants. Given this, we discuss the usefulness and the limitations of applying the Atkinson index to financial returns.

Another easy-to-grasp example in which the Sharpe ratio fails was proposed by Hodges (1998). Moreover, it is frequently taken as a litmus-test (cf. e.g. Zakamouline and Koekebakker (2009)) to benchmark alternative risk measures. Table 2 shows the payoff structure of two assets,  $C$  and  $D$ . The only difference is that asset  $D$  has a favorable state in which it delivers a higher return than asset  $C$ . Thus, it first-order dominates asset  $C$ . In terms of moments asset  $D$  has a higher mean, but also a higher variance. As the latter effect dominates, the Sharpe ratio of asset  $D$  is lower – erroneously indicating that asset  $C$  is more favorable. In terms of higher moments, asset  $D$  has a higher skewness, but also a higher kurtosis. Due to first-order stochastic dominance, both the Gini mean difference and the excess geometric certainty equivalent return correctly identify asset  $D$  as superior.

Finally, we want to compare our novel performance measure to some existing and popular measures. A very common measure – in particular to rank mutual funds – is the so-called *Morningstar* rating published by the investment research firm Morningstar (Morningstar, 2016). The rating itself is given in *stars* ranging from one to five given the overall investment category. 10% of the overall assets are labeled as one star (respectively 5 stars), 22.5% as 2 stars (respectively 4 stars) leaving a residual of 35% being 3 star assets. The Morningstar ratings are based on the certainty equivalent geometric excess

returns for a CRRA utility function with a coefficient of relative risk aversion of three<sup>9</sup> ( $\rho = 3$ ) and thus, there is a close connection between the Morningstar ratings and the Atkinson index. Below, we define  $M(\rho)$ , where  $M(3)$  corresponds to the Morningstar index:

$$M(\rho) \equiv R_E^{CE}(\rho) - 1 = (1 - A(\rho))E[\tilde{R}_E] - 1. \quad (13)$$

In Goetzmann et al. (2007), they propose a manipulation-proof performance measure that is in particular insensitive to dynamic trading strategies. Theoretically, the latter is defined as:

$$\hat{\theta}(\rho) = \frac{1}{1 - \rho} \ln \left( E[\tilde{R}_E^{1-\rho}] \right), \quad (14)$$

where they let

$$\rho = \frac{\ln \left( \frac{E[R_b]}{R_f} \right)}{\text{Var}(\ln R_b)}, \quad (15)$$

with  $R_b$  being the return on a benchmark portfolio. It is easy to see that we have

$$\hat{\theta}(\rho) = \ln(1 + M(\rho)) = \ln(1 - A(\rho)) + \ln E[\tilde{R}_E]. \quad (16)$$

Thus, we can decompose Goetzmann et al.'s (2007) measure into one part that is related to the Atkinson index, which measures risk, and one part related to expected returns.

---

<sup>9</sup>Note that in their classification Morningstar (2016) use an exponent of  $-\gamma$  rather than  $1 - \rho$  as suggested here. Thus, the value of  $\gamma = 2$  used by Morningstar (2016) translates into a value of  $\rho = 3$ .

### 3. Closed-form relations for the standard Atkinson index (CRRA utility)

In his seminal article, Atkinson (1970) assumes CRRA utility, which is convenient, because in this case, the certainty equivalent return,  $R^{CE}$  is independent of initial wealth, and thus, through equation (12), the Atkinson index is also independent of initial wealth. Below, we provide a general result, which links the standard Atkinson index to the cumulants of the return distribution. We then exemplify its properties for two specific and commonly assumed return distributions.

Given that the moment-generating function (MGF) and the cumulants exist,<sup>10</sup> we are able to express the Atkinson index and the certainty equivalent return in terms of the moment-generating function or the cumulants of the distribution.

**Proposition 3.1 (Relation between certainty equivalent and cumulants for CRRA utility)** *Suppose we have CRRA utility with a coefficient of relative risk aversion of  $\rho$  and that gross returns are given by  $R = e^r$ , where  $r$  follows some "well-behaved" distribution. Then, the Atkinson index equals*

$$A(\rho) = 1 - \frac{(\Psi_r(1 - \rho))^{\frac{1}{1-\rho}}}{\Psi_r(1)}, \quad (17)$$

where  $\Psi_r(t) \equiv E[e^{tr}]$  is the moment-generating function of  $r$ , and the continuously compounded certainty equivalent excess return is given by

$$r_E^{CE}(\rho) \equiv \ln R_E^{CE} = \frac{1}{1-\rho} k_r(1-\rho) - r_f, \quad (18)$$

---

<sup>10</sup>In particular, we exclude cases with non-finite moments. In the following we will refer to this class of moment-generating functions more briefly as "well-behaved".

where  $k_r(t) \equiv \ln \Psi_r(t) = \ln E[e^{tr}]$  is the cumulant-generating function of  $r$ .

The above expression can also be written as

$$r_E^{CE}(\rho) = \sum_{n=1}^{\infty} \frac{\kappa_n^r \cdot (-1)^{n-1} \cdot (\rho - 1)^{n-1}}{n!} - r_f, \quad (19)$$

where  $\kappa_n^r \equiv k_r^{(n)}(0)$  is the  $n$ th cumulant of  $r$ .

**Proof** A proof is given in Appendix A.2.

The first four cumulants are given by

$$\kappa_1^r = E[r] \quad (20)$$

$$\kappa_2^r = \text{Var}[r] = \sigma^2 \quad (21)$$

$$\kappa_3^r = \text{Skew}[r]\sigma^3 = E[(r - E[r])^3]\sigma^3 \quad (22)$$

$$\kappa_4^r = (\text{Kurt}[r] - 3\text{Var}[r]^2)\sigma^4 = (E[(r - E[r])^4] - 3\text{Var}[r]^2)\sigma^4. \quad (23)$$

Note that the fourth cumulant is the excess kurtosis relative to the normal distribution.

In fact, the cumulant approach is helpful in order to restrict the values of  $\rho$ . Consider the CRRA utility function underlying our performance measure

$$u(w) = \frac{w^{1-\rho} - 1}{1 - \rho}.$$

For integer-valued  $\rho > 1$  and  $n > 1$ , we can write the  $n$ th derivative as follows

$$\frac{\partial^n u(w)}{\partial w^n} = (-1)^{n-1} w^{1-\rho-n} (n-1)! \binom{\rho+n-2}{\rho-1},$$



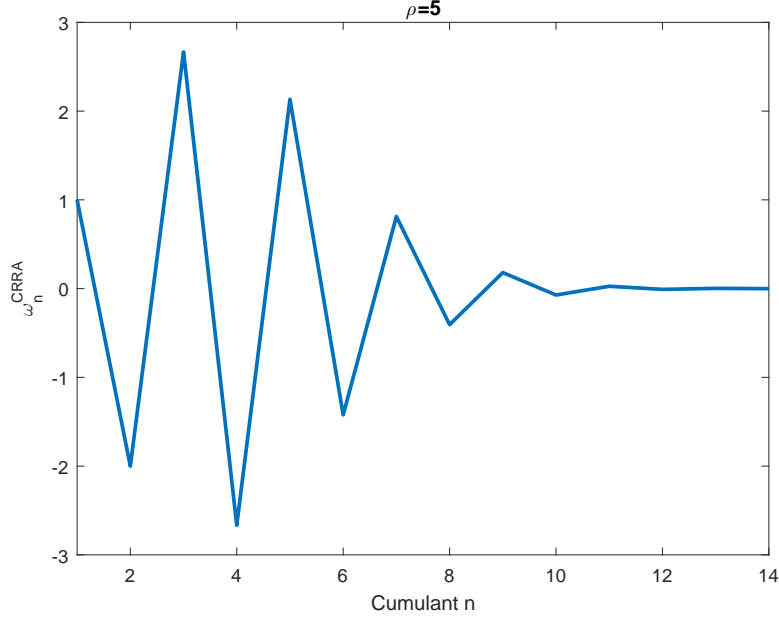


Figure 4: Weight  $\omega_n$  as function of cumulant  $n$  for CRRA utility with  $\rho = 5$ .

Weight  $\omega_n^{CRRRA}$  attached to the  $n$ -th cumulant when computing the certainty equivalent  $r^{CE}$  with CRRA utility. For the displayed case of relative risk aversion  $\rho = 5$  the largest weights are attached to cumulants 3 and 4. Weights alternate in sign starting with a positive value for  $n = 1$ .

with the last term being the binomial coefficient. The sign of this measure is positive for odd moments (mean, skewness, ...) and negative for even moments (variance, kurtosis, ...). In contrast to the quadratic utility function, the derivatives of the utility do not disappear for higher levels of  $n$  (for quadratic utility this is the case for  $n > 2$ ), but eventually explode in absolute terms.

A similar pattern emerges for the weights attached to the cumulants  $\omega_n^{CRRRA} = \frac{(-1)^{n-1}(\rho-1)^{n-1}}{n!}$ . Moreover, low moments are of higher importance than higher moments. In particular, the weight converges to zero for very high cumulants<sup>11</sup>

$$\lim_{n \rightarrow \infty} \frac{(-1)^{n-1}(\rho-1)^{n-1}}{n!} = 0.$$

<sup>11</sup>In contrast, the absolute value of the  $n$ th derivative explodes as  $n \rightarrow \infty$ .

The absolute weight is largest for

$$n^* = \rho - 1,$$

and  $n^* - 1$  with  $\omega_{n^*} = -\omega_{n^*-1}$ .

For example for  $\rho = 5$  the highest absolute weight is put on the cumulants  $n = 3$  (skewness) and  $n = 4$  (kurtosis) ( $\omega_3 = -\omega_4 > 0$ ). This case is displayed in Figure 4. As already discussed, odd (even) cumulants carry positive (negative) coefficients, which converge to zero as  $n$  goes to infinity.

For  $\rho = 3$  – as used in the Morningstar rating (Morningstar, 2016) – the largest weights are attached to the mean and variance. In fact,  $\rho = 3$  constitutes a lower barrier for this type of reasoning.

In order to illustrate our general findings we want to consider two specific return distributions. Even though it is widely rejected in empirical data, the assumption of normally distributed returns is still *the* working hypothesis in financial markets, presumably due to its mathematical convenience. If continuously compounded returns  $r$  are normally distributed,  $R = \exp(r)$  will be log-normally distributed. The following proposition summarizes the value of the Atkinson index in the case that  $R$  is log-normally distributed.

**Corollary 3.2 (Atkinson index for CRRA utility and log-normally distributed returns)** *Suppose we have CRRA utility with a coefficient of relative risk aversion of  $\rho$  and that gross returns are given by  $R = e^r$ , where  $r \sim N(\mu_r, \sigma_r^2)$ . Then, the Atkinson index equals*

$$0 < A(\rho) = 1 - \exp\left\{-\rho\sigma_r^2/2\right\} < 1. \quad (24)$$

**Proof** Here, we can use that  $E[R^{(1-\rho)}] = E[e^{(1-\rho)r}]$  and results regarding the moment-generating function for a normally distributed random variable. The complete proof is relegated to Appendix A.2.1.

Note that for a given variance, higher risk aversion  $\rho$  increases the Atkinson index, and for a given coefficient of relative risk aversion, a higher variance likewise increases the Atkinson index. It is also interesting to point out that in this case, risk only depends on the variance  $\sigma_r^2$ .

From equation (12), we can see that the continuously compounded certainty equivalent return is given by

$$r^{CE}(\rho) = \ln R^{CE} = \mu_r + \frac{1}{2}(1 - \rho)\sigma_r^2. \quad (25)$$

Note that the certainty equivalent depends on the value of  $\rho$  which is an exogenous risk aversion parameter. In the special case of log-utility ( $\rho \rightarrow 1$ ), the continuously compounded certainty equivalent return is equal to the geometric mean ( $r^{CE}(\rho = 1) = \mu_r$ ). In the case that returns follow a log-normal distribution this is also equal to the log of the median. Moreover, if  $\rho = 3$  we have that  $r^{CE}(\rho = 3) = \mu_r - \sigma_r^2$ , which in turn equals the log of the mode. An increase in risk aversion  $\rho$  increases the Atkinson index and decreases the value of the certainty equivalent. The latter is true for general distributions, not just the log-normal distribution.

We want to contrast the above with returns  $r$  following the NIG distribution implying that gross returns  $R = \exp(r)$  follow a log-NIG distribution. The NIG distribution was introduced by Barndorff-Nielsen (1997a) and since it allows for negatively skewed and fat-tailed return distributions, it plays an important role in modeling returns in finance (Barndorff-Nielsen, 1997b). Its probability density function is given by

$$f(r; \alpha, \beta, \mu, \delta) = \frac{\delta \alpha \exp(\delta \gamma + \beta[r - \mu])}{\pi \sqrt{\delta^2 + (r - \mu)^2}} K_1(\alpha \sqrt{\delta^2 + (r - \mu)^2})$$

with  $\gamma \equiv \sqrt{\alpha^2 - \beta^2}$  and  $K_1$  being the modified Bessel function of the third kind. Each of the four parameters has some intuitive interpretation. The parameter  $\mu$  is a location parameter implying that changing  $\mu$  translates the probability density function. Meanwhile,  $\delta > 0$  scales the overall distribution. The important parameter  $\beta$  captures asymmetry. In particular, for  $\beta < 0$  we have negative skewness, which in our financial application captures downside risk.<sup>12</sup> Finally, the parameter  $\alpha > 0$  captures the tail of the distribution. In particular, low values indicate fat tails and thus high kurtosis. The variables have to obey  $|\beta| < \alpha$  implying  $\gamma > 0$ .

The distribution can be easily summarized using the moment-generating function (MGF), which is given by

$$\Psi_r(t) \equiv E[\exp(tr)] = \exp(t\mu + \delta(\gamma - \sqrt{\alpha^2 - (\beta + t)^2})),$$

and generates an infinite number of moments. The first four moments, which are also the most important, amount to

$$\begin{aligned} E[\tilde{r}] &\equiv M = \mu + \delta \frac{\beta}{\gamma}, \\ \text{Var}[\tilde{r}] &\equiv \sigma^2 = \delta \frac{\alpha^2}{\gamma^3}, \\ \text{Skew}[\tilde{r}] &\equiv S = 3 \frac{\beta}{\alpha \sqrt{\delta \gamma}}, \\ \text{Kurt}[\tilde{r}] &\equiv K = 3 + \frac{3}{\delta \gamma} \left[ 1 + 4 \left( \frac{\beta}{\alpha} \right)^2 \right]. \end{aligned}$$

---

<sup>12</sup>This is also important as most parametric distribution functions are either symmetric or exhibit positive skewness (Evans et al., 2000).

With help of the above relations, we can solve for the four free parameters. Following Karlis (2002), we can thus fit the distribution with:

$$\begin{aligned}\hat{\alpha} &= \frac{3\sqrt{3\bar{K} - 9 - 4\bar{S}^2}}{\bar{\sigma}^2(3\bar{K} - 9 - 5\bar{S}^2)}, \\ \hat{\beta} &= \frac{3\bar{S}}{\bar{\sigma}(3\bar{K} - 9 - 5\bar{S}^2)}, \\ \hat{\mu} &= \bar{M} - \frac{3\bar{S}\bar{\sigma}}{3\bar{K} - 9 - 4\bar{S}^2}, \\ \hat{\delta} &= 3\bar{\sigma} \frac{\sqrt{3\bar{K} - 9 - 5\bar{S}^2}}{3\bar{K} - 9 - 4\bar{S}^2},\end{aligned}$$

for which the hat indicates an estimated value and the bar signifies the moment in the sample, with  $M$  denoting the mean,  $S$  skewness and  $K$  kurtosis, respectively. The function can only be fitted if  $\bar{K} - 3 > \frac{5}{3}\bar{S}^2$ .<sup>13</sup>

Another convenient feature of the NIG distribution is that we can retrieve the standard normal distribution as a special case by letting  $\beta = 0$  (symmetry),  $\delta \rightarrow \infty$ ,  $\alpha \rightarrow \infty$ , and  $\frac{\delta}{\alpha} = \sigma_r^2$ . Below, we solve analytically for the standard Atkinson index in the case when the continuously compounded returns follow a NIG distribution.

**Corollary 3.3 (Atkinson index for CRRA utility and log-NIG distributed returns)** *Suppose we have CRRA utility with a coefficient of relative risk aversion of  $\rho$ , and that gross returns are given by  $R = e^r$ , where  $r \sim NIG(\hat{\mu}_r, \delta_r, \alpha_r, \beta_r)$ . Then, the Atkinson index equals*

$$A(\rho) = 1 - \exp \left\{ \frac{\delta_r}{1 - \rho} \left( \rho \sqrt{\alpha_r^2 - \beta_r^2} + (1 - \rho) \sqrt{\alpha_r^2 - (\beta_r + 1)^2} - \sqrt{\alpha_r^2 - (\beta_r + (1 - \rho))^2} \right) \right\}. \quad (26)$$

---

<sup>13</sup>Otherwise, we would have  $|\beta| > \alpha$ .

**Proof** Here, we can use that  $E[R^{(1-\rho)}] = E[e^{(1-\rho)r}]$  and results regarding the moment-generating function for a NIG distributed random variable. We present a complete proof in Appendix A.2.2.

The log (continuously compounded) certainty equivalent return is given by

$$r^{CE}(\rho) = \hat{\mu}_r + \frac{\delta_r}{1-\rho} \left( \sqrt{\alpha_r^2 - \beta_r^2} - \sqrt{\alpha_r^2 - (\beta_r + (1-\rho))^2} \right) \quad (27)$$

The geometric mean (nested for  $\rho = 1$ ) is not well-defined in this case. Moreover, for the solutions in (26) and (27) to be well-defined, the parameter values must satisfy  $\alpha_r \geq \max\{|\beta_r|, |\beta_r + 1|, |\beta_r + (1-\rho)|\}$ .

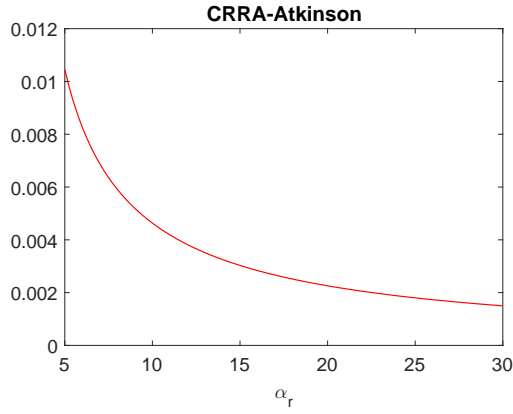
Due to the higher number of free parameters, the relation between the Atkinson index and the parametric distribution is more involved for the log-NIG distribution as compared to the log-normal distribution. To gauge the effects of parameter shifts around reasonable values we calibrate the first four moments to the average first four moments of hedge fund returns (covered in detail in section 5). Table 3 displays the results. As a benchmark we assume a parameter of risk aversion of three ( $\rho = 3$ ), in line with the calibration of Morningstar (2016). We provide (numerical) comparative statics by varying individual parameters around the benchmark calibration. As before, the Atkinson index does not depend on the translation parameter  $\hat{\mu}_r$ . However, the Atkinson index increases with the parameter of risk aversion  $\rho$  (cf. Figure 5c). As shown in Figure 5d the index is sensitive to scaling  $\delta_r$  as a larger value of  $\delta_r$  is accompanied by a higher Atkinson index. In general higher values of  $\delta_r$  increase the first two moments – being in the focus for a risk aversion of  $\rho = 3$ . More interestingly, the Atkinson index decreases with  $\alpha_r$  being our inverse tail parameter (cf. Figure 5a). Low values of  $\alpha_r$  increases the absolute values of the third and the fourth moment relative to the second moment. Finally, the parameter  $\beta_r$  in particular drives the skewness. More negative values implying more negative skewness also increase the Atkinson index, indicating

	<b>Return <math>r</math></b>	<b>Return <math>R</math></b>
Moment		
Mean	0.49 %	1.0042
Std. dev.	4.91 %	5.76%
Skewness	-0.41	-0.467
Kurtosis	8.27	8.07
Calibrated NIG	log-NIG	NIG
$\mu$	0.02	0.02
$\delta$	0.03	0.032
$\alpha$	12.52	12.32
$\beta$	-1.05	-1.19

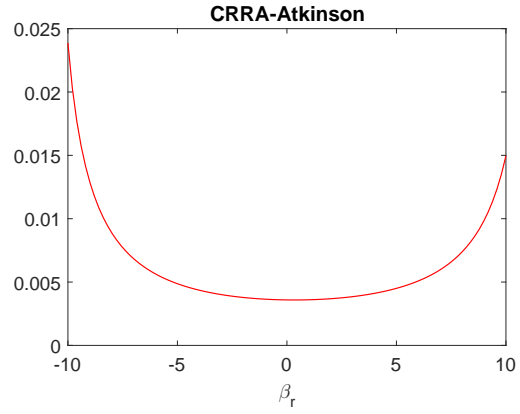
Table 3: Calibrated NIG model for average hedge fund data.

Using data from the Lipper Hedge Fund database of active funds from July 2007 to July 2017 (cf. Table 4), we calculate the first four moments of monthly returns  $r$  and transformed returns  $R = \exp(r) > 0$ , respectively. The left column reports the fit of the four parameters when fitting a NIG distribution on  $r$  (and thus a log-NIG distribution on  $R$ ), while the right column reports the parameter values when fitting a NIG distribution on  $R$ . We regard these values as reasonable steady state values for the comparative static exercises conducted in Figures 5 and 6, respectively.

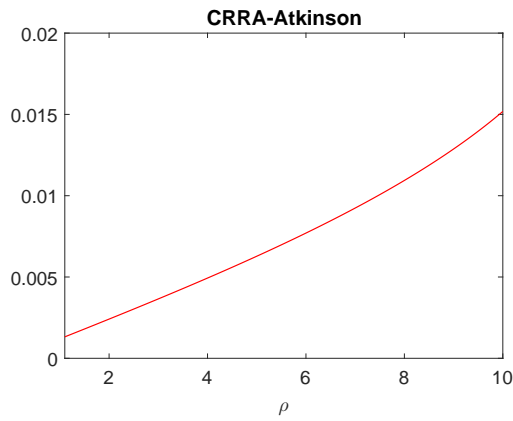
that the model is able to capture this effect. Interestingly, the partial derivative of the Atkinson index with respect to  $\beta_r$  is ambiguous (cf. Figure 5b). For high values of  $\beta_r$  implying high positive skewness and thus upside potential the Atkinson index is increasing in  $\beta_r$ . Note that this ambiguity stems from the fact that a higher  $\beta_r$  also increases the tails of the distribution and thus the kurtosis.



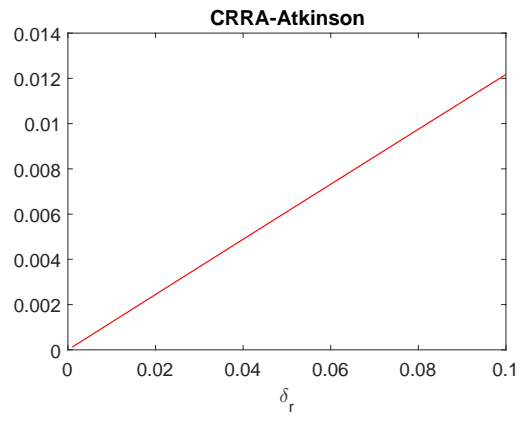
(a) Inverse tail parameter  $\alpha_r$



(b) Skewness parameter  $\beta_r$



(c) Relative risk aversion  $\rho$



(d) Scale parameter  $\delta_r$

Figure 5: Parameter variation for log-NIG distributed returns and CRRA utility. Atkinson index with CRRA utility with variation around the steady state values as reported in Table 3 (left column) for the parameters  $\alpha_r$ ,  $\beta_r$ ,  $\delta_r$  and relative risk aversion  $\rho$ .



## 4. Extensions of the Atkinson index within the class of HARA utility functions

In his seminal article, Atkinson (1970) uses CRRA utility, but his concept is open to essentially any monotone utility function. In order to generalize his concept, we can assume Hyperbolic Absolute Risk Aversion (HARA) utility which, e.g., nests CRRA utility as a special case;

$$u(w) = \frac{\rho}{1-\rho} \left( \frac{\lambda w}{\rho} + \phi \right)^{1-\rho}. \quad (28)$$

Obviously,  $\phi = 0$  results in CRRA utility. Assuming  $\phi = 1$  and letting  $\rho \rightarrow \infty$ , we obtain Constant Absolute Risk Aversion (CARA) utility with  $\lambda$  being the coefficient of absolute risk aversion. Further, in the special case when  $\rho = -1$ , we obtain quadratic utility.

**Proposition 4.1 (Atkinson index for general HARA utility)** *In the case of HARA utility (see equation 28), and a distribution of gross returns  $f(R)$ , we have that the Atkinson index is given by*

$$A(\rho, \lambda, \phi) = 1 + \frac{\rho}{\lambda w_0 \mu_R} \left( \phi - \left( \int_{-\infty}^{\infty} \left( \frac{\lambda w_0 R}{\rho} + \phi \right)^{1-\rho} f(R) dR \right)^{\frac{1}{1-\rho}} \right), \quad (29)$$

where  $w_0$  is initial wealth and  $\mu_R = \int_{-\infty}^{\infty} R f(R) dR$ . Further, the certainty equivalent geometric excess return is given by

$$R_E^{CE}(\rho, \lambda, \phi) = (1 - A) \frac{E[\tilde{R}]}{R_f} = \frac{\rho}{\lambda w_0 R_f} \left( \left( \int_{-\infty}^{\infty} \left( \frac{\lambda w_0 R}{\rho} + \phi \right)^{1-\rho} f(R) dR \right)^{\frac{1}{1-\rho}} - \phi \right). \quad (30)$$

**Proof** The result follows directly from applying the relations in (12) to the case of HARA utility (see equation 28).

In the following, we explore some interesting cases in which we can solve for the Atkinson index in closed form.

#### 4.1. Uniformly distributed returns and general HARA utility

The first case in which we can solve for the Atkinson index in closed form is the case of uniformly distributed returns and general HARA utility.

**Corollary 4.2 (Atkinson index for HARA utility and uniformly distributed returns)** *Suppose gross returns are uniformly distributed,*

$$f(R) = \frac{1}{b-a} \quad (a < R < b). \quad (31)$$

*Then, the Atkinson index is given by*

$$\begin{aligned} A(\rho, \lambda, \phi) = & 1 + \frac{2\rho\phi}{\lambda w_0(a+b)} \\ & - \frac{2}{(a+b)} \left( \frac{\rho}{\lambda w_0} \right)^{\frac{2-\rho}{1-\rho}} \left( \frac{1}{(2-\rho)(b-a)} \left( \left( \frac{\lambda w_0 b}{\rho} + \phi \right)^{2-\rho} - \left( \frac{\lambda w_0 a}{\rho} + \phi \right)^{2-\rho} \right) \right)^{\frac{1}{1-\rho}} \end{aligned} \quad (32)$$

*The corresponding certainty equivalent geometric excess return amounts to*

$$\begin{aligned} R_E^{CE}(\rho, \lambda, \phi) = & -\frac{\rho\phi}{\lambda w_0 R_f} \\ & + \frac{1}{R_f} \left( \frac{\rho}{\lambda w_0} \right)^{\frac{2-\rho}{1-\rho}} \left( \frac{1}{(2-\rho)(b-a)} \left( \left( \frac{\lambda w_0 b}{\rho} + \phi \right)^{2-\rho} - \left( \frac{\lambda w_0 a}{\rho} + \phi \right)^{2-\rho} \right) \right)^{\frac{1}{1-\rho}} \end{aligned} \quad (33)$$

**Proof** The result follows immediately from inserting the distribution in (31) into the expression in (29) and calculating the integral.

Here, we can note that, in general, initial wealth  $w_0$  and the risk aversion parameter  $\lambda$  enter the equation jointly. The latter determines the degree of absolute risk aversion for the case of a utility function of the Constant Absolute Risk Aversion (CARA) type. Thus, in this case, the level of initial wealth  $w_0$  matters for risk considerations.

The uniform distribution is certainly not a realistic description of asset returns. In particular, this distribution has zero skewness and negative excess kurtosis (amounting to -1.2). Yet, it has the convenient feature that it places both a lower bound (the variable  $a$ ) and an upper bound  $b$  on asset returns. A natural lower bound is  $a = 0$  implying  $R > 0$ .

Let us consider the case of CRRA utility already discussed so far. The latter is nested in the general result for  $\phi = 0$ . Let us also assume  $a = 0$  and  $0 < \rho < 2$ . In this case, the geometric certainty equivalent return simplifies to

$$R^{CE}(\rho) = (2 - \rho)^{\frac{1}{\rho-1}} b.$$

First of all, the result is independent of initial wealth  $w_0$  – which is always the case for CRRA utility. Secondly, the certainty equivalent increases with the maximum return  $b$  in a linear manner. Moreover, it decreases with risk aversion  $\rho$  for  $1 < \rho < 2$ .

We can also investigate the case of CARA utility. In this case, we have

$$R^{CE}(\lambda) = -\frac{1}{\lambda} \ln(1 - \exp(-\lambda b)) + \frac{1}{\lambda} \ln(\lambda b) \approx \frac{\ln(\lambda b)}{\lambda}.$$

The first part can be neglected for reasonably high values on both  $\lambda$  and  $b$ . Of course, the certainty equivalent decreases with  $\lambda$  while it increases with  $b$  in a concave manner.

## 4.2. Quadratic utility

Another case in which it is relatively easy to arrive at analytical solutions is the case of quadratic utility, which occurs when  $\rho = -1$ .

**Proposition 4.3 (Atkinson index for quadratic utility)** *Suppose the utility function is quadratic, i.e.,  $\rho = -1$  and that the return distribution has positive support only for  $R \in (0, \phi/(\lambda w_0))$ , where the upper bound guarantees nonsatiation. Then, the Atkinson index is given by*

$$A(\lambda, \phi) = 1 - \frac{1}{\lambda w_0 \mu_R} \left( \phi - \sqrt{\lambda^2 w_0^2 \sigma_R^2 + (\phi - \lambda w_0 \mu_R)^2} \right). \quad (34)$$

Here, we note that with the distributional restriction of positive support only for  $R \in (0, \phi/(\lambda w_0))$ , the Atkinson index lies between zero and one.<sup>14</sup> Further, if the variance of returns is zero ( $\sigma_R^2 = 0$ ), the Atkinson index takes a value of zero, meaning that we have total equality.

First of all, if we assume quadratic utility, we also assume away any impact of moments higher than the variance and thus, we face similar restrictions as for the Sharpe ratio. Second, marginal utility can become negative if we do not pick parameter values such that the returns stay below the satiation point. If we allow returns to have positive support beyond the satiation level, the Atkinson index can become greater than one even for positive expected returns, meaning that the certainty equivalent return can become negative.

Third, as we have in principle already shown, it is at most in line with second-order stochastic dominance. This is because the derivatives of the utility function for orders of  $n > 2$  are all zero. Thus, in cases in which there is no second-order stochastic dominance

---

<sup>14</sup>This is very intuitive, because over that range of returns, the utility function is concave (and increasing). To see it algebraically, it helps to note that  $R < \phi/(\lambda w_0)$  implies that  $\mu_R < \phi/(\lambda w_0)$ , and that  $R^2 < \phi^2/(\lambda^2 w_0^2)$  implies that  $\sigma_R^2 + \mu_R^2 < \phi^2/(\lambda^2 w_0^2)$ .

but, e.g., third-order stochastic dominance, the value of the certainty equivalent return based on quadratic utility might be lower.

The certainty equivalent return is given by

$$R_E^{CE}(\lambda, \phi) = \frac{1}{\lambda w_0 R_f} \left( \phi - \sqrt{\lambda^2 w_0^2 \sigma_R^2 + (\phi - \lambda w_0 \mu_R)^2} \right). \quad (35)$$

For  $\sigma_R^2 = 0$  we have that  $R_E^{CE} = \mu_R$ . Using the normalization  $w_0 = 1/R_f$ , the above expression simplifies to

$$R_E^{CE}(\lambda, \phi) = \frac{1}{\lambda} \left( \phi - \sqrt{\lambda^2 w_0^2 \sigma_R^2 + (\phi - \lambda w_0 \mu_R)^2} \right). \quad (36)$$

A specific and insightful case would be the one of a uniform distribution introduced in the previous section. In this case we have a mean of  $\mu_R = 0.5(a + b)$  and a variance of  $\sigma_R^2 = \frac{1}{12}(b - a)^2$  implying the following certainty equivalent if we normalize  $\lambda$  to unity ( $\lambda = 1$ )

$$R_E^{CE} = \phi - \frac{1}{\sqrt{3}} \sqrt{a^2 + b^2 + ab - 3\phi(a + b) + 3\phi^2}.$$

In order to avoid the satiation property of the quadratic utility function it is reasonable to assume  $\phi = b$  simplifying the previous relation to

$$R_E^{CE} = b - \frac{1}{\sqrt{3}}(b - a).$$

The above certainty equivalent increases with the maximum possible return  $b$  and decreases with the range  $b - a$  (also proportional to the standard deviation) as a measure of dispersion. For the realistic special case of  $a = 0$  we would have  $R_E^{CE} = \frac{\sqrt{3}-1}{\sqrt{3}}b \approx 0.42b$ . For the opposing case in which  $b = a$  - i.e. no dispersion - we would have  $R_E^{CE} = b$  as the return now is certain.

### 4.3. CARA utility

We want to contrast our results for Constant Relative Risk Aversion (CRRA) to the case of Constant Absolute Risk Aversion (CARA) which is a special case of the more general HARA utility: we can obtain CARA utility with a constant absolute risk aversion coefficient of  $\lambda$  by letting  $\phi = 1$  and  $\rho \rightarrow \infty$ .

As it turns out, under CARA utility, the Atkinson index is given by

$$A(\lambda) = 1 + \frac{1}{\lambda w_0 \mu_R} \ln \left( E[e^{-\lambda w_0 \tilde{R}_i}] \right), \quad (37)$$

It is easy to check that if  $R$  is constant, we get that  $A(\lambda) = 0$  for all  $\lambda$ , i.e., "total equality".

Further, the certainty equivalent geometric excess return is given by

$$R_E^{CE}(\lambda) = -\frac{1}{\lambda w_0 R_f} \ln \left( E[e^{-\lambda w_0 \tilde{R}_i}] \right). \quad (38)$$

Applying the normalization  $w_0 = 1/R_f$ , we get

$$R_E^{CE}(\lambda) = -\frac{1}{\lambda} \ln \left( E[e^{-\lambda \tilde{R}_E}] \right), \quad (39)$$

where  $R_E$  is the geometric excess return.

Similar to the CRRA case, we present a non-parametric performance measure using cumulants. Afterwards we exemplify this for the two parametric cases of a normal distribution and a NIG distribution.

**Proposition 4.4 (Relation between certainty equivalent and cumulants for CARA utility)** *Suppose we have CARA utility and that gross returns  $R$  follow some "well-behaved" distribution. Then, the Atkinson index equals*

$$A(\lambda) = 1 + \frac{k_R(-\lambda w_0)}{\lambda w_0 \mu_R}, \quad (40)$$

where  $k_R(t) \equiv \ln E[e^{tR}]$  is the cumulant-generating function of  $R$ . Further, the certainty equivalent geometric excess return is given by

$$R_E^{CE}(\lambda) = -\frac{k_R(-\lambda w_0)}{\lambda w_0 R_f}. \quad (41)$$

Applying the normalization  $w_0 = 1/R_f$  to the above expression, we get

$$R_E^{CE}(\lambda) = -\frac{k_R(-\lambda/R_f)}{\lambda} = \sum_{n=1}^{\infty} \frac{\kappa_n^R \cdot (-1)^{n-1} \lambda^{n-1} R_f^{-n}}{n!}, \quad (42)$$

where  $\kappa_n^R = k_R^{(n)}(0)$  is the  $n$ th cumulant of  $R$ .

**Proof** We present a formal proof in Appendix A.3.

As in the CRRA case, the sign of the derivative of the utility function is positive for odd  $n$ , corresponding to odd moments (mean, skewness, ...) and negative for even  $n$ , corresponding to even moments (variance, kurtosis, ...).<sup>15</sup> Moreover, low moments

---

<sup>15</sup>The  $n$ th derivative of the utility function  $u(w) = -\exp(-\lambda w)$  is given by  $\frac{\partial^n u(w)}{\partial w^n} = (-1)^{n-1} \lambda^n \exp(-\lambda w)$ .

are of higher importance than higher moments. In particular, the discount  $\omega_n^{CARA} = \frac{(-1)^{n-1} \lambda^{n-1} w_0^n}{n!}$  vanishes for very high cumulants:

$$\lim_{n \rightarrow \infty} \frac{(-1)^{n-1} \lambda^{n-1} w_0^n}{n!} = 0. \quad (43)$$

Similar to the case of CRRA utility there is a local maximum for the weights. The absolute maximum weight is attained for

$$n^* = \lambda - 1, \quad (44)$$

and  $n^* + 1$ . If – for example – we have  $\lambda = 3$  the largest weights are  $\omega_3 = -\omega_2 > 0$ . In fact, a similar relation as presented in Figure 4 holds when replacing  $\rho$  with  $\lambda - 1$ .

The CARA utility results are reminiscent of the approach proposed in Stutzer (2000). His performance measure is given by:

$$P_S = \max_{\eta > 0} \left\{ -\ln \left( E[e^{-\eta(\tilde{R}_E - R_b)}] \right) \right\}, \quad (45)$$

where  $R_b$  is the return on the benchmark portfolio.

Thus, using the normalization  $w_0 = 1/R_f$  and using the risk-free asset as a benchmark ( $R_b = R_f$ ), we can write the Stutzer index as

$$P_S = \max_{\lambda > 0} \left\{ \frac{1}{\lambda} R_E^{CE}(\lambda) - \lambda R_f \right\}. \quad (46)$$

However, in general, for stochastic benchmark portfolio returns  $R_b$ , the relation will not be as straight-forward.

In order to gain intuition into the general findings we want to consider specific parametric distributions starting with the assumption of normally distributed returns.



**Corollary 4.5 (Atkinson index for CARA utility and normally distributed returns)** *Suppose we have CARA utility and that gross returns are normally distributed, i.e.,  $R \sim N(\mu_R, \sigma_R^2)$ . Then, the Atkinson index is given by*

$$A(\lambda) = \frac{1}{2} \lambda w_0 \frac{\sigma_R^2}{\mu_R}, \quad (47)$$

*and the certainty equivalent geometric excess return is given by*

$$R_E^{CE}(\lambda) = \frac{\mu_R}{R_f} - \frac{1}{2} \frac{\lambda w_0 \sigma_R^2}{R_f}. \quad (48)$$

*Applying the normalization  $w_0 = 1/R_f$  to the above expression, we get*

$$R_E^{CE}(\lambda) = \frac{\mu_R}{R_f} - \frac{1}{2} \frac{\lambda \sigma_R^2}{R_f^2}. \quad (49)$$

**Proof** The proof is presented in Appendix A.3.1.

The result itself contains some interesting insights. As compared to the case with CRRA utility and positive gross returns, the Atkinson does not need to be smaller than one. This is because the distribution that we assume above has positive support for negative gross returns. In general, both the Atkinson index and the certainty equivalent returns increase with the variance  $\sigma_R^2$  as well as the degree of absolute risk aversion  $\lambda$ . In the case of CARA utility, however, the level of risk perception depends on initial wealth ( $w_0$ ). As individuals have a constant *absolute* level of risk endurance, this implies that — *ceteris paribus*— wealthier individuals (higher levels of  $w_0$ ) derive a higher Atkinson index respectively a lower certainty equivalent for a given asset. The index takes similar values as for the case with CRRA utility with  $\rho = \lambda$  if we normalize initial wealth to

one ( $w_0 = 1$ ) and make the (reasonable) assumption that expected net returns are small ( $\mu_R \approx 1$ ).<sup>16</sup>

Similar to the case of CRRA utility we can not only derive closed-form solutions for the normal distribution, but also for the more interesting NIG distribution.

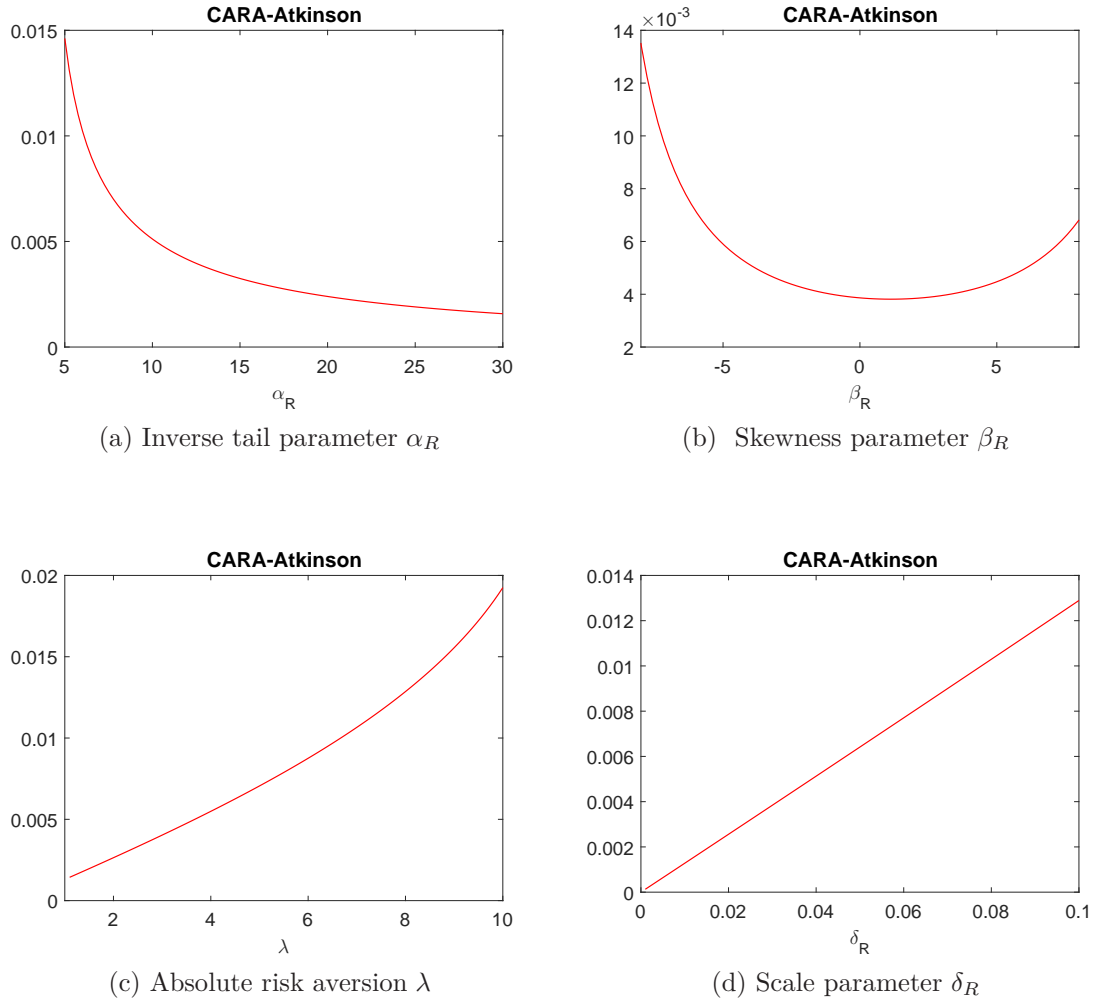


Figure 6: Parameter variation for NIG distributed returns and CARA utility. Atkinson index with CARA utility where we vary the parameter values  $\alpha_R$ ,  $\beta_R$ ,  $\delta_R$  and absolute risk aversion  $\lambda$  around the steady state values as reported in table 3 (right column).

<sup>16</sup>Moreover, we use the approximation  $\exp(x) \approx 1 + x$ .

**Corollary 4.6 (Atkinson index for CARA utility and NIG-distributed returns)** *Suppose we have CARA utility and that gross returns are NIG-distributed, i.e.,  $R \sim \text{NIG}(\hat{\mu}_R, \delta_R, \alpha_R, \beta_R)$ . Then, the Atkinson index is given by*

$$A(\lambda) = 1 + \frac{1}{\lambda} \frac{\left(-\lambda w_0 \hat{\mu}_R + \delta_R \left(\sqrt{\alpha_R^2 - \beta_R^2} - \sqrt{\alpha_R^2 - (\beta_R - \lambda w_0)^2}\right)\right)}{w_0 \left(\hat{\mu}_R + \frac{\delta_R \beta_R}{\sqrt{\alpha_R^2 - \beta_R^2}}\right)}, \quad (50)$$

and the certainty equivalent geometric excess return is given by

$$R_E^{CE}(\lambda) = \frac{1}{\lambda w_0 R_f} \left(\lambda w_0 \hat{\mu}_R - \delta_R \left(\sqrt{\alpha_R^2 - \beta_R^2} - \sqrt{\alpha_R^2 - (\beta_R - \lambda w_0)^2}\right)\right). \quad (51)$$

Here, we can also apply the normalization  $w_0 = 1/R_f$ :

$$R_E^{CE}(\lambda) = \frac{\hat{\mu}_R}{R_f} - \frac{\delta_R}{\lambda} \left(\sqrt{\alpha_R^2 - \beta_R^2} - \sqrt{\alpha_R^2 - (\beta_R - \lambda w_0)^2}\right). \quad (52)$$

**Proof** The proof is relegated to Appendix A.3.2.

Note that in order for the solutions in (51) and (52) to be well-defined, the parameter values need to satisfy  $\alpha_R \geq \max\{|\beta_R|, |\beta_R - \lambda w_0|\}$ .

In Figure 6, we repeat a similar exercise already undertaken for the case in which returns follow a log-NIG distribution and the utility function is of the standard CRRA type. In this case, however, we consider CARA utility and NIG-distributed returns. Once again, we use a realistic calibration of the four free parameters of the NIG distribution to fit hedge fund returns and vary the free parameters around it (see Table 3). As a benchmark, we assume a risk aversion parameter of  $\lambda = 3$  and normalize initial wealth to one ( $w_0 = 1$ ). The results not only point to the same qualitative result (as captured by the slopes of the curves in both Figure 6 and 5), but also the overall magnitudes of

both indexes are very close. The only major difference between the two utility functions is that for CARA utility, the parameter  $\hat{\mu}_R$  driving the mean matters. In fact, high values of  $\hat{\mu}_R$  decrease the Atkinson index, while increasing the certainty equivalent.

So far, we presented some general closed-form results. As demonstrated below, these results are of high importance for business practitioners.

## 5. Empirical tests

In this section, we bring the previously discussed methods to the data. We consider two cases. In the first application, we look at a large number of hedge funds showing that our approach yields new insights as compared to several other well-known performance measures by looking at rank correlations. We contrast this exercise with a large number of (opaque) investment strategies by having a close and detailed look at the most popular and well-studied market anomalies (size, value and momentum) and we show that these strategies lose their *glamour* once considered under the lens of our performance measure. In the second part we also show that the ranking is robust to measuring error for a high risk aversion.

### 5.1. Hedge funds

While low-risk securities might be well-described by a normal distribution, this assumption is highly questionable for more sophisticated investment strategies such as the ones undertaken by hedge funds. In fact, it is well documented that hedge funds exhibit non-normal returns (Malkiel and Atanu (2005), Brooks and Kat (2002)) in particular featuring left tail risk (Agarwal and Naik, 2004). Thus, in the literature evaluating alternative performance measures, they are frequently considered as appropriate data to compare different indexes (Eling and Schuhmacher, 2007; Zakamouline and Koekebakker, 2009).

Category	Obs. N	Mean [%]	Std. dev. [%]	Skewness	Kurtosis	Rejected JB [%]	Rejected AD [%]
Credit Focus	31	0.79	4.29	-0.35	10.83	100	100
Emerging Markets	73	0.51	10.03	-0.86	8.64	78.08	69.86
Equity Market Neutral	31	0.4	3.96	-0.54	7.19	77.42	64.52
Event Driven	64	0.46	4.23	-0.39	7.41	95.31	89.06
Fixed Income Arbitrage	26	0.59	3.93	-1.38	15.98	96.15	92.31
Global Macro	33	0.57	4.21	0.13	7.13	75.76	72.73
Long-Short Equity	232	0.48	5.12	-0.23	5.42	70.69	63.36
Managed Futures CTA	142	0.49	5.64	0.09	5.50	44.37	35.21
Multi Strategies	41	0.11	2.78	-0.90	6.36	80.49	70.73
All	673	0.49	4.91	-0.41	8.27	79.81	73.09

Table 4: Summary statistics for hedge fund returns.

Summary statistics for the first four moments of monthly returns for Hedge Funds covered in the Lipper Hedge Fund database from July 2007 to July 2017. The strategies are split up into 9 distinctive categories in which CTA stands for Commodity Trading Advisors. For each time series, we run both a Jarque-Bera (JB) and an Anderson-Darling (AD) test for normality with a significance level of 5%. The table reports the share of hedge fund time series failing this test.

We employ the so-called Lipper Hedge Fund (formerly TASS) database provided by Thomson Reuters. We consider monthly returns of funds active<sup>17</sup> between July 2007 and June 2017 (i.e. 120 observations in time). In the database, the funds are grouped into nine distinctive investment strategies. In Table 4, we present summary statistics for monthly fund returns. Most fund categories feature excess kurtosis and are also frequently characterized by negative skewness indicating downside risk. More formally, we compute both the Jarque-Bera test (JB) and the alternative Anderson-Darling (AD) test of normality with a significance level of 5%. The majority of the funds fail this test. While both tests yield similar results, the JB test is in general more strict, rejecting the hypothesis of normality for a larger share of hedge fund returns.

We use monthly data on the three-month Treasury bill provided by the Federal Reserve Bank of St. Louis as our proxy for the risk-free rate. We insert the excess returns<sup>18</sup>

<sup>17</sup>Note that the focus on funds that were active during the complete sample (in order to get a long observation period) reduces the overall sample substantially from approx. 7,500 funds to 673. Moreover, it imposes the issue of survivorship bias. Note that Zakamouline and Koekebakker (2009) perform a similar exercise with the TASS database, but with data ranging from 1994 until 2007, i.e., their sample ends before the financial crisis. In general, we find lower returns and more volatility.

<sup>18</sup>More specifically, we use the geometric excess returns  $R_{E,i,t} = \frac{R_{i,t}}{R_{f,t}} > 0$ .

	<b>Gini</b>	<b>CRRA 3</b>	<b>CRRA 10</b>	<b>Quad</b>	<b>Sharpe</b>	<b>ASSR</b>
Credit Focus	1	1	3	1	1	1
Emerging Markets	9	9	9	7	8	8
Equity Market Neutral	5	5	7	8	5	5
Event Driven	4	4	4	6	4	3
Fixed Income Arbitrage	2	2	6	2	2	7
Global Macro	6	3	2	3	3	2
Long-Short Equity	7	6	5	5	6	6
Managed Futures CTA	8	8	8	4	7	4
Multi Strategies	3	7	1	9	9	9

Table 5: Ranking hedge fund strategies using different performance measures.

Ranking of overall hedge fund investment strategies with data as reported in Table 4 from best (1) to worst (9) according to different performance measures: Gini mean difference (**Gini**), certainty equivalent from Atkinson index with constant relative risk aversion with risk aversion  $\rho = 3$  (Morningstar, 2016) and  $\rho = 10$ , respectively (**CRRA 3** and **CRRA 10**, respectively), certainty equivalent with quadratic utility (**Quad**), Sharpe ratio and Adjusted for Skewness Sharpe Ratio (**ASSR**) with a coefficient of constant relative risk aversion of 3 (Zakamouline and Koekebakker, 2009).

into our Atkinson- and Gini-based certainty equivalent measures. In contrast to well-known performance measures, the Atkinson index – from which we compute the certainty equivalent as our performance measure – entails a degree of freedom which controls the degree of risk aversion. As mentioned earlier, the popular Morningstar rating is a special case of the certainty equivalent with CRRA utility having a coefficient of relative risk aversion of three ( $\rho = 3$ ) (Morningstar, 2016). Thus, we consider this a good starting point in terms of risk aversion. We cross-check this with higher values of  $\rho$  which assign larger weights to higher-order cumulants. Other alternative measures that we consider are the Gini-ratio (Yitzhaki, 1982) and the Adjusted for Skewness Sharpe Ratio (ASSR) of Zakamouline and Koekebakker (2009) with a coefficient of constant relative risk aversion of three.

Using this approach, we can rank the nine distinctive hedge fund strategies identified by Thomson Reuters. The results are reported in Table 5. By construction, the approach of Zakamouline and Koekebakker (2009) taking skewness into account ranks the *fixed-income arbitrage* strategy, which exhibits a large amount of negative skewness

	<b>Gini</b>	<b>CRRA 3</b>	<b>CRRA 4</b>	<b>CRRA 5</b>	<b>CRRA 6</b>	<b>CARA 2</b>	<b>Quad</b>	<b>Sharpe</b>	<b>ASSR</b>
<b>Gini</b>	1.00								
<b>CRRA 3</b>	0.72	1.00							
<b>CRRA 4</b>	0.82	0.98	1.00						
<b>CRRA 5</b>	0.87	0.94	0.99	1.00					
<b>CRRA 6</b>	0.91	0.90	0.96	0.99	1.00				
<b>CARA 2</b>	0.57	0.96	0.90	0.84	0.78	1.00			
<b>Quad</b>	0.21	0.75	0.63	0.54	0.47	0.88	1.00		
<b>Sharpe</b>	0.61	0.92	0.87	0.83	0.78	0.94	0.85	1.00	
<b>ASSR</b>	0.47	0.84	0.79	0.75	0.70	0.88	0.84	0.93	1.00

Table 6: Rank correlation of hedge funds when evaluated according to different performance measures.

Spearman rank correlation of performance measures for all assets with data as reported in Table 4. We compare Atkinson-based certainty equivalents based on CRRA utility (with coefficient of relative risk aversion ranging between 3 and 6), CARA (risk aversion  $\lambda = 2$ ), and quadratic utility (**Quad**) with existing measures of Gini mean difference (**Gini**), Sharpe ratio (**Sharpe**), and Adjusted for Skewness Sharpe Ratio (**ASSR**, risk aversion 3).

considerably lower. The same holds true for the Atkinson approach if we use a large coefficient of risk aversion. Yet, the Atkinson approach also captures higher moments such as kurtosis. In particular, the highly non-normal strategy *credit focus* is not rated as the best strategy according to the Atkinson approach with a high risk aversion. Note that this simple analysis only takes into account the average group performance and consequently, it ignores the substantial within-group variation. Thus, in the following, we consider a more detailed approach.

In order to compare the methods, it is common in the literature to compute the rank correlation (Spearman, 1904) between the measures. If these are close to one (as e.g. reported in Eling and Schuhmacher (2007)), the performance measures yield little new insights. The correlation with the Sharpe ratio is of particular interest.

We compute the Spearman correlation matrix for all individual funds and report it in Table 6. The table contains several interesting insights. First of all – and in contrast to the study of Eling and Schuhmacher (2007) – in general the table features low correlations (below 0.95) indicating that the measures provide different rankings. For higher degrees of risk aversion for the Atkinson-based measures, the correlation with the standard Sharpe ratio decreases. This is in line with the notion that higher risk aversion implies a shift to higher cumulants beyond the second order, as considered in the Sharpe

ratio. In general, the ASSR has a lower correlation than the standard Sharpe ratio with the novel measures, yet the overall pattern for different degrees of risk aversion pertains. Meanwhile, the correlation with the Gini coefficient increases with assumed risk aversion in Atkinson-type measures. Constant absolute and relative risk aversion performance measures with similar risk aversion have high correlations. The latter is also due to the (implicit) assumption of a normalized value of wealth  $w_0 = 1$ . The measure based on the quadratic utility function has the lowest overall correlation with other measures.<sup>19</sup> Out of the performance measures that we consider, the Sharpe ratio and the CARA measure with an absolute risk aversion of two come closest to quadratic utility, which is natural, since they place their focus on the mean and the variance according to our results towards the end of Section 3.

In Figure 7 we plot the cumulants for the given observations up to an order of 6. In general, the median cumulant decreases with its order. Yet, with a higher order the measures feature more extreme observations. As presented in the upper panel of Figure 7, for a high order the mean cumulant exceeds the 95% quantile range indicating some substantial deviations for some individual observations. Thus, in general high order cumulants do not play an important role. Yet, they can take substantial values for certain assets and thus should be taken into account by investors.

So far, we considered the novel performance measures discussed respectively introduced in this paper and contrasted them with the standard Sharpe ratio. We can, however, also compare them to other more established performance measures such as those considered in Eling and Schuhmacher (2007) which are also standard in business practice. These measures all consist of ratios for which the numerator captures arithmetic excess returns ( $E(r_i - r_f)$ ) and the denominator is a measure of risk. As emphasized, e.g., by Hodges (1998), one major shortcoming of using the simple standard deviation as a risk measure is that it also punishes upside potential. Thus, in the following, we focus

---

<sup>19</sup>Note that, in order to avoid negative marginal utility, we set the satiation level to the highest observed monthly return in the dataset ( $r_{max} = 505\%$ ).



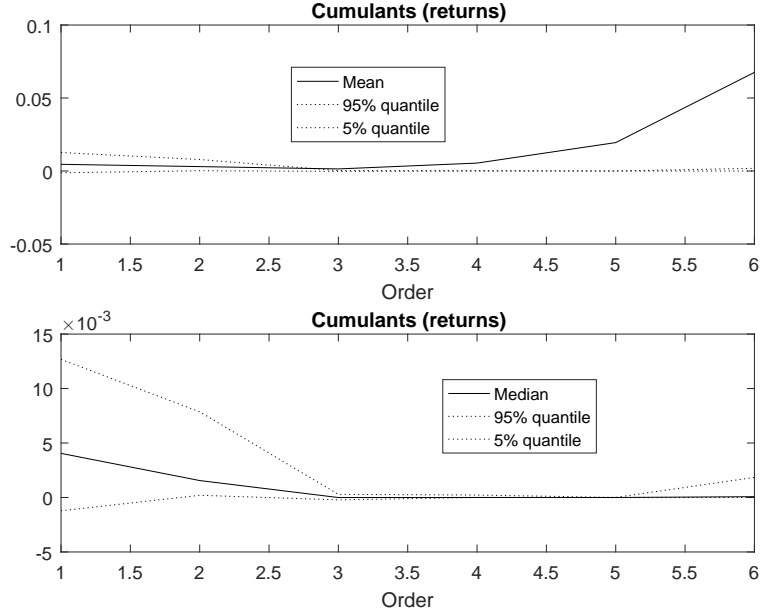


Figure 7: Cumulants in the dataset: Mean (upper panel) and median (lower panel) as well as 90% quantiles.

Values of the first 6 (integer) cumulants in the dataset as reported in Table 4. While the median cumulants decrease to a value close to zero for higher order (cf. lower panel), there are some extreme deviations for higher order cumulants as captured by high mean value exceeding the 90% interval (cf. upper panel).

on measures with an emphasis on downside risk as is common in the risk management literature.

We start with measures related to the lower partial moment. The latter is defined as

$$LPM_n(\bar{r}) = \int_{-\infty}^{\bar{r}} (\bar{r} - r)^n f(r) dr, \quad (53)$$

where  $n$  stands for the order of moment and  $\bar{r}$  is some threshold. Basically, the lower *partial* moment only considers values below the threshold value  $\bar{r}$ . We choose the mean ("average") risk-free rate  $E(r_f)$  as the threshold value. Starting from this partial mo-

ment several performance measures can be computed. The *Omega ratio* (Keating and Shadwick, 2002) is given by

$$Omega_i = \frac{E(r_i - r_f)}{LPM_1(E(r_f))} + 1. \quad (54)$$

The *Sortino ratio* (Sortino and van der Meer, 1991) is closely connected to the standard Sharpe ratio by considering the second moment

$$Sortino_i = \frac{E(r_i - r_f)}{\sqrt{LPM_2(E(r_f))}}. \quad (55)$$

The *Kappa3 ratio* (Kaplan and Knowles, 2004) uses the third partial moment

$$Kappa3_i = \frac{E(r_i - r_f)}{\sqrt[3]{LPM_3(E(r_f))}}. \quad (56)$$

A different approach is the *mean-absolute deviation* (MAD) approach (Caporin et al., 2014). Compared to the standard Sharpe ratio the excess returns are divided by the absolute deviation which is never larger than the standard deviation.<sup>20</sup> The performance measure thus reads

$$MAD_i = \frac{E(r_i - r_f)}{E[|r_i - E[r_i]|]}. \quad (57)$$

Due to the properties described in the above, it is never smaller than the Sharpe ratio.

Another measure, the Calmar ratio (Young, 1991) uses the worst (negative) return – the *maximum drawdown* (MD):

$$Calmar_i = \frac{E(r_i - r_f)}{-MD_i}. \quad (58)$$

---

<sup>20</sup>Note that if returns follow a normal distribution with standard deviation  $\sigma_i$  there is a simple linear relation:  $\frac{MAD_i}{\sigma_i} = \sqrt{\frac{2}{\pi}} \approx 0.8 < 1$  (Geary, 1935).

The most common measure used in business practice to consider tail events is the *Value at Risk*. The latter measures best expected return for the lower  $\alpha\%$  tail ( $VaR^\alpha$ ). From this measure the so-called *Dowd ratio* can be computed (Dowd, 2000):

$$Dowd_i^\alpha = \frac{E(r_i - r_f)}{VaR_i^\alpha}. \quad (59)$$

Usually, it is assumed that returns follow a normal distribution to make a statement about the right tail.<sup>21</sup> In particular, this is the case for the implementation in the software tool Matlab, which is commonly used by business practitioners. In fact, all of the above-mentioned additional risk measures are already implemented in Matlab. Thus, it is very convenient and easy to compute these performance measures.

Once again, we compute the Spearman rank correlation for the different measures for our sample of hedge funds. The results are summarized in Table 7. We are able to replicate similar results as reported in Eling and Schuhmacher (2007). That is, the standard measures (displayed in the lower right part of the table) are all highly correlated among each other ( $> 95\%$ ). Moreover, these more *sophisticated* measures exhibit a substantial correlation with the standard Sharpe ratio. Note that some of the measures have an implicit underlying normality assumption (e.g. the Dowd ratio). As compared to the data sample considered in Eling and Schuhmacher (2007) our data is more frequently rejected by formal normality tests (cf. Table 4). Nevertheless, the high correlations between the measures remain. The measures related to the lower partial moment (Omega, Sortino, Kappa3) are highly correlated among each other, yet this correlation diminishes as the gap between the order of the moments increases (e.g., comparing Omega and Kappa3). By construction the Sharpe ratio is highly correlated to the Sortino ratio as well as the MAD approach. The same holds true for the correlation between MAD and the Omega ratio.

---

<sup>21</sup>In this case the tail is computed as  $VaR_i^\alpha = -(E(r_i) + \sigma_i z_\alpha)$  where  $z_\alpha$  represents the  $\alpha$ -quantile of the standard normal distribution.

	<b>Gini</b>	<b>CRRA 3</b>	<b>CRRA 10</b>	<b>Quad</b>	<b>Sharpe</b>	<b>Calmar</b>	<b>MAD</b>	<b>Omega</b>	<b>Sortino</b>	<b>Kappa3</b>	<b>Dowd 5%</b>	<b>Dowd 1%</b>
Gini	1.00											
CRRA3	0.72	1.00										
CRRA10	0.95	0.77	1.00									
Quad	0.21	0.75	0.28	1.00								
Sharpe	0.61	0.92	0.65	0.85	1.00							
Calmar	0.53	0.89	0.61	0.87	0.98	1.00						
MAD	0.61	0.91	0.64	0.85	0.99	0.97	1.00					
Omega	0.60	0.91	0.63	0.85	0.99	0.97	1.00	1.00				
Sortino	0.57	0.91	0.64	0.87	0.99	0.99	0.99	0.99	1.00			
Kappa3	0.55	0.90	0.63	0.87	0.98	0.99	0.97	0.97	1.00	1.00		
Dowd 5%	0.58	0.90	0.63	0.85	0.97	0.96	0.97	0.97	0.97	0.96	1.00	
Dowd 1%	0.60	0.91	0.65	0.86	0.99	0.97	0.99	0.99	0.99	0.98	0.98	1.00

Table 7: Rank correlations between Atkinson-based and standard performance measures.

Spearman rank correlations between performance measures for all assets with data as reported in Table 4. We compare Atkinson-based certainty equivalents based on CRRA utility (risk aversion  $\rho = 3$  and  $\rho = 10$ ), CARA (risk aversion  $\lambda = 2$ ), and quadratic utility (**Quad**) with existing measures of Gini mean difference (**Gini**), Sharpe ratio (**Sharpe**), and standard downside-risk based performance measures. The latter include the Calmar ratio, Mean Absolute Deviation (**MAD**), measures related to lower partial moments (**Omega**, **Sortino**, **Kappa3**), and the Dowd ratio with Value at Risk values of  $\alpha = 5\%$  and  $1\%$ , respectively.

Compared to the above-mentioned levels of correlation, the correlations with the novel measures introduced in this paper are low. The standard Atkinson case (CRRA utility) with a risk aversion of 3 – in line with Morningstar (2016) – is in the area of 90%. As we increase the coefficient of relative risk aversion to ten ( $\rho = 10$ ), the correlations between the Atkinson index and the standard measures decrease – in line with the theoretical discussion below Proposition 3.3. The Gini-based performance assessment has the lowest overall correlation with standard measures. The simple quadratic utility – only taking into account the first two moments – has a low correlation with the Gini- and Atkinson-based measures, while exhibiting a higher correlation with the standard measures. All in all, our results thus suggest that the measures we propose in this paper contain new information about asset performance that is not captured by the standard approaches.

## 5.2. Market anomalies

So far we presented an example in which our novel approach presented new insights as compared to standard approaches. Given the large number of funds – and in line with the existing literature – we focused on the rank correlation as an easy-to-interpret measure of value-added. Of course, this obstructs from many details. In particular, the differential ranking might be subject to measurement noise.

In order to get a more detailed look at the workings of our measure, we consider a different and more sparse data set: rather than looking at several hundred (opaque) hedge fund strategies we consider the three most common market anomalies (size, value, and momentum) and compare them to each other as well as the market strategy.<sup>22</sup>

The deviations from the standard Capital Asset Pricing Model were introduced in Fama and French (1993) (size and value) respectively Carhart (1997) (momentum). Fama and French (1993) show that investors can earn excess returns by having a port-

---

<sup>22</sup>A similar exercise was conducted in Kadan and Liu (2014) in order to benchmark their newly proposed performance measure.

	Market	Size	Value	Momentum
<b>Mean</b> [%]	0.66	0.21	0.38	0.66
<b>Std. Dev.</b> [%]	5.35	3.21	3.50	4.71
<b>Skewness</b>	0.19	1.94	2.17	-3.06
<b>Kurtosis</b>	10.84	22.46	22.06	30.81

Table 8: Summary statistics for market anomalies.

Summary statistics for the first four moments of monthly returns for market anomalies and the market portfolio for the US stock universe from 1927 to 2017.

folio with a long position in firms with low market capitalization while shorting those with a high market capitalization (size effect). In addition, they show that one can generate excess returns by having a long position in high book-to-market firms and a corresponding short position in those with low book-to-market values (value strategy). Finally, Carhart (1997) demonstrates that a portfolio with long positions in assets with high recent returns (*winner*s) together with a short position in assets with recent low returns (*loser*s) is also able to generate abnormal returns.

Kenneth French constructs these portfolios as well as the market portfolio (returns in excess of the risk-free rate) for the US stock market and makes them freely available on his homepage.<sup>23</sup> We employ his database and consider monthly returns from the years 1927 (beginning of the datasample) until February, 2018 (the most recent observation).<sup>24</sup> Table 8 provides descriptive statistics for the four strategies and shows that all strategies exhibit excess kurtosis. Moreover, the momentum strategy also displays negative skewness making it especially prone to downside risk. All time series fail both the formal Jarque-Bera and Anderson-Darling test for normality at a 1% level.

We investigate the strategies in terms of performance using both our newly derived Atkinson index (with different utility specifications) and compare it to the classic Sharpe ratio. The results are reported in Table 9. First of all, the Sharpe ratio considers the

<sup>23</sup>It can be found at [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

<sup>24</sup>Note that in their exercise, Kadan and Liu (2014) only consider a sub-sample lasting from 1962 to 2009 avoiding extreme stock price movements such as the Great Depression. Nevertheless, we recover their main conclusion of the value strategy being more attractive under alternative performance measure and the momentum strategy losing in attractiveness.

		Market	Size	Value	Momentum
<b>Sharpe</b>		0.123 [2]	0.066 [4]	0.109 [3]	0.141 [1]
<b>CRRA</b>	3	0.222 [1]	0.067 [4]	0.210 [2]	0.196 [3]
	5	-0.095 [3]	-0.026 [2]	0.103 [1]	-0.418 [4]
	10	-1.078 [3]	-0.254 [2]	-0.156 [1]	-7.173 [4]
<b>CARA</b>	2	0.374 [2]	0.113 [4]	0.263 [3]	0.416 [1]
	4	0.080 [3]	0.019 [4]	0.154 [1]	0.084 [2]
	9	-0.753 [3]	-0.205 [2]	-0.103 [1]	-1.837 [4]
<b>Quad</b>		0.298 [2]	0.083 [4]	0.226 [3]	0.384 [1]

Table 9: Performance measures – Sharpe ratio vs. Atkinson-based.

The table reports the performance of each strategy using the Sharpe ratio as well as the Atkinson-based measures. Values within brackets show the ranking with 1 indicating the best performance.

momentum strategy the most attractive and the size strategy the least attractive of them all. The very same ranking of the four strategies is obtained with the Atkinson-based performance measure under both quadratic utility as well as CARA preferences with a risk aversion of 2. This is not surprising as they only consider the first two moments respectively place the largest weight on these moments. This, however, is not the case for CRRA utility which in general is considered as a more realistic utility function. Moreover, the ordering changes depending on the assumed value of risk aversion. In fact, for high values of risk aversion CARA and CRRA utility rankings become identical. For high levels of risk aversion, the momentum strategy eventually turns out to be the least attractive, whereas the value strategy delivers the best performance. Thus, the more sophisticated measures punish the presence of substantial downside risk (negative skewness) more heavily.

Given that we only consider four strategies in total, we end up with six possible pairwise comparisons. This also allows us to assess whether the performance measures can be statistically distinguished from each other. In Table 10, we report the pairwise comparison.<sup>25</sup> In general, the performance measure is more able to distinguish different

---

<sup>25</sup>They are computed by the ratio of the gap between the two measures and the respective standard deviation. The standard errors are bootstrapped. For the Sharpe ratio and non-normal returns the standard error is eventually known and given by  $\sqrt{\frac{1}{T}(1 + 0.5S^2 - M_3 \cdot S + 0.25(M_4 - 3)S^2)}$  with  $S$

		MA-S	MA-V	MA-MO	S-V	S-MO	V-MO
<b>Sharpe</b>		0.057 (3.151) *	0.015 (0.827)	-0.018 (-0.824)	-0.042 (-2.467) *	-0.075 (-3.561) **	-0.033 (-1.580)
<b>CRRA</b>	3	0.155 (1.752)	0.012 (0.135)	0.026 (0.199)	-0.143 (-2.365) *	-0.129 (-1.179)	0.014 (0.122)
	5	-0.069 (-0.727)	-0.198 (-2.051) *	0.322 (1.366)	-0.129 (-2.187) *	0.392 (1.755)	0.520 (2.326) *
	10	-0.825 (-5.641) **	-0.922 (-6.291) ***	6.095 (3.714) **	-0.098 (-1.555)	6.920 (4.230) **	7.017 (4.290) **
<b>CARA</b>	2	0.261 (3.085) *	0.111 (1.277)	-0.042 (-0.403)	-0.150 (-2.512) *	-0.304 (-3.536) **	-0.153 (-1.793)
	4	0.061 (0.678)	-0.074 (-0.816)	-0.004 (-0.026)	-0.135 (-2.293) *	-0.065 (-0.550)	0,070 (0.594)
	9	-0.548 (-4.802) **	-0.650 (-5.668) **	1.084 (2.192) *	-0.102 (-1.678)	1.632 (3.365) **	1.734 (3.575) **
<b>Quad</b>		0.215 (2.684) *	0.072 (0.886)	-0.086 (-0.833)	-0.143 (-2.533) *	-0.301 (-3.563) **	-0.158 (-1.841)

Table 10: Performance measures measures – Sharpe ratio vs. Atkinson-based.  
The table reports the difference between the strategies (MA: market, S: size, V: value, MO: Momentum) and the associated t-statistics (in parentheses). Asterisks denote statistical significance at the 0.1% (\*\*\*), 1% (\*\*), and 5% (\*) levels.



		<b>Market</b>	<b>Size</b>	<b>Value</b>	<b>Momentum</b>
<b>Std. Dev.</b>	[%]	5.35 [4]	3.21 [1]	3.50 [2]	4.71 [3]
<b>CRRA</b>	3	0.436 [4]	0.145 [1]	0.170 [2]	0.465 [3]
	5	0.751 [3]	0.237 [1]	0.276 [2]	1.075 [4]
	10	1.727 [3]	0.464 [1]	0.534 [2]	7.786 [4]
<b>CARA</b>	2	0.284 [4]	0.099 [1]	0.117 [2]	0.246 [3]
	4	0.576 [3]	0.192 [1]	0.225 [2]	0.577 [4]
	9	1.404 [3]	0.416 [1]	0.481 [2]	2.485 [4]
<b>Quad</b>		0.360 [4]	0.129 [1]	0.154 [2]	0.279 [3]

Table 11: Risk measures – Standard deviation vs. Atkinson-based.

The table reports the measure of risk for each strategy differentiating between the standard deviation (as used in the Sharpe ratio) and the Atkinson-based measures (all in %). Values in parentheses give the respective ranking with 1 indicating the lowest risk.

strategies under a high assumed risk aversion. For the case of CRRA utility with risk aversion of  $\rho = 3$  (as implicitly assumed in the Morningstar rating), the performance measures cannot be distinguished from each other. In general, it is difficult to distinguish the market portfolio from both the value and the momentum strategy other than for high values of risk aversion. This is also in line with the example displayed in Figure 3 for which the gap between the certainty equivalents for two given distributions substantially increases with the risk aversion. On the other hand, with a low risk aversion it might be hard to distinguish the two given distributions for example in small samples.

So far, we only considered the certainty equivalent. In Table 11, we perform a similar exercise for the corresponding risk measures, i.e., the standard deviation and the respective Atkinson indexes. In this case, only the ranking of the market and the momentum strategy switches for CRRA and CARA utility with high levels of risk aversion. For high levels of risk aversion, the momentum strategy is classified as riskier and thus loses its glamour. As shown in Table 12 it is generally harder to distinguish between the riskiness of the size and the value effect. More interestingly, both the difference between the riskiness of the market and the momentum effect is significant for the standard de-

---

being the Sharpe ratio,  $T$  the number of observations, and  $M_3$  and  $M_4$  representing the skewness respectively kurtosis (Mertens, 2002).

		MA-S	MA-V	MA-MO	S-V	S-MO	V-MO
<b>Std. dev.</b>		2.146 (13.515) ***	1.853 (11.733) ***	0.646 (3.246) **	-0.294 (-1.957)	-1.500 (-6.701) ***	-1.207 (-5.559) **
<b>CRRA</b>	3	0.291 (14.861) ***	0.266 (13.439) ***	-0.029 (-0.505)	-0.025 (-2.214) *	-0,320 (-5.708) **	-0,295 (-5,261) **
	5	0.513 (15.242) ***	0.474 (13.697) ***	-0.324 (-1.776)	-0.039 (-2,318) *	-0.838 (-4.651) **	-0.799 (-4.431) **
	10	1.263 (12.479) ***	1.193 (11.799) ***	-6.059 (-3.772) **	-0.069 (-2.352)*	-7.321 (-4.566) **	-7.252 (-4.523) **
<b>CARA</b>	2	0.186 (15.355) ***	0.168 (13.151) ***	0.038 (1.620)	-0.018 (-2.140) *	-0.147 (-6.861) ***	-0.130 (-5.896) ***
	4	0.384 (16.041) ***	0.351 (14.355) ***	0.000 (-0.007)	-0.033 (-2.371) *	-0.384 (-6.035) ***	-0.351 (-5.501) **
	9	0.988 (14.371) ***	0.923 (13.330) ***	-1.081 (-2.404) *	-0.065 (-2.442) *	-2.069 (-4.647) ***	-2.004 (-4.501) **
<b>Quad</b>		0.231 (13.110) ***	0.206 (11.351) ***	0.081 (3.146) *	-0.025 (-2.013) *	-0.150 (-6.761) ***	-0.125 (-5.529) **

Table 12: Risk measures – Standard deviation vs. Atkinson-based.

The table reports the difference between the strategies (MA: market, S: size, V: value, MO: Momentum; all in percentage points) and the associated t-statics (in parentheses). Asterisks denote statistical significance at the 0.1% (\*\*\*), 1% (\*\*), and 5% (\*) levels.

viation and an Atkinson index with high Constant Relative Risk Aversion, but the sign – indicating the ranking - is different. Consequently, the (rank) correlation between the different risk measures is higher than between the performance measures.

## 6. Conclusion

In this paper, we explore the usefulness of applying the Atkinson (1970) index – well-known in the literature on economic inequality – to financial time series of returns in order to measure financial risk. Combining the Atkinson index with the expected return and the risk-free rate, we obtain a very general performance measure, which encompasses the Morningstar index (Morningstar, 2016) as a special case. We derive closed-form solutions for a large number of combinations of preferences and return distributions, along with general formulae based on the cumulants of the return distribution. Finally, we apply our risk and performance measures to hedge fund data and well-known market anomalies (Fama-French and momentum-based strategies), and we find that our proposed class of performance measures contains additional information, not captured by existing ones.

A possible avenue for future research would be to use the Atkinson index in combination with expected returns in order to more conveniently arrive at optimal portfolios for expected utility maximizers when returns cannot plausibly be assumed to be normally distributed.

## References

- Agarwal, V. and N. Y. Naik (2004). Risks and Portfolio Decisions Involving Hedge Funds. *Review of Financial Studies* 17(1), 63–98.
- Atkinson, A. B. (1970). On the Measurement of Inequality. *Journal of Economic Theory* 2, 244–263.
- Barndorff-Nielsen, O. E. (1997a). Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. *Scandinavian Journal of Statistics* 24(1), 1–13.
- Barndorff-Nielsen, O. E. (1997b). Processes of Normal Inverse Gaussian Type. *Finance & Stochastics* 2(1), 41–68.
- Brockett, P. L. and Y. Kahane (1992). Risk, Return and Skewness Preference. *Management Science* 38, 851–866.
- Brooks, C. and H. M. Kat (2002). The Statistical Properties of Hedge Fund Index Returns. *Journal of Alternative Investments* 5(2), 26–44.
- Caporin, M., G. M. Jannin, F. Lisi, and B. B. Maillet (2014). A Survey on the Four Families of Performance Measures. *Journal of Economic Surveys* 28(5), 917–942.
- Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance* 52(1), 57–82.
- Chamberlain, G. (1983). A Characterization of the Distributions that Imply Mean-Variance Utility Functions. *Journal of Economic Theory* 29, 185–201.
- Cont, R. (2001). Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues. *Quantitative Finance* 1(2), 223–236.

- Cowell, F. (2000). Measurement of Inequality. In A. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Volume 1 of *Handbook of Income Distribution*, Chapter 2, pp. 87–166. Amsterdam: Elsevier.
- Dowd, K. (2000). Adjusting for Risk: An Improved Sharpe Ratio. *International Review of Economics & Finance* 9(3), 209–222.
- Eeckhoudt, L. and H. Schlesinger (2006). Putting Risk in Its Proper Place. *American Economic Review* 96, 280–289.
- Eeckhoudt, L., H. Schlesinger, and I. Tsetlin (2009). Apportioning of Risks via Stochastic Dominance. *Journal of Economic Theory* 144, 994–1003.
- Ekern, S. (1980). Increasing Nth Degree Risk. *Economics Letters* 6, 329–333.
- Eling, M. and F. Schuhmacher (2007). Does the Choice of Performance Measure Influence the Evaluation of Hedge Funds? *Journal of Banking & Finance* 31(9), 2632–2647.
- Evans, M., N. A. J. Hastings, and J. B. Peacock (2000). *Statistical Distributions* (3rd ed.). New York: Wiley.
- Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.
- Geary, R. C. (1935). The Ratio of the Mean Deviation to the Standard Deviation as a Test of Normality. *Biometrika* 27(3/4), 310–332.
- Gini, C. (1912). *Variabilità e Mutuabilità. Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche*. Bologna: C. Cuppini.
- Goetzmann, W., J. Ingersoll, M. Spiegel, and I. Welch (2007). Portfolio Performance Manipulation and Manipulation-proof Performance Measures. *Review of Financial Studies* 20(5), 1503–1546.

- Goetzmann, W., J. Ingersoll, M. I. Spiegel, and I. Welch (2002). Sharpening Sharpe Ratios. NBER Working Paper 9116, National Bureau of Economic Research.
- Hodges, S. (1998). A Generalization of the Sharpe Ratio and its Applications to Valuation Bounds and Risk Measures. Working paper, University of Warwick.
- Kadan, O. and F. Liu (2014). Performance evaluation with high moments and disaster risk. *Journal of Financial Economics* 113(1), 131–155.
- Kaplan, P. and J. Knowles (2004). Kappa: A Generalized Downside Risk-Adjusted Performance Measure. *Morningstar Associates and York Hedge Fund Strategies*.
- Karlis, D. (2002). An EM Type Algorithm for Maximum Likelihood Estimation of the Normal-Inverse Gaussian Distribution. *Statistics & Probability Letters* 57(1), 43 – 52.
- Keating, C. and W. Shadwick (2002). A Universal Performance Measure. *Journal of Performance Measurement* 6(3), 59–84.
- Kraus, A. and R. H. Litzenberger (1976). Skewness Preference and the Valuation of Risky Assets. *Journal of Finance* 31, 1085–1100.
- Laffont, J.-J. (1989). *The Economics of Uncertainty and Information*. Cambridge, MA: MIT Press.
- Levin, J. (2006). Choice under Uncertainty. Lecture notes, Stanford University.
- Lorenz, M. O. (1905). Methods of Measuring the Concentration of Wealth. *Publications of the American Statistical Association* 9(70), 209–219.
- Lundtofte, F. and A. Wilhelmsson (2013). Risk Premia: Exact Solutions vs. Log-Linear Approximations. *Journal of Banking & Finance* 37, 4256–4264.
- Malkiel, B. G. and S. Atanu (2005). Hedge Funds: Risk and Return. *Financial Analysts Journal* 61(6), 80–88.

- Martin, I. (2013a). Consumption-Based Asset Pricing with Higher Cumulants. *Review of Economic Studies* 80, 745–773.
- Martin, I. (2013b). The Lucas Orchard. *Econometrica* 81, 55–111.
- Mertens, E. (2002). Comments on variance of the IID estimator in Lo (2002). Working paper, University of Basel.
- Morningstar (2016). The Morningstar Rating for Funds. *Morningstar Methodology*, 1–19.
- Owen, J. and R. Rabinovitch (1983). On the Class of Elliptical Distributions and their Applications to the Theory of Portfolio Choice. *Journal of Finance* 38(3), 745–752.
- Schuhmacher, F. and M. Eling (2011). Sufficient Conditions for Expected Utility to Imply Drawdown-Based Performance Rankings. *Journal of Banking & Finance* 35(9), 2311–2318.
- Sharpe, W. F. (1966). Mutual Fund Performance. *Journal of Business* 39(S1), 119–138.
- Sharpe, W. F. (1994). The Sharpe Ratio. *Journal of Portfolio Management* 21(1), 49–58.
- Shorrocks, A. F. (1983). Ranking Income Distributions. *Economica* 50(197), 3–17.
- Smetters, K. and X. Zhang (2013). A Sharper Ratio: A General Measure for Correctly Ranking Non-Normal Investment Risks. NBER Working Paper 19500, National Bureau of Economic Research.
- Sortino, F. A. and R. van der Meer (1991). Downside Risk. *The Journal of Portfolio Management* 14(4), 27–31.
- Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *American Journal of Psychology* 15(1), 72–101.

- Stutzer, M. (2000). A Portfolio Performance Index. *Financial Analysts Journal* 56(3), 52–61.
- Yitzhaki, S. (1982). Stochastic Dominance, Mean Variance, and Gini's Mean Difference. *American Economic Review* 72(1), 178–185.
- Yitzhaki, S. and E. Schechtman (2013). *The Gini Methodology: A Primer on a Statistical Methodology*. New York: Springer.
- Young, T. (1991). Calmar Ratio: A Smoother Tool. *Futures* 20(1), 40.
- Zakamouline, V. and S. Koekebakker (2009). Portfolio Performance Evaluation with Generalized Sharpe Ratios: Beyond the Mean and Variance. *Journal of Banking & Finance* 33(7), 1242–1254.



## A. Additional Proofs

### A.1. Certainty equivalent returns and $n$ th-order stochastic dominance

Suppose a distribution  $F_A$   $n$ th-order stochastically dominates a distribution  $F_B$  and that the elementary utility function satisfies  $(-1)^k u^{(k)}(w) < 0$  for  $k = 1, 2, \dots, n$  and all  $w$ . Then, it follows from Section 5 in Ekern (1980) (see also Theorem 1 in Eeckhoudt et al. (2009)) that

$$E[u(w_0 \tilde{R}_A)] \geq E[u(w_0 \tilde{R}_B)]. \quad (60)$$

Since, by assumption,  $u$  is strictly increasing, the above implies that

$$R_{E,A}^{CE} = \frac{1}{w_0 R_f} u^{-1}(E[u(w_0 \tilde{R}_A)]) \geq \frac{1}{w_0 R_f} u^{-1}(E[u(w_0 \tilde{R}_B)]) = R_{E,B}^{CE}. \quad (61)$$

### A.2. Atkinson index for CRRA utility

For CRRA utility we know that the Atkinson index is defined as

$$A(\rho) = 1 - \frac{E[\tilde{R}^{1-\rho}]^{\frac{1}{1-\rho}}}{E[\tilde{R}]}. \quad (62)$$

It is easy to express this relation using the moment-generating function (MGF)  $\Psi_r(t) = E[\exp(tr)]$  for which  $R = \exp(r)$ . Due to the exponential transformation of the variable of interest – the geometric returns  $R$  – we can write the Atkinson index as

$$A(\rho) = 1 - \frac{E[\exp(r(1-\rho))]^{\frac{1}{1-\rho}}}{E[\exp(r)]}. \quad (63)$$

We can then exploit the relation between the MGF and the cumulant-generating function (CGF)  $k_r(t)$  given by  $k_r(t) = \ln(\Psi_r(t)) \leftrightarrow \exp(k_r(t)) = \Psi_r(t)$ . Therefore, the Atkinson index can be rewritten as

$$\begin{aligned} A(\rho) &= 1 - \frac{\Psi_r(1-\rho)^{\frac{1}{1-\rho}}}{\Psi_r(1)} = 1 - \frac{\exp(k_r(1-\rho))^{\frac{1}{1-\rho}}}{\exp(k_r(1))} \\ &= 1 - \frac{\exp\left(\frac{k_r(1-\rho)}{1-\rho}\right)}{\exp(k_r(1))} = 1 - \exp\left(\frac{k_r(1-\rho)}{1-\rho} - k_r(1)\right). \end{aligned} \quad (64)$$

Given that the certainty equivalent equals  $R^{CE}(\rho) = (1 - A(\rho))E[\tilde{R}]$ , we can write the latter as

$$R^{CE}(\rho) = \exp\left(\frac{k_r(1-\rho)}{1-\rho}\right) \quad (65)$$

and we have that

$$r^{CE}(\rho) \equiv \ln R^{CE}(\rho) = \frac{k_r(1-\rho)}{1-\rho}. \quad (66)$$

Log returns in excess of the risk-free rate  $r_f$  are thus given by

$$r_E^{CE}(\rho) = \frac{k_r(1-\rho)}{1-\rho} - r_f. \quad (67)$$

■

The specific case for returns following a log-normal distribution respectively a log-NIG distribution are easy to establish from this general result.

### A.2.1. Atkinson index for CRRA utility and log-normally distributed returns

If we assume that geometric returns  $R$  follow a log-normal distribution, arithmetic returns  $r$  will follow a normal distribution  $r \sim N(\mu_r, \sigma_r)$ . For this case, the MGF is well-known and given by

$$\Psi_r(t) = \exp(\mu_r t + 0.5\sigma_r^2 t^2), \quad (68)$$

implying the following cumulant generating function:

$$k_r(t) = \mu_r t + 0.5\sigma_r^2 t^2. \quad (69)$$

Taking this result and inserting it in the general relation given in equation (64), it is easy to verify that

$$A(\rho) = 1 - \exp(-0.5\rho\sigma_r^2). \quad (70)$$

■

### A.2.2. Atkinson index for CRRA utility and log-NIG distributed returns

In the case where geometric returns  $R$  follow a log-NIG distribution, arithmetic returns  $r$  are described by a NIG distribution ( $r \sim NIG(\hat{\mu}_r, \delta_r, \alpha_r, \beta_r)$ ) with an MGF

$$\Psi_r(t) = \exp(\hat{\mu}_r t + \delta_r(\sqrt{\alpha_r^2 - \beta_r^2} - \sqrt{\alpha_r^2 - (\beta_r + t)^2})), \quad (71)$$

implying a CGF of

$$k_r(t) = \hat{\mu}_r t + \delta_r(\sqrt{\alpha_r^2 - \beta_r^2} - \sqrt{\alpha_r^2 - (\beta_r + t)^2}). \quad (72)$$

Similar to the case with log-normally distributed returns we can insert the latter function into the general result in equation (64) to arrive at

$$A(\rho) = 1 - \exp \left\{ \frac{\delta_r}{1 - \rho} \left( \rho \sqrt{\alpha_r^2 - \beta_r^2} + (1 - \rho) \sqrt{\alpha_r^2 - (\beta_r + 1)^2} - \sqrt{\alpha_r^2 - (\beta_r + (1 - \rho))^2} \right) \right\}. \quad (73)$$

Note that the result is independent of  $\hat{\mu}_r$ . ■

### A.3. Atkinson index for CARA utility

For a utility function of the CARA type the Atkinson index is given by

$$A(\lambda) = 1 + \frac{\ln(E[\exp(-\lambda w_0 \tilde{R})])}{\lambda w_0 \mu_R}. \quad (74)$$

Similar to the proof in A.2.1 we can exploit the properties of the CGF  $k_R(t)$ . It is important to keep in mind that in this case we make assumptions about the gross returns  $R$  rather than the log transform  $r = \ln(R)$  as considered for CRRA utility. Using the definition  $k_R(t) = \ln(\Psi_R(t)) = \ln(E[\exp(tR)])$  the Atkinson index can be written as

$$A(\lambda) = 1 + \frac{k_R(-\lambda w_0)}{\lambda w_0 \mu_R}. \quad (75)$$

Using the relation  $R_E^{CE} = (1 - A(\lambda))E[\tilde{R}_E]$  together with  $E[\tilde{R}_E] = \frac{\mu_R}{R_f}$ , we can express the certainty equivalent geometric excess return as

$$R_E^{CE}(\lambda) = -\frac{k_R(-\lambda w_0)}{\lambda w_0 R_f}. \quad (76)$$

#### A.3.1. Atkinson index for CARA utility and normally distributed returns

In the case in which returns follow a normal distribution,  $R \sim N(\mu_R, \sigma_R)$ , we can insert the definition of the CGF

$$k_R(t) = \mu_R t + 0.5 \sigma_R^2 t^2, \quad (77)$$

together with the mean  $E[\tilde{R}] = \mu_R$  in the general equation presented in (75) to arrive at

$$A(\lambda) = \frac{1}{2} \lambda w_0 \frac{\sigma_R^2}{\mu_R}. \quad (78)$$

The geometric certainty equivalent excess return is thus given by

$$R_E^{CE}(\lambda) = \frac{\mu_R}{R_f} - 0.5 \frac{\sigma_R^2 \lambda w_0}{R_f}. \quad (79)$$

■

### A.3.2. Atkinson index for CARA utility and NIG distributed returns

Now, we assume that returns follow a NIG distribution,  $R \sim NIG(\hat{\mu}_R, \delta_R, \alpha_R, \beta_R)$ . Once again, we employ the general result (75) jointly with the CGF of the NIG distribution

$$k_r(t) = \hat{\mu}_R t + \delta_R (\sqrt{\alpha_R^2 - \beta_R^2} - \sqrt{\alpha_R^2 - (\beta_R + t)^2}) \quad (80)$$

as well as the general mean

$$E[\tilde{R}] = \hat{\mu}_R + \delta_R \frac{\beta_R}{\sqrt{\alpha_R^2 - \beta_R^2}} \quad (81)$$

to arrive at

$$A(\lambda) = 1 + \frac{1}{\lambda} \frac{\left( -\lambda w_0 \hat{\mu}_R + \delta_R \left( \sqrt{\alpha_R^2 - \beta_R^2} - \sqrt{\alpha_R^2 - (\beta_R - \lambda w_0)^2} \right) \right)}{w_0 \left( \hat{\mu}_R + \frac{\delta_R \beta_R}{\sqrt{\alpha_R^2 - \beta_R^2}} \right)} \quad (82)$$

The certainty equivalent geometric excess returns thus amounts to

$$R_E^{CE}(\lambda) = \frac{\hat{\mu}_R}{R_f} - \frac{\delta_R (\sqrt{\alpha_R^2 - \beta_R^2} - \sqrt{\alpha_R^2 - (\beta_R - \lambda w_0)^2})}{\lambda w_0 R_f} \quad (83)$$

■