Rationalization of Investment Preference Criteria

Jacques Pézier

ICMA Centre
Henley Business School
University of Reading

July 21, 2011

ICMA Centre Discussion Papers in Finance DP2011-12

Copyright 2011 Pézier. All rights reserved.
ABSTRACT

The majority of risk adjusted performance measures (RAPM) currently in use – e.g., Treynor ratio, $(\alpha/\beta)$ ratio, Omega index, RoVaR, ‘coherent’ preference criteria, etc. – are incompatible with any sensible utility function and would be best avoided. We argue instead for the assessment of a maximum certainty equivalent excess return ($CER^*$) criterion, or equivalent criteria, adapted to investment circumstances: alternative investments, return forecasts, and risk attitude. We explain the assessment of $CER^*$s and give three applications: performance comparisons among traditional and alternative funds, optimal design of structured products, and explanation of the credit risk premium puzzle.

Keywords: Certainty equivalent return, risk-adjusted performance measure, risk aversion, HARA utility functions, coherent risk measures, spectral indices, Sharpe ratio, generalized Sharpe ratio, information ratio, Treynor ratio, Jensen alpha, skewness, kurtosis, volatility skew, optimal structured products, credit risk premium.

JEL Classification Codes: D80, D81, G10, G11, G13

Authors Details:
Jacques Pézier
Visiting Professor, ICMA Centre, Business School,
Email: j.pezier@icmacentre.ac.uk

The University of Reading, PO Box 242
Reading RG6 6BA, United Kingdom
Tel: +44 (0)1183 786675 (ICMA Centre) Fax: +44 (0)1189 314741

Disclaimers:
The contents of this paper are presented in good faith, and neither the authors, the ICMA Centre nor the University, will be held responsible for any losses, financial or otherwise, resulting from actions taken on the basis of its contents. Any persons reading the paper are deemed to have accepted this.
BEST IS relative especially when it comes to personal preferences; there is little point debating in abstract what is the best opera, or the best wine. Even in the more mundane world of investment management where attention is mainly focused on returns, it would be unreasonable to look for a best asset mix, or a best fund, for all investors, regardless of their personal preferences and circumstances. We revisit the much-studied problem of defining performance criteria for risky investments, be they in traditional securities, hedge funds or other alternative investments, advocate a universal approach adaptable to personal circumstances, develop analytical approximations, and apply the results not only to the assessment of funds performance, but also to the design of structured products and the pricing of credit risks. We also explain why most criteria currently in use are incompatible with this approach.

Global investment performance standards (GIPS®) have been agreed to enable fair comparison of funds returns. But past returns are notably unreliable for predicting future returns. Investors are generally risk-averse and are concerned about the uncertainty in future returns. Various risk adjusted performance measures (RAPMs) have been designed to satisfy this twin purpose: to reflect statistical characteristics of past returns that are likely to persist in the future and to provide a preference scale for ranking investment opportunities. Two types of question ensue: (i) How reliable is the ex post estimate of a RAPM in predicting future performance and (ii) how relevant are specific RAPMs as ex ante preference criteria? We address in detail the latter question and only mention the first in some empirical applications.

The design of preference criteria has been approached from two directions: heuristic and axiomatic. The heuristic approach, exemplified by Sharpe and Treynor, consists in combining intuitively chosen measures of reward (desirable attribute) and risk (undesirable attribute) into a RAPM and then examining under which circumstances this criterion leads to apparently reasonable preference rankings. The appreciation of what may be desirable and undesirable attributes, as well as what may be reasonable preference rankings, is left to intuition and is therefore debatable. The other direction, the axiomatic approach, exemplified by utility theory and ‘coherent’ metrics, postulates a few axioms of rational choice and then seeks criteria that satisfy these axioms. Debates focus on whether the chosen axioms are necessary and sufficient to characterize rational behavior. In both cases, however, a preference criterion should reflect the best use an investor can make of an investment opportunity given her investment alternatives, her return forecasts, and her risk attitude. It would be futile to look for an objective ‘best’, independently of these conditions.

Investment problems are becoming ever more diverse in our growing market economies. More people of all ages, with different needs and risk appetites, have financial assets and liabilities

---

1There are many other relevant considerations such as liquidity, transparency, strategies, management experience, risk management, ethics, etc.

2(GIPS®) are published by the Chartered Financial Accountants Institute, a global not-for-profit association of investment professionals. It has a long-standing history of and commitment to establishing a broadly accepted ethical standard for calculating and presenting investment performance. The first publication of (GIPS®) dates back to April 1999. The most recent edition was issued in April 2010.

3The Squam Lake ReportFrench (2010) goes as far as recommending (Recommendation 4, p.61) that ‘Past returns should not be reported in the standardized disclosure label’ because of the ‘large body of research that past returns in general, and short-term returns in particular, are almost useless in forecasting subsequent investment performance’. Regulators often impose a warning to that effect in performance reports.

4The Squam Lake Report also recommends (Recommendation 3, p.60) that ‘The standardized disclosure should present simple but meaningful measures of risk’.
to manage. The range of investment products offered to them has widened far beyond the
traditional cash, bond, and equity classes to include actively managed funds (traditional or
hedge funds), private equity, REITs, structured products, and derivative products facilitating
the trading of more exotic risks - commodities, credit risks, volatilities and correlations, weather,
carbon emissions, etc. - some of which offer return characteristics markedly different from those
of traditional asset classes. The recent proliferation of RAPMs - Cogneau and Hübner (2009)
report 101 - may be a response to these varied circumstances, but as there are no agreed
standards stipulating which RAPM should be used under which circumstance, there is room
for confusion.

In Section I, we re-examine the best-known and most commonly used RAPM, the Sharpe
ratio (Sharpe (1966, 1994)), to explain its intuitive appeal as well as its limitations. The
Sharpe ratio, like many other RAPMs, does not refer explicitly to an allocation of resources
among alternative investments; it does not require a statement of risk aversion either. It has
the appearance of a universal, incontrovertible preference measure. But the reality is that if
some stringent conditions are not met, it can go against common sense, like not favoring the
investment that always guarantees greater returns.

In Section II, we link the Sharpe ratio to a primitive and universal satisfaction index for risky
investments, namely, the minimum sure excess return over the risk-free rate on total wealth
an investor would be ready to accept in exchange for giving up a risky investment; we call
it the certainty equivalent excess return (CER). It is a personal judgment that encapsulates
the investor’s views and risk attitude, investment opportunities and constraints. We take it as
tautological that the larger the CER, the better the investment. If an investor’s only choice
were to optimize the allocation of her entire wealth to a single risky asset and a risk-free asset,
the corresponding maximum certainty equivalent excess return, based on her personal forecast
and risk attitude would characterize the attractiveness of the risky asset to her.

A CER* is expressed on the familiar rate of return scale. Any monotonic positive transforma-
tion of a CER* generates an equivalent criterion, but on a less familiar scale. Hodges (1998)
was the first to note that a generalized Sharpe ratio (GSR) can be defined as an increasing
function of a CER* that equates the original Sharpe ratio when the return forecast for the
risky asset is normally distributed and the investor exhibits constant absolute risk aversion.
We propose in Section III a more general definition of GSR relying only on the current coef-
ficient of local absolute risk aversion of the investor and discuss the pros and cons of using a
GSR scale rather than a rate of return scale.

In Section IV, we introduce a benchmark portfolio as an investment alternative. We assess
the potential improvement from adding a risky asset (active portfolio) to the benchmark both
in terms of marginal and optimal increases in CER* and show that these measures trans-
form into equivalent generalized Jensen alpha (GJA) and generalized information ratio (GIR),
respectively. But other RAPMs related to a benchmark portfolio or to a multifactor return
model may not be equivalent to increases in CER*s. In particular, we do not find investment
conditions and utility functions that would relate the Treynor ratio and the (α/β) ratio to a
CER*.

The systematic evaluation of a CER* is based on an optimal wealth allocation, which generally
requires a numerical procedure. That is a necessity for any criterion that takes into account the personal circumstances of an investor. To avoid the numerical procedure and gain insight into preferences for skewness and excess kurtosis of returns, we develop analytical approximations for \(CER^*\)'s based on two-parameter \(HARA\) utility functions and knowledge of the first four moments of return distributions. But these approximations are not reliable when return distributions deviate markedly from normality as is almost certainly the case, for example, with options, CPPI strategies, and credit risk sensitive instruments.

In Section VI, we critically review several popular RAPMs. They may owe their popularity to their simplicity: as a rule, they do not require an explicit statement of risk attitude nor of investment alternatives, not to mention an optimization. Indeed, many appear to be universal and independent of investment size (scale invariant). But we find that most are incompatible with any sensible utility function. Many do not even reflect risk aversion or, like the expected shortfall (ES) criterion, display a kind of risk aversion that is incompatible with utility theory. We show in particular that properties used by Artzner, Delbean, Eber, and Heath (1999) to define ‘coherent’ risk metrics, are either not necessary (convexity and sub-additivity) or unsuitable (homogeneity degree one in investment size) when translated into equivalent properties for preference criteria.

In Sections VII and VIII, we apply \(CER^*\)-based criteria to ex post performance measurements. We find in Section VII that returns on major asset classes or indices, even exotic ones (e.g., hedge funds, private equity, commodities) are not so abnormally distributed that they would require more than our four-moment approximation. But we show in Section VIII that full \(CER^*\)-based criteria are critical when comparing individual funds. We assess the performance of four famous hedge funds and two well-diversified traditional funds taken individually and in combination with a total return S&P500 index fund. We find that GSRs and GIRs are more discriminating than ordinary Sharpe and information ratios and lead to significantly different preference rankings.

In Section IX we compare the performance of derivatives to that of their underlying assets and design optimal return payoffs. In Section X we find to what degree the apparently higher risk premiums observed with credit risks compared to market risks are explained by the negative skewness of returns on credit sensitive instruments.

We summarize our findings in Section XI and make two suggestions. First, that standards for fair reporting of fund performance promoted by professional associations should be extended to include \(CER^*\)-based performance criteria. These associations should define relevant standard investment conditions for various classes of investors and specify the calculation rules. Second, that risk adjusted performance criteria instead of results could serve as the basis for designing executive compensation schemes that would better align the interests of executives, investors, and shareholders.
I. Limitations of the Ordinary Sharpe Ratio

The grandfather of all RAPMs is the Sharpe ratio (Sharpe (1966)). The Sharpe ratio of a risky asset is defined as the absolute value of the expected excess return divided by the standard deviation of return for that asset at some investment horizon. It does not depend on the amount invested in the risky asset, nor on the degree of risk aversion of the investor. On the other hand, its value depends on the choice of investment horizon, so, to facilitate comparisons, a one-year horizon is usually chosen as standard. Investors use the Sharpe ratio either as an ex post measure of investment performance (subject to various assumptions and statistical errors) or an ex ante assessment (based on personal forecasts) of the merits of alternative investment opportunities. But the Sharpe ratio can be justified as a preference criterion only when the following conditions are met:

(A1) The investor is only concerned with return on total wealth at the end of the investment period and always prefers larger returns to smaller returns.

(A2) The total wealth of the investor can be optimally and freely allocated to a risk-free asset and any one of the risky assets to be compared.

(A3) The investor’s preferences are entirely determined by the knowledge of expected returns and standard deviations of returns.

(A4) Investors are risk-averse in the sense that between two investments with same expected return they always prefer the investment with the lower standard deviation of return (i.e., risk).

Then, if the Sharpe ratio of a risky asset $B$ is larger than that of a risky asset $A$, for any allocation to $A$ and the risk-free asset, there is an allocation to $B$ and the risk-free asset that yields the same expected return with a lower standard deviation. Hence, given the choice to

---

5In this paper, ‘return’ should be understood as ‘excess’ return above the risk-free rate during the relevant investment period, unless ‘nominal’ return is specified.

6The traditional definition of the Sharpe ratio refers to an expected excess return and not to its absolute value. It is consistent with the constraint of no short positions in risky assets and indicates that assets with negative Sharpe ratios are less attractive than the risk-free asset. Our definition with an absolute value is consistent with unconstrained allocations, long or short, to all assets, and recognizes that shorting some risky assets to invest more in the risk-free asset may be attractive. It is also consistent with our subsequent definitions of CER and generalized Sharpe ratio.

7It has been written that the use of Sharpe ratios is justified by either quadratic utilities or Normal return distributions. The first condition is incorrect. Quadratic utilities – which are not acceptable utility functions because they are not monotonically increasing and therefore do not satisfy the principle of non-satiation – would justify criteria that are increasing functions of the expected value of the squared return, not criteria that are homogeneous degree 1 in the expected return. The second condition is not necessary. It may also be that quadratic utility is the name given, incorrectly, to a mean-variance criterion.

8The risky and the risk-free assets can be held long or short in any amount. What constitutes a risk-free asset depends on the choice of return scale and investment horizon. For example, over the long term, one may be more concerned with real returns relative to inflation than with nominal returns, but there may be no asset yielding the chosen real return over the investment horizon.

9Alternatively, all risky assets return distributions have the same shape within a positive linear transformation, meaning that all returns $r_i$ can be written as positive linear transformations of the same return $x$: $r_i = a_i x + b_i$, $a_i > 0$. 

5
construct a portfolio with either A or B (exclusively) and the risk-free asset, any risk-averse investor should prefer to use B rather than A.

Conditions (A1) to (A4) are rarely met in practice. Condition (A1) states that the final return distribution is all that matters. In reality, the consequences of an investment are usually multifaceted and spread out over time. The investor may care about the maximum draw-down over the investment period, a regular income flow for consumption, the personal enjoyment and influence gained from the investment, ethical values, etc. Returns are usually appreciated in the context of external factors (e.g., inflation or benchmark returns) thus necessitating relative scales of appreciation. (A1) also requires that the return scale should satisfy the principle of non-satiation, that is, more is always better.

Condition (A2) states that, bar exceptional circumstances, the Sharpe ratio should not be used to compare alternative incremental investments to a current risky portfolio. Only the risk-free asset and one of the alternative investments at a time is allowed.

Condition (A3) is certainly met when comparing two-parameter return distributions of the same family (whether Normal distributions or not), but it is most unlikely to be met otherwise.

The last condition, (A4), is the most benign of the four. Most investors, whether individual or institutional, are risk-averse in the sense that they would accept to let go of a risky opportunity for less than its expected value, or, conversely, are willing to buy insurance at a premium larger than the expected covered loss. Exceptions (gambling in casinos, or failing to insure against exceptional risks) can generally be traced to incomplete valuations of consequences (e.g., failure to take into account an entertainment value, a charitable motive, an assumption of automatic insurance, . . . ) or misperceptions of risks.

If any of conditions (A1)-(A4) is not met, the Sharpe ratio may violate some basic axioms of choice under uncertainty such as the axiom of stochastic dominance, even in its strong form, as the following example illustrates.

Each of the two risky assets A and B in Figure 1 offer three possible excess returns depicted by a node with three branches. The probabilities of the outcomes are marked on the branches and the corresponding excess returns (in percentages of the investment amount) are at the tips. The outcomes of A and B are identical except for the largest excess return which is greater

---

10 The Sharpe ratio can be used to compare the merits of incremental investments only if incremental investments are not correlated with existing investments, the investor has constant absolute risk aversion, and conditions (A1), (A3) and (A4) are satisfied.

11 When return distributions are of the same functional form characterized by two parameters only, all their characteristics are the same for the same choice of parameters, but that is not the case otherwise.

12 There is strong (order 0) stochastic dominance of risky asset B over risky asset A when, under any scenario, B yields a better result than A. There is weak (order 1) stochastic dominance of risky asset B over A when, for any level of return, B offers at least the same probability and for some levels a greater probability than A of reaching these levels. In other words, under weak stochastic dominance B is not guaranteed to yield a better return than A, but is never a worse bet and sometimes a better bet than A. The axiom of stochastic dominance states that if asset B stochastically dominates asset A at least weakly, then all investors, regardless of their personal risk attitudes, should strictly prefer B to A. This is an axiom because it is not a demonstrable consequence of more fundamental principles of rational behavior; it is, in effect, taken as one of the defining characteristics of rational behavior. It is one of the axioms of choice on which utility theory is based.
with $B$ than with $A$. Therefore $B$ stochastically dominates (weakly) $A$ and $B$ should always be preferred to $A$, no matter how risk-averse the investor is.

Nonetheless, the Sharpe ratios are $10/\sqrt{200} = 0.707$ for $A$ and $14/\sqrt{568} = 0.587$ for $B$, indicating a preference for $A$ over $B$! The Sharpe ratio would still favor $A$ over $B$ if all $A$ returns were reduced by one percentage point and $A$ and $B$ were perfectly dependent; in this case $B$ would guarantee a strictly better result than $A$ (strong dominance). In this case, recommending $A$ over $B$ could be regarded as unethical. Clearly, expected values and standard deviations are not sufficient to characterize the attractiveness of these two assets because their two trinomial return distributions are not entirely characterized by the same two parameters - they have different shapes. RAPMs not based on full return distributions can lead to such irrational conclusions.

**II Special Relationship between the Ordinary Sharpe Ratio and the Maximum Certainty Equivalent Excess Return**

An investor allocates a fraction $\omega_A$ of her wealth to a risky asset $A$ and a fraction $(1-\omega_A)$ to a risk-free asset $F$. We call certainty equivalent excess return ($CER$) of this allocation over a certain investment horizon the minimum guaranteed excess return above the risk-free rate on total wealth that she would find equally attractive. Among all possible allocations of her wealth between the risk-free asset and the risky asset – assuming that both assets are available in unlimited (positive or negative, i.e., long or short) amounts – there must be an optimal allocation $\omega^*_A$ that maximizes her $CER$.\(^{13}\) To stress that the optimal allocation and the corresponding maximum certainty equivalent excess return depend on her personal forecast, $p$, and personal utility function, $u$, we write the latter as $CER^*(\omega^*_A|F; p, u)$ and specify $p$ and $u$ when appropriate.\(^{14}\) Inasmuch as a greater sure return is always preferred to a smaller sure return, $CER$'s provide a cardinal preference ranking of risky assets on the familiar rate of return scale.

\(^{13}\)Any risky asset should offer both positive and negative excess returns with finite probabilities, otherwise there would be strong stochastic dominance between the risky and the risk-free returns, hence an arbitrage situation.

\(^{14}\)The symbol $p$ stands for the investor’s personal probabilistic forecast (also called ’physical’, or ’market’, or ’real world’ probability as opposed to ’risk neutral’ probability). The symbol $u$ stands for the investor’s risk attitude as described by a utility function on the PV of final wealth. Because we consider cash flows rather than consumption patterns, and we assume the availability of a risk-free asset, we can calculate a PV by discounting at the risk-free rate and address risk preferences separately from time preferences. In Section IV we
A CER* is independent of any leverage built into a risky asset (such as a leveraged hedge fund); it is the investor who determines the optimal leverage of the risky asset according to her risk attitude. It follows that a CER* cannot be negative since the optimal allocation, long or short, to the risky asset cannot be worse than a nil allocation, i.e., a full investment in the risk-free asset. To facilitate comparisons among CER*s, it is advisable to express them on a per annum basis, as is usually done for interest rates or Sharpe ratios.\textsuperscript{15}

A CER is a primitive concept; CERs of simple investments can be assessed intuitively. But to ensure consistency in the assessment of complex investments, it is advisable to quantify return forecasts into probability distributions and risk attitudes into utility functions. This ensures that CERs can be calculated systematically and that the results are coherent with the axioms of choice underpinning utility theory.\textsuperscript{16} Intuitive assessments of CERs could violate some of these axioms and, as a result, expose the investor to being arbitraged.

The calculation of a CER is particularly straightforward under the following two conditions:

(A5) The forecast of the risky asset excess return is Normal: \( r \sim N(\mu, \sigma^2) \).

(A6) The investor’s risk attitude is described by a negative exponential utility function (exponential utility, for short) on the PV of future wealth: \( u(w) = -\exp(-w/\lambda) \), where \( \lambda > 0 \) is the current value of the coefficient of local absolute risk tolerance \( (LART) \), that is \( \lambda = L(0) = -u'(0)/u''(0) \), a constant in this instance.\textsuperscript{17,18}

We refer to conditions (A5)-(A6) as the Normal distribution, exponential utility \( (N,E) \) case, and to the corresponding CER as \( CER(\omega|F;N,E) \). With an optimal allocation \( \omega^* = \lambda \mu/\sigma^2 \) to the risky asset and \( (1 - \omega^*) \) to the risk-free asset (see Appendix A), the \( CER(\omega|F;N,E) \) reaches a maximum equal to:

\[
CER^*(\omega^*|F;N,E) = \frac{1}{2}\lambda(\mu/\sigma)^2
\]

Therefore the ordinary Sharpe ratio, \( SR \), is a monotonically increasing function of \( CER^*(\omega|F;N,E) \), namely:

\[
SR \equiv (\mu/\sigma) = \sqrt{(2/\lambda)CER^*(\omega|F;N,E)}.
\] (1)

The two criteria are equivalent in the sense that they always lead to the same preference

\textsuperscript{15}As with interest rates, CER*s may vary with the investment horizon but can always be expressed on an equivalent per annum scale.

\textsuperscript{16}For a presentation of the axioms of choice supporting utility theory see Von Neumann and Morgenstern (1944) or Savage (1954). Utility theory can be regarded as a method consistent with these axioms for inferring CERs of complex investments from CERs of simple investments.

\textsuperscript{17}We favor this parameterization of the exponential utility function with a \( LART \) coefficient expressed in units of present wealth (or as a fraction of initial wealth if, without loss of generality, initial wealth is taken as one unit of wealth). Nonetheless, many authors prefer to use the inverse coefficient, the Arrow-Pratt coefficient of local absolute risk aversion, defined as \( \gamma = -u''/u' \). A finite positive \( \lambda \) indicates risk aversion.

\textsuperscript{18}Alternatively, one could estimate the utility of future wealth at some investment horizon with the relevant coefficient of risk tolerance at this horizon. Both analyses are equivalent if the coefficient of risk tolerance increases in time at the rate used for discounting.
rankings among risky assets. Condition (A5) implies that \( CER^*(\cdot|F; N, E) \) is a function of only two distributional characteristics: expected return and standard deviation of return. Condition (A6) specifies this function and the role played by the coefficient of absolute risk tolerance. Both criteria are consistent with conditions (A1) to (A6).

For example, if the excess return for asset \( A \) in Figure 1 were normally distributed (which it is not), an investor with a coefficient of absolute risk tolerance of, say, \( \lambda = 0.16 \) (i.e., 16% of total wealth) would want to invest \( \omega_A^* = 0.16(0.10/0.02) = 0.80 \), that is, 80% of her wealth in \( A \) to achieve

\[
CER^*(\omega_A^*|F; N, E) = \frac{1}{2} (0.16) (0.1)^2/0.02 = 4%
\]

meaning that she would require a minimum riskless excess return of 4% on her total wealth for giving up the optimal investment of 80% of her wealth in \( A \). Alternatively, if the investment in \( A \) was constrained to a sub-optimal size, say \( \omega_A = 40\% \) of wealth, the \( CER^*(\omega_A|F; N, E) \) of this investment would be less – only 3% in this case – whereas the Sharpe ratio of asset \( A \) remains the same.

Figure 2 displays \( CER(\omega_A|F; N, E) \) and the expected return of an investment \( \omega_A \) (as a fraction of total wealth) in asset \( A \). It confirms that for small investments compared to risk tolerance \( (\omega_A \ll \lambda) \) the marginal increase in \( CER(\omega_A|F; N, E) \) is equal to the expected return of the investment; therefore any asset with positive expected return is attractive in small quantities. As the investment amount increases, the difference between expected return and \( CER(\omega_A|F; N, E) \), or risk premium, increases like the square of the investment amount and reaches a maximum for an optimal allocation, in this case \( CER^*(\omega_A^*|F; N, E) = 4\% \) for \( \omega_A^* = 80\% \) of wealth. This maximum characterizes the attractiveness of asset \( A \), assuming it can be invested in optimally. For larger allocations to \( A \), \( CER(\omega_A|F; N, E) \) decreases and eventually becomes negative, indicating that larger investments in \( A \) eventually become unattractive because too risky.

A positive risk premium for any (non-zero) amount invested in a risky asset is the defining mark of risk aversion (a more general definition than condition (A4)). It implies concavity of the utility function (decreasing marginal utilities). The exponential utility function (A5) is an example of concave utility function. Combined with a normally distributed risky asset return, it produces a risk premium that is proportional to the square of the invested amount. This is a special case of a general property: with all twice-differentiable, concave utility functions, the risk premium of a small risky investment increases with the square of the size of the investment (see proof in Appendix B).

It is of the essence of a preference criterion to be able to reflect risk attitude; risk aversion, in particular, explains the role of insurance, the observation of positive implied credit spreads, the popularity of capital guaranteed investment products, and the general observation that the riskier a security, the higher the anticipated expected excess return is. Cognitive psychologists have confirmed that risk aversion is common among investors, at least when they consider investments with positive expected returns (Kahneman and Tversky (1979)).
III Relationship between Generalized Sharpe Ratio ($GSR$) and Maximum Certainty Equivalent Excess Return

The ordinary Sharpe ratio is consistent with conditions (A1) to (A6) because it is a monotonically increasing function of $CER^*(F; N, E)$, but it may lead to blatant contradictions when some of these conditions are not met, which is often the case in reality. For example, some investors are prepared to take bigger risks when they become wealthier, and vice versa, thus contradicting the constant absolute risk tolerance characteristic of exponential utility functions. There are also many reasons why return forecasts may not be normally distributed; for example, they could be driven by the occurrence or non-occurrence of a single major event (a discovery, a bankruptcy, an election, . . .) and be bimodal. There is evidence that return distributions for traditional securities (e.g., shares) may not differ much from log-normality over the long term, but exhibit jumps and high excess kurtosis over the short term. Returns on alternative investments often exhibit greater departures from normality. For example, default risks on corporate bonds cause negative skewness. Conversely, call options or capital guaranteed structured investment products are specifically designed to provide positively skewed return.

To illustrate, the $CER^*(F; N, E)$s, of assets $A$ and $B$ in Figure 1, calculated with a coefficient of absolute risk tolerance $\lambda = 0.16$, were found to be 4% and 2.76%, respectively, showing, wrongly, that $A$ should be preferred to $B$. In this case, the assumption of normality of return distributions is clearly incorrect. When the forecast return distributions of $A$ and $B$ in Figure 1 are taken into account, together with the same exponential utility, the $CER^*(F; N, E)$s, are 3.82% and 4.13% for $A$ and $B$, respectively, showing correctly that $B$ should be preferred to $A$. 

Figure 2: Certainty equivalent excess return of investing in risky asset $A$, assuming normally distributed return and exponential utility ($\lambda = 0.16$)
<table>
<thead>
<tr>
<th>Performance Measures</th>
<th>Asset A</th>
<th>Asset B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asset Certainty Equivalent returns:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$CER^*(\cdot</td>
<td>F;N,E)$ (optimal investment)</td>
<td>4.00% (80%)</td>
</tr>
<tr>
<td>$CER^*(\cdot</td>
<td>F;p,E)$ (optimal investment)</td>
<td>3.82% (74%)</td>
</tr>
<tr>
<td>Equivalent Generalized Sharpe ratios:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$GSR(\cdot</td>
<td>F;N,E)$</td>
<td>0.707</td>
</tr>
<tr>
<td>$GSR(\cdot</td>
<td>F;p,E)$</td>
<td>0.587</td>
</tr>
</tbody>
</table>

These results are summarized at the top of Table I, together with the corresponding optimal investment amounts.

Hodges (1998) defines an exponential utility-based Generalized Sharpe Ratio ($GSR$) as

$$GSR(\cdot|F;p,E) \equiv \sqrt{-2\log (-EU^*(\cdot|F;p,E))},$$

where $EU^*(\cdot|F;p,E)$ stands for maximum expected utility with personal distribution and exponential utility. This definition relies on the relationship between expected utility and certainty equivalent for an exponential utility function.

We propose the following more general definition of a $GSR$:

$$GSR(\cdot|F;p,u) \equiv \sqrt{(2/L(0)) CER^*(\cdot|F;p,E)},$$

(2)

where $L(0)$ stands for the current value of the LART of the investor, that is, $L(0) = -u'(0)/u''(0)$. With a Normal return distribution and an exponential utility function, the $GSR$ is equal to the ordinary Sharpe ratio but in general will take on different values under different conditions. For example, returning to the example in Figure 1, with an exponential utility the $GSR$s are 0.587 and 0.718 for $A$ and $B$, respectively, as shown at the bottom of Table I.

Remarkably, $EU^*(\cdot|F;p,E)$ and $GSR(\cdot|F;p,E)$ are independent of the investor’s constant LART equal to $\lambda$ whilst $CER^*(\cdot|F;p,E)$ is proportional to it. Indeed, the expected utility of investing an amount $\omega$ in a risky investment yielding excess return $r$ depends on $\omega/\lambda$ and not on $\omega$ and $\lambda$ separately. Therefore, for any level of risk tolerance $\lambda$ there is an optimal investment $\omega^*$ leading to the same maximum expected utility, independent of $\lambda$. This maximum expected utility is, by definition, the utility of $CER^*(\cdot|F;p,E)$ which is a function of $CER^*(\cdot|F;p,E)/\lambda$; $CER^*(\cdot|F;p,E)$ is therefore proportional to $\lambda$.

This is true for all utility functions of excess return $r$ and current LART $L(0)$ that depend on ($r/L(0)$) and not on $r$ and $L(0)$ separately. The class of hyperbolic absolute risk aversion (HARA) utility functions we use later in Section VI is a sub-set of these utility functions.
Therefore the corresponding $GSR$s are independent of the investor’s $L(0)$. With $HARA$ utilities, $GSR$s depend only on the sensitivity of the $LART$ to changes in wealth. If this sensitivity were the same for all investors and all investors made the same return forecasts, then they would all prefer the same global risky asset mix; it would be the ‘market’ portfolio of all investable assets; Tobin’s two-fund separation theorem would hold as Cass and Stiglitz (1970) proved.

Ordinary Sharpe ratios are popular, despite being expressed on a dimensionless scale, because they are simple to calculate: they do not require a statement of risk attitude nor an optimization. They appear, superficially, to be usable by any risk-averse investor to gauge the attractiveness of any risky asset. But this is true only in the special $(N, E)$ case. Conversely, $GSR$s have general applicability, but are more difficult to calculate. Yet, for investors with $HARA$ utilities, they do not require a statement of $L(0)$. $CER*$s are equivalent criteria to $GSR$s expressed on the more familiar rate of return scale, but they require a statement of $L(0)$. An individual investor with known $L(0)$, may therefore prefer to use a $CER*$, whereas the $GSR$ scale may be preferred for public reporting.

Unfortunately, few firms or individuals are prepared to make quantitative probabilistic forecasts and to encode their risk attitudes in utility functions, the two necessary ingredients for systematic calculations (as compared to intuitive assessments) of $CER*$-based criteria. It is not so much the logic of the approach$^{19}$ that is in question as the reluctance of decision makers to express their personal risk attitudes and forecasts.$^{20}$

### IV Maximum $CER$ with a Benchmark Portfolio: Generalized Information Ratio and Jensen Alpha

The $CER*$s and $GSR$s evaluated so far express the attractiveness of a risky asset when the only alternative is a risk-free asset. But investors often consider new opportunities as modifications to their current investment portfolio or to a benchmark investment rather than as exclusive alternatives. Active portfolios are defined as departures from a benchmark portfolio (a market index or a peer average). The $CER*$ concept can be extended to assess the maximum contribution of an active allocation to a benchmark portfolio.

In Appendix C we determine the optimal allocation between a risk-free asset $F$, a benchmark portfolio $B$, and an active portfolio $A$, all available in any long or short amounts. At first we assume the $(N, E)$ case: a joint Normal distribution for the returns of the active portfolio and the benchmark portfolios with marginal distributions $r \sim N(\mu, \sigma^2)$ and $r_B \sim N(\mu_B, \sigma^B)$.

---

$^{19}$Few investors would argue against the axioms of utility theory, but their instinctive choices may violate them. Cognitive psychologists have shown that many investors behave as if they were more concerned about changes in wealth than absolute wealth (see Kahneman and Tversky (1979), on prospect theory and the review by Machina (1987)). More generally, the scope of utility theory may be limited because decision problems in the real world are not well-defined but re-formulated as they are re-solved. Alternatives and consequences are discovered in a continuous quest and themselves create new objectives, or levels of aspiration. A theory of bounded rationality was developed by Simon (1957) to address these features and then further developed by Sauerman and Selten (1962) as a theory on non-optimizing boundedly rational behavior that has become known as aspiration adaptation theory (Selten (1998)). However, the investment performance measures we are addressing are sufficiently simple and well-defined to fall within the scope of utility theory.

$^{20}$Ironically, it is often the proponents of behavioral finance who criticize rational methods that rely on subjective inputs and favor instead ad hoc methods that do not require personal judgments.
respectively and correlation coefficient \( \rho \). We call \( \alpha \) and \( \beta = \rho \sigma / \sigma_B \) the intercept and the slope of the OLS regression line of the active return against the benchmark return and call \( \sigma_x = \sigma \sqrt{1 - \rho^2} \) the standard deviation of the active return relative to the benchmark portfolio return (specific risk). The investor has constant absolute risk tolerance, \( \lambda \). In this case, the optimal allocations to the active and benchmark portfolios are found to be \( \omega^* = \lambda \alpha / \sigma_x^2 \) and \( \omega_B^* = \lambda \mu_B / \sigma_B^2 - \beta \sigma_x^2 \), respectively. The maximum increase in certainty equivalent return contributed by the active portfolio, denoted \( \Delta CER^*(A|F; B; N; E) \) to specify the conditions including the presence of a benchmark portfolio, is:

\[
\Delta CER^*(\omega_A^*|F; B; N; E) \equiv CER^*(\omega_A^*|F; B; N; E) - CER^*(\omega_B^*|F; N; E) = \frac{1}{2}(\alpha / \sigma_x)^2, \tag{3}
\]

where \( \alpha / \sigma_x \) is recognized as the information ratio\(^{21} \) of the risky asset first introduced by Treynor and Black (1973). So, in the same way as we defined a generalized Sharpe ratio in (8), we define a generalized information ratio as:

\[
GIR(\cdot|F; B; p; u) \equiv \sqrt{(2/L(0))\Delta CER^*(\cdot|F; B; p; u)}, \tag{4}
\]

where \( \Delta CER^*(\cdot|F; B; p; u) \) is the increase in maximum \( CER \) achieved by adding an active portfolio to the existing optimal allocation to a benchmark \( B \) and the risk-free asset \( F \). Thus, the GIR is a preference criterion for ranking the attractiveness of active portfolios combined with a given benchmark portfolio, whatever their joint return distribution may be. In the special \((N, E)\) case the GIR equals the ordinary information ratio.

But several other RAPMs relative to a benchmark are being used and we can seek whether there are investment conditions under which they would be equivalent to a \( CER^* \). There are three particularly well-known relative performance criteria when the benchmark portfolio is a highly diversified ‘market’ portfolio so that specific risks of individual assets in the portfolio can be regarded as negligible compared to the total (‘systematic’) risk of the portfolio. Jensen (1969) proposed \( \alpha \) as a measure of absolute attractiveness of a risky asset in this diversified market portfolio context. Under market equilibrium conditions, we would expect the Jensen \( \alpha \) to be nil for all assets, nonetheless, in a dynamic environment some assets might be forecast to yield momentarily positive or negative alphas. The \( (\alpha / \beta) \) ratio, which measures \( \alpha \) per unit of systematic risk, is another intuitive measure of attractiveness for assets out of equilibrium, at least if added in moderate quantities to the market portfolio so as not to compromise its high degree of diversification. It has been hailed as a natural extension of the Sharpe ratio when a diversified, near equilibrium market portfolio is available. The Treynor ratio (Treynor (1965)), in its simplest form \( (\mu / \beta) \), is yet another relative performance RAPM still much in use.

We have just shown that, when adding a risky asset to a market portfolio and a risk-free asset and re-optimizing the three allocations, the contribution of the additional risky asset to total systematic risk is offset so that the additional risky asset contributes only its specific risk. That is why we find the information ratio rather than the Treynor or the \( (\alpha / \beta) \) ratio to be a relevant performance measure in these circumstances. Would we obtain the Treynor or the \( (\alpha / \beta) \) ratio if we did not re-optimize the allocation to the market portfolio but kept it at its optimum \( M^*_p \) when the risk-free asset is the only alternative? We show in Appendix C that the optimal

\(^{21}\) Also known as appraisal ratio.
allocation to the new risky asset would be only $\omega^* = \lambda \alpha / \sigma^2$ and the increase in $CER$ only:

$$\Delta CER^*(\cdot | F, M_F^*; N, E) = \frac{1}{2} (\alpha / \sigma)^2.$$ 

Therefore $(\alpha / \sigma)$ would be the relevant preference criterion in these circumstances. Indeed, we have not been able to find generic investment conditions under which the Treynor or the $(\alpha / \beta)$ ratio would reflect the optimum contribution of a new risky asset to a market portfolio.

If instead of looking for optimal $CER$ contributions with different strategies, we consider the marginal $CER$ contribution of adding a small quantity of risky asset to an optimal market portfolio and risk-free asset mix, denoted $(M_F^*; F)$, we find that the Jensen $\alpha$ is the marginal contribution measure,

$$\alpha = CER^*(\cdot | F, M_F^*; N, E).$$

A generalized Jensen alpha ($GJA$) can be defined by removing the condition of Normal joint return distributions and exponential utility, that is:

$$GJA(\cdot | F, M_F^*; p, u) \equiv CER^*(\cdot | F, M_F^*; p, u). \quad (5)$$

V Analytical Approximations of $CER^*$ and $GSR$ with HARA Utility Functions and First Four Moments of Return Distributions

Closed form analytical solutions for the optimization of a $CER$-based criterion are available only in special circumstances, the $(N, E)$ case in particular. Analytical approximations for small deviations from the $(N, E)$ case can be developed in terms of a few moments of the return distribution and a few parameters describing risk attitude. To this end we use the general class of two-parameter HARA utility functions and characterize return distributions with their first four moments.

A two-parameter HARA utility function on the PV of future wealth, $w$, can be written as follows:

$$u(w) \equiv \text{sign} (\eta - 1) \left( 1 + \left( \frac{\eta}{\lambda} \right) (w - w(0)) \right)^{(1 - 1/\eta)}; \quad \lambda > 0, \quad \eta \geq 0. \quad (6)$$

$\lambda$ is equal to $L(0)$, the LART at the current expected wealth $w(0)$; $\eta$ represents the sensitivity of the LART to changes in wealth. As a HARA utility is a function of wealth per unit of $\lambda$, the corresponding $CER^*$s are proportional to $\lambda$ and the corresponding $GSR$s are independent of it; but both are affected by $\eta$.\(^{23}\)

\(^{22}\)In the following relationships, we indicate that the allocation to the market portfolio is maintained at its optimum with the risk-free asset by writing it as $M_F^*$.

\(^{22}\)This marginal contribution is independent of any small change in the market allocation around $M_F^*$. The marginal contribution of a risky asset varies linearly with the initial allocation to the market portfolio. It is equal to the expected excess return $\mu$ when the initial portfolio is entirely in the risk-free asset, that is, $\mu = CER^*(\cdot | F, M_F^*; N, E)$. It could be nil or negative, even if the risky asset has positive $\alpha$, if the initial portfolio $(F, M)$ is not optimal.

\(^{23}\)With HARA utilities, the LART is proportional to wealth, therefore its inverse, the local absolute risk aversion coefficient, is inversely proportional to wealth, hence the name hyperbolic absolute risk aversion utility function.
Without loss of generality, the initial expected wealth can be taken as the unit of wealth and \( \lambda \) expressed per unit of wealth. Then, in terms of period excess return, \( r = w - 1 \), a HARA utility is expressed as:

\[
u(r) = \text{sign} (\eta - 1) \left( 1 + \left( \frac{\eta}{\lambda} \right) r \right)^{(1-1/\eta)}, \quad \eta \geq 0, \quad \lambda > 0.
\]

With this parameterization the LART is \( L(r) = (\lambda + \eta r) \). The exponential utility \( u(r) = -\exp(-r/\lambda) \) is a special case of HARA utility when \( \eta \downarrow 0 \). Other special cases include the hyperbolic \( (\eta = 0.5) \), the logarithmic \( (\eta \rightarrow 1) \), and the square root \( (\eta = 2) \) utility functions.

The conditions \( \eta \geq 0, \quad \lambda > 0 \) are necessary to ensure that a HARA utility function is monotonically increasing in \( r \). When \( \eta > 0 \), there is a low level of wealth/return at which risk tolerance becomes nil. The CER of any distribution giving a finite probability of reaching this level or falling below it - as, for example, a Normal distribution would - is infinitely negative. Thus a wide range of frequently assumed return distributions would be considered worse than the risk-free asset.\(^{24}\) However, any return distribution with positive expected excess return and a finite lower limit could be an attractive investment, at least in some small quantity. We illustrate this by comparing the attractiveness of risky assets \( A \) and \( B \) in Figure 1 with a two-parameter HARA utility function.

Table II shows the \( CER^* (A; p, HARA) \) and \( GSR (A; p, HARA) \) together with the optimal investment amount \( \omega^* \) for HARA utility functions with \( \lambda = 0.16 \) and \( \eta = 0, \quad 0.5, \quad 1, \quad \) and \( 2 \), respectively. With an exponential utility function \( (\eta = 0) \) the results are identical to those in Table I. For \( \eta \geq 0 \), asset \( B \) ranks higher than asset \( A \), as one would expect from the stochastic dominance of \( B \) over \( A \). The maximum \( CER \) for \( A \) decreases when \( \eta \) increases. The maximum \( CER \) for \( B \) increases at first and then decreases when \( \eta \) increases.

\(^{24}\)An economic agent cannot sustain losses larger than his net worth; other agents have to bear extra losses. HARA utility functions force us to consider return distributions with finite lower limits and to dismiss others as too risky. However, it could be more realistic to assume that utility functions should have a finite lower limit reflecting the maximum pain that can be sustained by the agent when losses exceed his net worth.

### Table II: Preference criteria with HARA utilities, personal distributions, and four-moment approximations

<table>
<thead>
<tr>
<th>HARA utilities</th>
<th>Exponential ( (\eta=0) )</th>
<th>Hyperbolic ( (\eta=0.5) )</th>
<th>Logarithmic ( (\eta=1) )</th>
<th>Sq. root ( (\eta=2) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda = 0.16 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Asset A:</strong> ( \omega^*_A )</td>
<td>0.739</td>
<td>0.673</td>
<td>0.566</td>
<td>0.376</td>
</tr>
<tr>
<td>( GSR (A; F; p, HARA) )</td>
<td>0.691</td>
<td>0.686</td>
<td>0.660</td>
<td>0.585</td>
</tr>
<tr>
<td>( CER^* (A; F; p, HARA) )</td>
<td>3.82%</td>
<td>3.76%</td>
<td>3.48%</td>
<td>2.74%</td>
</tr>
<tr>
<td>( ACER^* (A; F; p, HARA) )</td>
<td>3.75%</td>
<td>3.25%</td>
<td>2.00%</td>
<td>-2.75%</td>
</tr>
<tr>
<td><strong>Asset B:</strong> ( \omega^*_B )</td>
<td>0.700</td>
<td>0.700</td>
<td>0.577</td>
<td>0.381</td>
</tr>
<tr>
<td>( GSR (B; F; p, HARA) )</td>
<td>0.718</td>
<td>0.748</td>
<td>0.744</td>
<td>0.678</td>
</tr>
<tr>
<td>( CER^* (B; F; p, HARA) )</td>
<td>4.13%</td>
<td>4.48%</td>
<td>4.42%</td>
<td>3.33%</td>
</tr>
<tr>
<td>( ACER^* (B; F; p, HARA) )</td>
<td>3.62%</td>
<td>3.60%</td>
<td>3.63%</td>
<td>3.79%</td>
</tr>
<tr>
<td>( ACER^* (B; F; p, HARA) )</td>
<td>3.44%</td>
<td>3.47%</td>
<td>3.07%</td>
<td>0.94%</td>
</tr>
</tbody>
</table>
To characterize small departures of a return distribution from normality, we use its first four moments. Returns from traditional asset classes show minor deviations from normality (see Section VII), but large deviations can be expected with credit sensitive assets (rare extreme losses) and some rule-based dynamic investment strategies (e.g., stop-loss, option replication, CPPI funds).

We now approximate the expected utility of investing a quantity $\omega$ of a risky asset returning $r$ using a HARA utility function and the moments of $r$. We take a Taylor series expansion of the utility around the expected return $\mu = E[r]$, calculate the expected utility, maximize it over the choice of $\omega$, and derive values for $CER^*$ ($|F;p,HARA)$ and $GSR$ ($|F;p,HARA)$. Calling $\mu_i$ the $i^{th}$ order ($i > 1$) centered moment of $r$, that is $\mu_i = E[(r - \mu)^i]$, the expected HARA utility of investing an amount $\omega$ in the risky asset is, to the $k^{th}$ moment:

$$E[u(\omega r)] = \frac{z^{(1-\eta/\eta)}}{(\eta - 1)} \left[ 1 + k \sum_{i=2}^{k} \left( \prod_{j=1}^{i} (1 + (j - 2) \eta) \right) \left( \frac{-\omega}{\lambda z} \right)^{i \mu_i} \right],$$

with $z = 1 + \eta \omega \mu / \lambda$. The expected utility is a function of $(\omega / \lambda)$; therefore, maximizing it over the investment amount $\omega$ yields a maximum expected utility independent of $\lambda$, a $CER^*$ proportional to $\lambda$ and an equivalent $GSR$ independent of $\lambda$. Explicitly, to the fourth moment, the expected utility is:

$$E[u(\omega r)] = \frac{z^{(1-\eta/\eta)}}{(\eta - 1)} \left[ 1 + (1 - \eta) \left( \frac{\omega}{\lambda z} \right)^2 \mu_2 - (1 - \eta) (1 + \eta) \left( \frac{\omega}{\lambda z} \right)^3 \mu_3 + (1 - \eta) (1 + \eta) (1 + 2\eta) \left( \frac{\omega}{\lambda z} \right)^4 \mu_4 \right].$$

From its maximum, $EU^*$, we derive

$$CER^* (|F;m_4,HARA_4) = (\lambda/\eta) \left[ (EU^*)^{\eta/(\eta-1)} - 1 \right]$$

and, according to (8),

$$GSR (|F;m_4,HARA_4) = \sqrt{2.CER^* (|F;m_4,HARA)/\lambda}.$$

In (7) and (8) the return distribution and utility function previously denoted generically by $p$ and $u$ are now specified as $m_4$ and $HARA_4$ to indicate the first four-moment approximation and the choice of utility function.

These preference criteria for assets $A$ and $B$ are also presented in Table II. We observe that the four-moment $CER^*$'s and $GSR$s are closer to the values obtained with the personal distributions than to those obtained with incorrectly assumed Normal distributions. However, they still indicate, disappointingly, that investment $A$ is more attractive than investment $B$; therefore the four-moment approximation violates the axiom of stochastic dominance. Replacing the numerical optimization by an analytical approximation cannot be expected to improve these results.
This should not surprise us. The infinite Taylor series expansion converges towards the exact expected utility for any finite investment, \( \omega \), but a finite series may be a poor approximation. There is no guarantee that a Taylor series expansion of expected utility to the fourth order is accurate and that the resulting criterion does not violate the axiom of stochastic dominance. There is no guarantee either that the accuracy of the method would increase by taking a few higher order terms into account. As a counter-example, with exponential utilities and a fifth order expansion, the maximum expected utility for asset \( A \), which has a symmetrical return distribution, remains unchanged, but the fifth moment of asset \( B \) is sufficiently large that the approximate maximum expected utility becomes positive – which should be impossible with an exponential utility function. Consequently, there are no five-moment \( CER^* \) and \( GSR \) for asset \( B \).

There are indeed two intrinsic difficulties with moment-based criteria. First, a probability distribution may not be sufficiently well characterized by its first few moments. For example, one could fit four moments with a trinomial, a Johnson, a Normal-inverse Gaussian, a Gram-Charlier density (Jondeau and Rockinger (1999)), a maximum entropy distribution, or any number of other distributions with finite or infinite higher moments. Second, calculating expected utilities based on a \( k \)th order Taylor series expansion amounts to using a \( k \)th order polynomial utility function; but there is no sensible risk-averse polynomial utility function, unless returns are capped. If \( k \) is odd and greater than 1, the utility function displays risk-averse behavior at one end of the range of returns and risk seeking behavior at the opposite end. If \( k \) is even and greater than 1, the term of order \( k \) would have to have a negative coefficient to reflect risk aversion for increasing returns, but beyond some large return, utilities would start decreasing, which would contradict the monotonicity requirement for utility functions.\(^{25}\) The domain of such utility functions should therefore be limited on the upside.

Knowing that moment-based utility criteria are intrinsically inadequate, one might as well use a closed form analytical approximations to the four-moment maximum \( CER^* \) and \( GSR \) to avoid a numerical optimization and shed some light on preferences for skewness and excess kurtosis. Keeping to \( 4 \)th order terms in \( \omega \) and \( z \), we obtain the approximation:

\[
ACER(\omega | F; m_4, HARA_4) = \omega \mu - \lambda z \left\{ \left( \frac{\omega}{\lambda z} \right)^2 \frac{\mu_2}{2!} - \left( \frac{\omega}{\lambda z} \right)^3 \frac{(1 + \eta) \mu_3}{3!} + \left( \frac{\omega}{\lambda z} \right)^4 \left[ \frac{(1 + \eta) (1 + 2\eta) \mu_4 - 3 \mu_2^2}{4!} \right] \right\}.
\]

If the return forecast distribution is near Normal and the Sharpe ratio of the risky asset is smaller than 1, we can assume that the optimal investment is not too different from the \((N, E)\) case, that is, \( \omega^* = \lambda \mu / \mu_2 \). With the notations \( \sigma^2 = \mu_2 \) for variance, \( \xi = \mu_3 / \sigma^3 \) for skewness, \( \kappa = (\mu_4 / \sigma^4 - 3) \) for excess kurtosis, and \( SR = \mu / \sigma \) for the ordinary Sharpe ratio, we obtain

\(^{25}\)The principles of non-satiation of the return objective and of weak stochastic dominance imply that utility functions must be increasing functions of wealth or return. The axioms of utility theory do not impose other conditions on the shape of a utility function. However, since most investors are risk-averse, their utility functions should be concave, at least down to a point of maximum sustainable loss. It should also be continuous in the absence of critical levels of wealth. We therefore consider as sensible for the purpose of gauging investment performance, any utility function that is twice-differentiable with 1st order derivative \( u' \geq 0 \) and 2nd order derivative \( u'' \leq 0 \).
Table III: Approximate generalized Sharpe ratios for four HARA utility functions

<table>
<thead>
<tr>
<th>HARA utility function</th>
<th>Approximate generalized Sharpe ratio (AGSR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential (η = 0)</td>
<td>$SR[1 + SR\xi/6 - SR^2\kappa/24]$</td>
</tr>
<tr>
<td>Hyperbolic (η = 0.5)</td>
<td>$SR[1 + SR\xi/4 - SR^2\kappa/8]$</td>
</tr>
<tr>
<td>Logarithmic (η = 1)</td>
<td>$SR[1 + SR\xi/3 - SR^2(2\kappa + 1)/8]$</td>
</tr>
<tr>
<td>Square root (η = 2)</td>
<td>$SR[1 + SR\xi/2 - SR^2(5\kappa + 6)/8]$</td>
</tr>
</tbody>
</table>

the following approximation for the maximum CER:

$$ACER^* (|F; m_4, HARA_4) \approx \frac{1}{2} \lambda SR^2 \left\{1 + \frac{1}{3} SR (1 + \eta) \xi - \frac{1}{12} SR^2 \left[(1 + \eta) (1 + 2\eta) \kappa + 3\eta (2\eta - 1)\right]\right\}. \tag{9}$$

The corresponding GSR, according to (8), is approximately:

$$AGSR (|F; m_4, HARA_4) \approx SR \left\{1 + \frac{1}{6} SR (1 + \eta) \xi - \frac{1}{24} SR^2 \left[(1 + \eta) (1 + 2\eta) \kappa + 3\eta (2\eta - 1)\right]\right\}. \tag{10}$$

Table III gives the GSRs for η = 0, 0.5, 1, and 2.

The approximations for CER* and GSR in (9) and (10) favor positive skewness and penalize excess kurtosis. This is in line with intuition and earlier arguments like those of Scott and Horvath (1980). These formulae also show that preferences for positive skewness and negative excess kurtosis increase with the sensitivity of risk tolerance to wealth.

However, these preferences for skewness and excess kurtosis fall short of indicating that, in our first example, asset B should always be preferred to A. The approximate values of CER* (|F; m_4, HARA_4) for assets A and B are shown in Table II and are represented graphically in Figure 3 for risk tolerance sensitivities to wealth ranging from η = 0 to η = 2. Asset B is shown with solid black lines and A with dashed lines. The CER*s calculated with the physical probability distributions are plotted in bold; the CER*s calculated with the first four moments and with approximation (9) are plotted with thinner lines.

We found that in the (N, E) case the CER*s for A and B are 4% and 2.76%, respectively, showing incorrectly that A would be superior to B. Instead, the CER*s with the physical distributions and HARA utilities show correctly that B is always preferred to A. But the exact four-moment-based CER*s (intermediate weight lines in Figure 3) are sufficiently different from those obtained with the physical distributions to show, wrongly, that A should be preferred to B. Finally, the approximate CER*s calculated with formula (9) tend to be further away from the genuine CER*s and give mixed indications: A is preferred to B when η < 1/3 and vice versa when η > 1/3.

In conclusion, for portfolio performance assessment, and indeed for portfolio optimization, it

\[26\) (9) was proposed by Pézier (2004) for the exponential case (η = 0) as a Sharpe ratio adjustment for skewness and excess kurtosis.\]
is always better to use an empirical distribution (ex post) or a full probabilistic forecast (ex ante) of returns and a utility function to optimize a $CER^*$ rather than to rely on a four-moment characterization of the return distribution. Moment-based criteria, especially their analytical approximations, should only be used when deviations from the $(N, E)$ case are small and the Sharpe ratio is smaller than 1. But even in these circumstances, there are no guarantees that moments-based criteria will not violate such basic axioms of choice as stochastic dominance.

VI Limitations of Heuristic Risk Adjusted Performance Measures (RAPMs)

Public reporting of investment performance rarely takes into account investor specific circumstances. What is reported in the absence of such information, and to satisfy a large body of heterogeneous investors in an objective way, are summary statistics of past returns and, sometimes, comparisons with benchmark returns and sensitivities to market risk factors, usually with the warning that past performance offers no guarantee of future returns. Analysts may also express their personal views, but, ultimately, investors must make their own forecasts.

Yet, as already noted, many investors would like to know what is best for them without having to state their individual views, investment circumstances, and risk attitude. This started a quest for objective and universal RAPMs that aspire to be more reliable guides for future investment decisions than ordinary Sharpe and information ratios. More than a hundred RAPMs have been proposed since the introduction of the Treynor (1965) and Sharpe (1966) ratios. Many are seen peppering the performance reports of hedge funds and other actively managed funds.

27A standard for the calculation of standard deviation of historical returns has been introduced in the latest release of Global investment performance standards (GIPS®), which came into effect on 1st January 2011. the GIPS® do not go further in terms of risk assessment.
that have been singled out - not always fairly - as yielding non-normally distributed returns (see for example Brooks and Kat (2002), Agarwal and Naik (2004), and Malkiel and Saha (2005)).

Cogneau and Hübner (2009) report 101 RAPMs. Meucci (2005) gives a pedagogical presentation of what he calls ‘satisfaction indices’ and analyzes their key properties. Sharma (2005) critically reviews several commonly used RAPMs in the context of hedge fund performance assessment (he also argues in favor of a particular CER-based measure, which we briefly discuss at the end of this section). Scholz and Wilkens (2005) show the links between a few of the more traditional RAPMs and also introduce their own. We complement these reviews by showing how the most popular RAPMs have been designed and analyze in Appendix D their relations with utility functions, if any. A simple taxonomy helps understand the motivations behind the main families of RAPMs. First, we can distinguish RAPMs that relate to the performance of a single risky asset from those which consider performance relative to a benchmark. Second, within each of these two families, we can distinguish a few RAPMs that are path-dependent from the majority that are path-independent (and therefore depend only on final returns).

Among the single asset path-dependent RAPMs we have the Calmar (Young (1991)) and Sterling (Bacon (2008)) ratios; they are similar to the Sharpe ratio except that the standard deviation at the denominator of the ratio is replaced by a maximum drawdown or an average of maximum drawdowns over successive periods. These criteria are affected by dynamic features of the return process such as heteroskedasticity or mean reversion. If the return process has independent, identically distributed (i.i.d.) increments, there is no reason to use path-dependent criteria as they can be mapped into equivalent criteria defined on the return distribution at the investment horizon. Path-dependent RAPMs on returns relative to a benchmark could be designed similarly.

But the majority of RAPMs are path-independent.\textsuperscript{28} The older ones use symmetrical measures of uncertainty, probably because modern portfolio theory assumed initially that asset returns were normally distributed. RAPMs on one branch use absolute asset returns whereas, following the development of CAPM and arbitrage pricing theory, RAPMs on another branch use returns relative to a market portfolio or to a number of risk factors. Along the first branch one finds the Sharpe ratio ($\mu/\sigma$), the Fama (1972) total risk alpha (TRA), the normalized TRA (or $M^2$) and the Modigliani and Modigliani (1997) RAP. Along the second branch, in parallel as it were, one finds the Treynor ratio ($\mu/\beta$), the Jensen $\alpha$, the normalized Jensen $\alpha$ (or $T^2$) and the Scholz and Wilkens (2005) MRAP.

With the rapidly growing role played by alternative investments since the early 1990s (structured products, hedge funds, private equity, etc.) greater attention has been given to the treatment of asymmetrical uncertainties. New RAPMs have been defined with downside-risk statistics; upside returns are considered separately, if at all. The choice of downside statistics can be viewed as a means of expressing a degree of risk aversion, an ersatz for a utility function. For example, excess expected return on value-at-risk at a significance level $\alpha$,\textsuperscript{29} or $RoVaR(\alpha)$, has become popular since the mid 1990s when the Basel Committee on Banking Supervision

\textsuperscript{28}When path-independent RAPMs are estimated from a series of historical returns, one assumes implicitly that these returns are i.i.d.

\textsuperscript{29}The value-at-risk at the significance level $\alpha$ (or confidence level $(1 - \alpha)$), denoted $VaR(\alpha)$, is usually defined as the negative of the $\alpha$-quantile of the empirical distribution of return at the investment horizon.
proposed to set standards for minimum capital requirements based on VaR measures. The logic is that potential losses should not exceed a firm’s capital at the end of a certain period with a probability greater than the frequency of default observed for good credit quality financial firms (typically, slightly less than 0.1% per year for single 'A' rated financial firms). RoVaR can therefore be regarded as a measure of expected excess return relative to some threshold, usually a risk-free rate or a cost of capital, per unit of capital requirement (Gregoriou and Gueyie (2003)). Between two distributions with same standard deviation, RoVaR favors the distribution with the larger expected return above threshold and penalizes the distribution with a fatter lower tail. VaR may be inferred from historical returns using a variety of methods including analytical estimates. For example, VaR can be approximated by using the Cornish-Fisher expansion for quantiles of a non-Normal distribution based on the variance, skewness, and excess kurtosis of that distribution.

More generally, downside risks can be characterized by lower partial moments of various order. The lower partial moment of order $n$ ($n > 0$) of a nominal return distribution given a threshold $\tau$, is defined as:

$$LPM(n, \tau) \equiv E[(\max(\tau - r, 0))^n],$$

where $r$ is a nominal return and $E[.]$ the expectation operator under personal probabilities or the corresponding ex post estimator.

The zero order lower partial moment is the probability of the return not exceeding the threshold. It is often used as a probabilistic constraint in portfolio optimization problems where this constraint is linked to a capital requirement. But it has also been proposed as an objective (to be minimized) either directly or indirectly as in Stutzer (2000). Stutzer remarks that, asymptotically, the probability for a normally distributed nominal return $r \sim N(\mu, \sigma^2)$ to fall below a low threshold is proportional to $\exp[-\frac{1}{2}(\mu/\sigma)^2]$. He also notes that this expression is the opposite of the maximum expected utility of an investment in the risky asset calculated with an exponential utility function. Because Stutzer believes that a reasonable objective for a fund manager is to minimize the probability of not achieving a minimum threshold (a 'failure' in his terms or 'disaster' in those of Roy (1952) and Fishburn (1977), he concludes that it must also be reasonable to maximize the expected utility of an investment in the risky asset calculated with an exponential utility function - an amusingly roundabout way of justifying the use of utility theory and, more specifically, of exponential utility functions. His proposed portfolio performance index is an increasing function of the resulting maximum utility, which can be written in our notations as:

$$I_p \equiv -\log(-EU^* (|F; p, E)) = CER^* (|F; p, E)/\lambda = \frac{1}{2} GSR^2 (|p, E).$$

In other words, the Stutzer index is proportional to $CER^* (|F; p, E)$ and equal to half of Hodges’ generalized Sharpe ratio squared.

The first order lower partial moment is the value of a put option struck at the threshold, calculated with a personal probability distribution. It is used in the definition of several criteria such as the Omega ratio (Keating and Shadwick (2002)):

$$\Omega(\tau) \equiv \frac{E[\max(r - \tau, 0)]}{E[\max(\tau - r, 0)]} = 1 + \frac{(\mu - \tau)}{LPM(1, \tau)},$$

In other words, the Stutzer index is proportional to $CER^* (|F; p, E)$ and equal to half of Hodges’ generalized Sharpe ratio squared.
where \( \mu = E[r] \) is the personal expected nominal return.

A related criterion is the expected shortfall (ES), also called tail conditional expectation (TCE), or conditional value at risk (CVaR) in the financial literature. It is the personal expected return conditional on the return being less than a certain \( \alpha \)-quantile (Kusuoka (2001) and Tasche (2002)). In terms of lower partial moment of order 1,

\[
ES(\alpha) \equiv r_\alpha - LPM(1, r_\alpha) / \alpha,
\]

where the threshold \( r_\alpha \) is the \( \alpha \)-quantile of the personal return distribution.

Interestingly, the expected shortfall satisfies the 'coherence' properties proposed by Artzner, Delbean, Eber, and Heath (1999) for risk metrics (after translation into similar properties for preference metrics) as well as two additional properties proposed by Acerbi (2002) making it a risk-averse spectral index.

A popular class of 'coherent' (but not spectral) indices, is of the form:

\[
Coh(n) \equiv \mu - \nu \cdot [LMP(n, \mu)]^{1/n}, \; \nu > 0,
\]

where \( \mu \) is again used for the personal expected return and \([LMP(n, \mu)]^{1/n}\) is the normalized, centered lower partial moment of order \( n \), for example, the lower semi-standard deviation when \( n = 2 \). As far back as 1959, Markowitz suggested using a downside semi-variance rather than a variance to determine optimal portfolios, but, to his chagrin, the mathematics proved intractable.

Sortino and van der Meer (1991) suggested using the lower semi-standard deviation instead of the standard deviation in the Sharpe ratio and thus introduced what is now referred to as the Sortino ratio. It has become part of the wider group of Kappa indices proposed by Kaplan and Knowles (2004) where the denominator is a centered lower partial moment of order \( n \). Ziemba (2005) prefers to use a lower partial moment of second order with respect to a nil threshold (not the expected return) in his definition of a Symmetric Downside-Risk Sharpe ratio.

Many of these RAPMs depend on the choice of at least one exogenous parameter such as a threshold value or a confidence level. The choice of this parameter is implicitly related to the risk attitude of the investor, but it is hardly a substitute for a full utility function. Indeed, we show in Appendix D that none of the families of preference criteria defined above is compatible with any sensible utility function, except, possibly, for some of the criteria based on lower partial moments (e.g., Sortino). Indeed many do not even reflect risk aversion. Some of those that do, like expected shortfall, display a type of risk aversion that is incompatible with the risk aversion produced by any concave utility function. This argument is developed in Appendix D as part of a wider discussion of similarities and differences between CERs, coherent, and spectral indices.

Rather than specifying all defining parameter(s) of a RAPM some authors suggest an extensive form of analysis whereby the RAPM is calculated over a range of values of the defining parameter(s). Then investors can decide whether, over the range of parameter values that is relevant to them, the RAPM gives a clear preference ranking among the investment opportunities they
want to compare. If not, it means that they should not have strong preferences among such investments. Either way, one should question the need for an extensive form of analysis when direct inspection of cumulative return distributions provides at least as much information. For example, there is a one-to-one relationship between an $\Omega(\tau)$ curve and a cumulative return distribution, so there is no more information in the first than in the second, and one might as well compare cumulative return distributions rather than $\Omega(\tau)$ curves.

The same problems would appear with even greater complexity if one were to consider downside risks relative to a benchmark portfolio or in a multi-risk factor context. One would have to judge how bad is a downside risk relative to the benchmark compared to a downside risk on the benchmark itself. So, there have been few attempts to develop downside-risk RAPMs on relative returns.

If several RAPMs, or a range of parameters values for one RAPM, appear to be sensible choices and none of them emerges as a clear favorite, it is always possible to try them all and to compare their rankings of risky assets. If they all lead to the same ranking then the choice is immaterial. If they produce different rankings, some authors have suggested combining them in a global performance measure. That is the approach commonly used in new product reviews when many attributes have to be taken into account (e.g., for a television: image quality, sound quality, connectivity, radiation, energy consumption, price, etc.). Each attribute is scored, weighted, and added to the others to obtain a global score. The same logic could be applied to compare funds using attributes like size, years of operations, management experience, risk management organization, jurisdiction, liquidity, frequency of reporting, expense ratio, past performance, etc. But the logic of this approach is dubious when applied to the assessment of a single characteristic, the return, and all that is known about it is either a series of historical returns or a probability forecast.

Nonetheless, modern forms of data envelopment analysis (DEA), introduced by Charnes, Cooper, and Rhodes (1978) to combine input and output measures in operational research problems, have been used to compare fund performance (Gregoriou, Rouah, Stephen, and Diz (2005) and Ammann and Moerth (2008)). A number of return characteristics are chosen and classified intuitively as either ‘good’ (e.g., expected return above threshold) or ‘bad’ (e.g., lower partial moment). Then a weighted sum of good measures, called efficiency measure, is maximized for a risky asset subject to two conditions: with the optimal set of weights, (i) the weighted sum of bad measures for this asset is equal to 1 and (ii) the weighted sum of good measures for all other assets is less or equal to the weighted sum of bad measures. As a result: the maximum efficiency score for an asset is capped at 1; several assets may have a score of 1; each score depends on optimal weights specific to the asset. So DEA does not discriminate among maximum efficiency assets and the rankings do not depend on a single set of weights. Improvements have been suggested to remove some of these weaknesses, for example, Gregoriou, Sédzro, and Zhu (2005) introduce an average of cross-efficiency scores (efficiencies of one asset when using optimal weights from others); Beadorf (2010) uses cross-efficiency scores to analyze the performance of private banking portfolios. She concludes that the multi-criteria DEA approach can produce more discriminatory scores than any single criterion, but the results are very sensitive to the choice of component criteria and the choice of weights, hence require careful interpretation. There is certainly no assurance that a DEA efficiency score would be compatible with $CER$ maximization.
In conclusion, most of the single-asset, end of investment period RAPMs that we have reviewed do not correspond to specific investment conditions. Weighted scores combining several RAPMs are even less likely to be related to specific circumstances and are even more difficult to interpret.

This problem is recognized by Sharma (2005) who, like us, argues in favor of using CERs to compare the performance of assets with non-Normal returns. Where Sharma’s analysis differs from ours is in the choice of the investment context. Sharma calculates a CER, which he calls AIRAP (Alternative Investment Risk Adjusted Performance), based on allocating the entire wealth of the investor to a risky asset with no alternative investment or leveraging and therefore no optimization. He also favors the use of a single parameter power utility function and suggests a standard value for this parameter. We think these conditions are too restrictive and unusual. Allocating one’s entire wealth to a single asset is generally neither optimal nor realistic. AIRAP varies when the risky asset is leveraged; it favors low-risk assets that can be taken in large quantities and penalizes high-risk assets that could be very attractive taken in small quantities. In contrast, a CER maximizing the investor’s expected utility function ought to be more realistic; its calculation is also more informative as it leads to the composition of the optimal portfolio for this investor.

VII  Ex post Performance Evaluation of Main Asset Classes: No need for Sophisticated RAPMs

Many investment information services report estimates of skewness and excess kurtosis in addition to means and standard deviations. These statistics depend critically on the duration of the return period. To illustrate, in Table IV we compare the performance of three traditional asset classes – Bond, Equity, and Real Estate (RE) – and five alternative asset classes – Volatility (Vol), Commodity (Com), Hedge Funds (HF), Funds of Hedge Funds (FoHF), and Private Equity (PE) – using moments, ordinary Sharpe ratios, and GSRs. The data are the 51 monthly returns from October 2002 to December 2006 examined by Pézier and White (2008) in a study of the relative merits of alternative investments. These statistics are excess returns adjusted for transaction costs and de-autocorrelated to eliminate possible reporting biases for the less liquid asset classes. Each asset class is represented by an index and, as such, monthly returns are closer to normality and have lower volatility than if they pertained to a single representative asset in each class. The GSRs are calculated with two HARA utility functions, an exponential, \( GSR(p, E) \), and a logarithmic, \( GSR(p, L) \).

Monthly returns of all asset classes, except Equity, show negative skewness and all, except Hedge Funds, show excess kurtosis. Returns from RE, (Vol) – taken short – and, to a lesser extent, FoHF depart markedly from normality. This is confirmed by the Jarque-Berra statistic. With 51 observations (small sample), this statistic has a critical value of about 6.7 at the 95% confidence level; thus monthly returns for RE, (Vol) and FoHF have less than 5% probability of arising from a normal distribution, whereas monthly returns from other asset classes do not depart markedly from normality. The ordinary Sharpe ratios and the GSRs are close for all asset classes except for RE and (Vol) for which the effects of negative skewness and positive excess kurtosis result in markedly lower GSRs, especially with a logarithmic utility. The four-moment approximations of GSRs (Table III) are given immediately below the corresponding GSRs. They are close to the exact GSRs except for RE and (Vol), the two asset classes with returns
<table>
<thead>
<tr>
<th></th>
<th>Bond</th>
<th>Equity</th>
<th>RE</th>
<th>(Vol)</th>
<th>Com</th>
<th>FoHF</th>
<th>HF</th>
<th>PE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>0.13%</td>
<td>1.08%</td>
<td>1.38%</td>
<td>8.74%</td>
<td>0.49%</td>
<td>0.40%</td>
<td>0.72%</td>
<td>1.85%</td>
</tr>
<tr>
<td><strong>Std-deviation</strong></td>
<td>0.90%</td>
<td>3.23%</td>
<td>3.30%</td>
<td>10.89%</td>
<td>2.25%</td>
<td>1.37%</td>
<td>1.80%</td>
<td>5.13%</td>
</tr>
<tr>
<td><strong>Skewness</strong></td>
<td>-0.41</td>
<td>0.25</td>
<td>-1.35</td>
<td>-0.63</td>
<td>-0.52</td>
<td>-0.85</td>
<td>-0.29</td>
<td>-0.22</td>
</tr>
<tr>
<td><strong>Excess kurtosis</strong></td>
<td>0.32</td>
<td>1.09</td>
<td>2.60</td>
<td>2.64</td>
<td>0.22</td>
<td>1.20</td>
<td>-0.41</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>JB (6.7 @ 5%)</strong></td>
<td>1.64</td>
<td>3.05</td>
<td>29.92</td>
<td>18.18</td>
<td>2.38</td>
<td>9.16</td>
<td>1.06</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>Sharpe ratio</strong></td>
<td>0.15</td>
<td>0.33</td>
<td>0.42</td>
<td>0.80</td>
<td>0.22</td>
<td>0.29</td>
<td>0.40</td>
<td>0.36</td>
</tr>
<tr>
<td><strong>GSR(p, E)</strong></td>
<td>0.15</td>
<td>0.34</td>
<td>0.39</td>
<td>0.73</td>
<td>0.22</td>
<td>0.28</td>
<td>0.40</td>
<td>0.36</td>
</tr>
<tr>
<td><strong>AGSR(m₄, E)</strong></td>
<td>0.15</td>
<td>0.34</td>
<td>0.37</td>
<td>0.68</td>
<td>0.22</td>
<td>0.28</td>
<td>0.40</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>GSR(p, L)</strong></td>
<td>0.15</td>
<td>0.33</td>
<td>0.35</td>
<td>0.62</td>
<td>0.21</td>
<td>0.27</td>
<td>0.39</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>AGSR(m₄, L)</strong></td>
<td>0.14</td>
<td>0.33</td>
<td>0.28</td>
<td>0.26</td>
<td>0.21</td>
<td>0.26</td>
<td>0.39</td>
<td>0.34</td>
</tr>
</tbody>
</table>

|                        | Mean | 1.6%   | 13.7%| 17.8% | 173.6%| 6.1% | 4.9% | 9.0% | 24.3%|
| **Std-Deviation**      | 3.1% | 12.3%  | 13.4%| 95.9% | 8.3%  | 5.0% | 6.6% | 21.6%|
| **Skewness**           | 0.06 | 0.36   | 0.02 | 0.75  | 0.15  | -0.10| 0.13 | 0.45 |
| **Excess kurtosis**    | 0.06 | 0.08   | -0.18| 1.02  | -0.16 | 0.10 | 0.09 | 0.46 |
| **JB (5.99 @ 5%)**     | 0.82 | 21.84  | 1.41 | 137.00| 4.65  | 2.07 | 3.21 | 42.94|
| **Sharpe ratio**       | 0.51 | 1.11   | 1.33 | 1.67  | 0.74  | 0.99 | 1.36 | 1.13 |
| **GSR(p, E)**          | 0.52 | 1.19   | 1.35 | 2.19  | 0.75  | 0.97 | 1.39 | 1.23 |
| **AGSR(m₄, E)**        | 0.52 | 1.18   | 1.35 | 1.82  | 0.75  | 0.97 | 1.39 | 1.20 |
| **GSR(p, L)**          | 0.48 | 0.92   | 1.11 | 2.17  | 0.67  | 0.74 | 1.00 | 1.03 |
| **AGSR(m₄, L)**        | 0.50 | 1.06   | 1.15 | 0.59  | 0.73  | 0.81 | 1.07 | 0.98 |
departing most from normality. For all other asset classes, the four-moment approximations can still be regarded as useful improvements on the ordinary Sharpe ratios.

But investors are generally interested in longer than monthly investment periods, especially when investing in illiquid assets. The inference of annual return distributions from monthly returns requires a probabilistic model. The monthly returns are already de-autocorrelated; if we assume further that they are i.i.d., we can bootstrap annual returns by compounding 12 monthly returns chosen at random in the 51 monthly return data set. Statistics derived from a set of 1,000 such simulations of annual returns are reported in the bottom half of Table IV. These annual statistics bear little resemblance to the monthly statistics. Two main factors affect the annual returns: the sum of i.i.d returns would tend toward a normal distribution but compounding generates positive skewness and excess kurtosis associated with Log-normal distributions, especially when monthly return volatilities are high. As a result, all annual skews are positive except for FoHF, which has the most negative monthly skew, and excess kurtoses are small and mainly positive. At the 95% confidence level, returns for three asset classes depart significantly from normality: (Vol) and PE, with both higher than normal skewness and excess kurtosis, and Equity, with a higher than normal skewness.

Compared with ordinary Sharpe ratios, the $GSR$s reflect the balance between the preference for positive skewness and the aversion against excess kurtosis, the latter playing a greater role with logarithmic than exponential utilities. On balance, because of positive skewness, exponential $GSR$s are slightly larger than ordinary Sharpe ratios for all asset classes except FoHF. Because of positive excess kurtosis, logarithmic $GSR$s are smaller than ordinary Sharpe ratios for all asset classes except (Vol). Four-moment approximations of $GSR$s depart from the ordinary Sharpe ratios in the right direction in all cases except for (Vol) with a logarithmic utility. They are closer to corresponding $GSR$s than Sharpe ratios except when there is an extreme departure from normality of returns as is the case for (Vol) and, to a lesser extent, PE and Equity. It seems safe to say that a four-moment approximation is more relevant than the ordinary Sharpe ratio when the difference between the two is less than 20% as would be the case when departures from normality are small, the coefficient of sensitivity of risk tolerance to wealth is not greater than 1 and the Sharpe ratio itself is not greater than 1.

VIII Ex post Performance Evaluation of Famous Funds Using $GSR$s and $GIR$s with HARA utilities

We now compare the performance of six well-known funds, the total return S&P500 index, and the long US Treasury Bond using ordinary and generalized Sharpe ratios and information ratios. The data are the yearly return figures from Dec 1986 to Dec 1999 studied by Ziemba (2005).

Table V and Figure 4 display the ordinary Sharpe ratios and the generalized Sharpe ratios calculated with exponential and logarithmic utilities. The bottom line of V shows the geometric average excess annual return of each fund. Ignoring risks, Berkshire-Hathaway clearly comes first with an annual average of 19.37%, followed by Quantum (16.54%). They are the only two funds outperforming the S&P500 index (12.62%) over the period, although the Tiger Fund (12.11%) comes close behind. When taking risks into account, different rankings emerge. Har-
Table V: Performance comparison of famous funds using GSRs and CER*s

<table>
<thead>
<tr>
<th>Fund</th>
<th>Sharpe ratio</th>
<th>$GSR(p, E)$</th>
<th>$GSR(p, L)$</th>
<th>$CER^*(N, E)$</th>
<th>$CER^*(p, E)$</th>
<th>$CER^*(p, L)$</th>
<th>Avg. Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tiger</td>
<td>1.12</td>
<td>1.40</td>
<td>2.17</td>
<td>9.96%</td>
<td>15.65%</td>
<td>37.54%</td>
<td>12.11%</td>
</tr>
<tr>
<td>Quantum</td>
<td>1.11</td>
<td>1.33</td>
<td>1.67</td>
<td>9.81%</td>
<td>14.05%</td>
<td>22.21%</td>
<td>16.54%</td>
</tr>
<tr>
<td>Ford</td>
<td>1.01</td>
<td>1.04</td>
<td>1.18</td>
<td>8.18%</td>
<td>8.73%</td>
<td>11.20%</td>
<td>9.01%</td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>1.05</td>
<td>1.08</td>
<td>1.15</td>
<td>8.80%</td>
<td>9.32%</td>
<td>10.51%</td>
<td>12.62%</td>
</tr>
<tr>
<td>Harvard</td>
<td>1.14</td>
<td>1.12</td>
<td>1.15</td>
<td>10.42%</td>
<td>10.11%</td>
<td>10.50%</td>
<td>9.73%</td>
</tr>
<tr>
<td>Berk.-H.</td>
<td>0.77</td>
<td>0.79</td>
<td>0.81</td>
<td>4.70%</td>
<td>4.96%</td>
<td>5.25%</td>
<td>8.33%</td>
</tr>
<tr>
<td>Windsor</td>
<td>0.67</td>
<td>0.67</td>
<td>0.61</td>
<td>3.56%</td>
<td>3.28%</td>
<td>2.93%</td>
<td>2.39%</td>
</tr>
<tr>
<td>US-Tr.</td>
<td>1.07</td>
<td>1.10</td>
<td>0.41</td>
<td>1.37%</td>
<td>1.42%</td>
<td>1.40%</td>
<td>2.39%</td>
</tr>
</tbody>
</table>

vard, Tiger, Quantum, S&P500, and Ford have ordinary Sharpe ratios greater than one, whilst Berkshire-Hathaway and Windsor trail with 0.77 and 0.67, respectively; only US-Treasuries with 0.41 have a worse Sharpe ratio. These results would be meaningful if returns were normally distributed and investors had constant absolute risk aversion. The optimal amounts they would allocate to these funds if the only alternative was the risk-free asset (first line in Table V labeled $\omega^*(N, E)$) range from 207% of wealth for Harvard down to 41% of wealth for Berkshire-Hathaway.

But these funds’ returns are not normally distributed. Assuming they will continue to similarly to their past returns, the Tiger and Quantum funds are the most attractive, especially when the GSRs are calculated with a logarithmic utility ($GSR(p, L)$) rather than an exponential utility ($GSR(p, E)$). With a logarithmic utility, investors would find it optimal to invest almost six times their wealth in the Tiger fund because of its limited downside risk, whereas they would invest only 39% of their wealth in Berkshire-Hathaway.

It is also informative to consider the $CER^*$s in the lower part of Table V and to compare them with the average returns on the bottom line. The $CER^*$s are affected by two factors (i) a risk discount and (ii) the optimal leveraging of the investments. For example, with the Tiger fund, the $CER^*$ corresponding to the ordinary Sharpe ratio ($CER^*(N, E)$) is 9.96% only, compared to an average return of 12.11%, because of risk and despite an optimal leveraging of 58% (158% invested in Tiger implies a borrowing of 58%). But based on the historical return distribution of this fund, the low downside risk allows for much higher optimal leveraging and therefore much higher exponential and logarithmic $CER^*$s.

We now consider the more realistic situation where potential investors in hedge funds would already be equity investors having access to an S&P500 total return fund. These investors seek the maximum increase in $CER$ from an optimal combination of the risk-free asset, the S&P500, and an alternative fund. The attractiveness of the alternative fund is measured in terms of potential increase in certainty equivalent return, $\Delta CER^*$ or, equivalently, in terms of GIR.
Table VI and Figure 5 display the composition of the optimal allocations $\omega^*$ for each alternative fund and $\omega_{S&P}^*$ for the benchmark S&P500 fund, the GSRs, and the $\Delta CER^*$s that each fund contributes. These results are calculated under the three conditions: $(N, E)$, $(p, E)$, and $(p, L)$ and with $\lambda = 0.16$ for the calculation of $\Delta CER^*$s. For example, the last column in Table VI indicates that a logarithmic utility investor would be optimally invested 71% in US Treasuries and 133% in the S&P500 total return fund, a total investment of 204% of wealth that requires 104% funding. This allocation does not reflect the market equilibrium with nil net funding and an international bond market about twice the size of the international equity market; it highlights the exceptional 12.62% geometric average yearly excess return of the S&P500 during the December 87 to December 99 period under review. In contrast, Dimson, Marsh, and Staunton (2006) report a lower equity premium average of 4.11% for 17 countries over the period 1900-2005, and certainly the last 10 years have produced no significant equity premia in developed countries. The risk tolerance coefficient of 0.16 we have used for our illustrations is compatible with a forecast excess return for equities of 3% and current market capitalizations.

The differences in performance among funds are striking. Tiger, which is positively correlated with S&P500, is so attractive that our investor would rather short the S&P500 to leverage a long position in Tiger. But Quantum, which is less correlated with S&P500 and has a high alpha, easily wins the contest. It should be combined with the S&P500 in a highly leveraged portfolio. In the middle range, Ford, Harvard, and Berkshire-Hathaway would make only minor contributions to an S&P500 investor. Finally the Windsor performance is so poor that, it would be advantageous to short it (for example, with a total return swap) to leverage an investment in the S&P500.

It is also worth noting that the relative attractiveness of funds changes when replacing an

---

**Figure 4: Performance comparison of famous using GSRs and CER*s**

![Performance comparison of famous using GSRs and CER*s](image-url)
Incorrect Normal distribution assumption with the empirical return distributions and when changing risk attitude from constant absolute to constant relative risk aversion. For example, the attractiveness of Harvard decreases under these changes whereas the attractiveness of Tiger and Quantum greatly increases.

This highlights the critical issue of choosing representative conditions for public reporting of performance measures. One has to specify: (i) the return distributions of the funds at some investment horizon, (ii) investment alternatives, and (iii) the risk attitude of investors.

Depending on the fund, returns may be recorded monthly, weekly, daily or at even higher frequencies over a historical period of months or years. For public reporting, these data must be transformed as objectively as possible into a return distribution at some investment horizon. The standard horizon is 1 year for traditional asset classes, though longer horizons in line with typical holding periods could be more relevant for illiquid assets. There are generally too few yearly return observations to construct a yearly return distribution without making some assumptions; old data, when available, might not be relevant anyway.\(^\text{30}\) The simplest and most common assumption is that short-term returns are i.i.d. and longer-term returns should be calculated with simple compounding. This is a weaker assumption than a choice of specific price dynamics (e.g., geometric Brownian) and allows for more realistic evaluations of non-normal characteristics (as we have seen in Section VII).

The investment context that justifies the use of GSRs - that is, an optimally leveraged position in one risky asset only - has the advantage of simplicity but seems relevant only for the comparison of global investment portfolios. Otherwise, one or a few benchmark portfolios of general interest to investors should be specified, for example: a worldwide 'market’ benchmark including bonds

\(^\text{30}\)For example, three years of monthly historical data is recommended by the Chartered Financial Accountants Institute in their GIPS®.
and equities, and a few indices for major asset classes such as an investment grade bond index, the MSCI World Equity Index, the S&P500 index, . . . ). Performance of single assets or funds would be expressed as GIRs relative to these benchmarks.

Risk attitudes, the current LARTs in particular, vary enormously among investors. As we have seen, CER* s are strongly affected by $L(0)$; with HARA utilities, they are proportional to $L(0)$. On the other hand, with HARA utilities, GSRs and GIRs are independent of $L(0)$ and only affected (to a lesser degree) by the sensitivity of LARTs to changes in wealth. These criteria are therefore more suitable for public reporting. Since Bernoulli introduced the concept of utility nearly 300 years ago, it has been argued that absolute risk tolerance should increase with wealth. Bernoulli suggests that a thousand dollar risk is to a millionaire what a million dollar risk is to a billionaire; he therefore advocates the use of a logarithmic utility function with constant relative risk tolerance, but this condition may be too restrictive. With the two-parameter HARA utility function (6), the LART is $L(r) = (\lambda + \eta r)$ and $\eta$ can be chosen to fit representative risk attitudes. The analysis of optimal option profiles in the next section suggests that, on average across investors, $\eta = \sigma^2/\mu$ with $\mu$ and $\sigma^2$ representing the expected excess return and the variance of the ‘market’ portfolio. This places the average $\eta$ in a range of approximately 0.15 to 0.50.

For a survey of the literature on logarithmic utilities see Hakansson and Ziemba (1995).
IX Optimal Return Profiles with HARA Utilities

Markowitz’s mean-variance portfolio optimization and Sharpe-Mossin-Lintner’s subsequent equilibrium asset pricing model assume the Normal return distribution, exponential utility \( (N, E) \) conditions for a static investment over a finite period. Different conditions would lead to different results. Simply transposing to continuous time and with constant views, Markowitz’s solution implies a constant mix strategy (constant percentage value in each asset), i.e., a contrarian dynamic strategy with a total return that is no longer a linear combination of the component asset returns. The Sharpe ratio is a rational preference criterion for single risky assets in the \( (N, E) \) case and should be interpreted with caution in other situations. Some authors have been troubled by the realization that the Sharpe ratio of a fairly priced long call option on a log-normally distributed asset price is lower than the Sharpe ratio of the underlying asset, whereas the Sharpe ratio of a short put option is higher. They argue that there are no management skills involved in choosing to go long a call or short a put rather than to invest in the underlying asset and therefore that the Sharpe ratios should be the same for these alternative investments. Leland (1999) proposes to take into account higher moments of the return distribution to adjust the betas for options within the equilibrium asset pricing model, so as to equate the Sharpe ratios of options to those of their underlying assets.

Leland is correct in pointing out that the ordinary Sharpe ratio is not suited for the comparison of risky assets with different shapes of return distributions. But it does not follow that a suitable criterion should always indicate indifference between investing in an asset or in an option (i.e., a non-linear payoff) on this asset, simply because it does not take management skills to do one or the other. An investor might prefer one or the other for at least two reasons. First, the return forecast of the investor, her volatility forecast in particular, may differ from the forecast implied by the market price of the derivative. Second, assuming homogeneity of views, it is still possible that an investor would prefer a payoff that is non-linear in the asset price because of her personal risk attitude. Such preferences are reflected by differences in \( CER^* \)'s between investment in derivatives and in their underlying assets.

Following Brennan and Solanki (1981) and Constantinides (1982), Pézier (2007) analyzes the general problem of maximizing the expected utility of the PV of future wealth, \( w (r, T) \), for an investor with utility function \( u(w) \) making a probability forecast \( p(r) \) for the vector \( r \) of asset excess returns at the investment horizon \( T \), when the market implied risk-neutral distribution – the distribution pricing any asset, including derivatives, as the expected PV of its final payoff discounted at the risk-free rate – is \( q(r) \). Maximizing the personal expected utility of future wealth, \( E_P [u(w(r, T))] \), subject to the risk-neutral expected PV of future wealth, \( E_Q [w] \), being equal to initial wealth, yields the condition:

\[
u' (w(r, T)) \propto q(r) / p(r).
\]

Analytical solutions to (11) can be found for a variety of combinations of utility functions and return distributions.

The solution is particularly simple in the \( (N, E) \) case. The risky asset price dynamics are:

\[
S(T) = S(0) (1 + rT) \circ (u + rT),
\]
with risk-free rate \( r_f \), a unit vector \( \mathbf{u} \), and multivariate Normal distributions of excess returns over a unit period\(^{32} \) \( \mathbf{r} \sim N(\mu, \Sigma) \) under personal probabilities and \( \mathbf{r} \sim N(\mathbf{0}, \Sigma) \) under implied risk-neutral probabilities.\(^{33} \) When these return distributions are combined with an exponential utility function \( u(w) = -\exp(-w/\lambda) \), condition (11) becomes:

\[
\exp(-w(\mathbf{r}, T)/\lambda) \propto \exp(-\mu^T\Sigma^{-1}\mathbf{r}).
\]

Setting \( E_Q[w(\mathbf{r}, T)] = 1 \) yields the PV of final wealth

\[
\frac{w(\mathbf{r}, T)}{T} = 1 + \lambda \mu^T\Sigma^{-1}\mathbf{r}.
\]  \( \text{(12)} \)

This PV is generated by the static allocation to the risky assets:

\[
\omega = \lambda \Sigma^{-1}\mu,
\]  \( \text{(13)} \)

which is the unconstrained Markowitz’s optimal portfolio allocation, as one would expect in the \((N, E)\) case. The corresponding portfolio excess expected return, variance, Sharpe ratio, and \( CER^*\) per unit of time are,\(^{34} \) respectively:

\[
\mu = \lambda \mu^T\Sigma^{-1}\mu,
\]

\[
\sigma^2 = \lambda^2 \mu^T\Sigma^{-1}\mu = \lambda \mu,
\]

\[
SR = \sqrt{\mu^T\Sigma^{-1}\mu} = \sqrt{\lambda/\lambda},
\]

\[
CER^* = \frac{1}{2} \lambda \mu^T\Sigma^{-1}\mu = \frac{1}{2} \mu.
\]

The generalized Sharpe ratio of any portfolio non-linear in \( \mathbf{r} \) should therefore be no greater than that of the Markowitz’s portfolio. Figure 6 confirms this by comparing the generalized Sharpe ratios \( GSR(\text{long call}, F; N, E) \) of long call and short put options on an asset with expected return \( \mu = 10\% \) and volatility \( \sigma = 20\% \) for a range of strikes from 3 standard deviations in-the-money to 3 standard deviations out-of-the-money.

With the chosen return distribution, the Sharpe ratio of the underlying asset is \( \mu/\sigma = 0.5 \). A long call position at least two standard deviations in-the-money has a \( GSR \) almost equal to 0.5. However, the \( GSR \) of an at-the-money long call is only 0.433 and decreases further out-of-the-money. The results are similar for a short put option. More than two standard deviations in-the-money, its \( GSR \) is almost equal to that of the underlying asset, but is only 0.421 at-the-money and decreases further out-of-the-money. With a coefficient of risk tolerance \( \lambda = 0.16 \), the optimal size of an investment in the underlying asset is \( \omega^* = \lambda \mu/\sigma^2 = 40\% \) of wealth. The optimal underlying notional size of an investment in an option on this asset varies

\(^{32}\)Normally distributed returns are considered over a single period of time since normality is not preserved when compounding returns over multiple periods.

\(^{33}\)The column vector \( \mu \) represents the expected excess returns of the assets in the portfolio and \( \Sigma \) the variance-covariance matrix of returns, per unit time period; the inverse matrix is denoted \( \Sigma^{-1} \) and \( \mu^T \) is the transpose of vector \( \mu \). In subsequent equations, we shall also use a vector of volatilities, \( \sigma \) and a vector of correlated Wiener processes \( \mathbf{W} \) with correlation matrix \( \mathbf{R} \), leading to the variance-covariance matrix \( \Sigma = \text{diag}(\sigma) \mathbf{R} \text{diag}(\sigma) \). The symbol \( \circ \) denotes the Hadamard (term by term) product of two vectors or matrices.

\(^{34}\)In the \((N, E)\) case, the \( CER^* \) is defined like other rates so that, over a unit period, \( w = 1 + CER^* \). See Appendix F for the calculation.
from a minimum of 40\% of wealth for deep in-the-money options to several time this size for deep out-of-the-money options.

With the exception of the \((N, E)\) case, the optimal wealth function determined by condition (11) is generally not linear in the underlying assets returns and therefore can only be realized with a dynamic investment strategy or the purchase of suitable options. We derive analytical solutions for combinations of HARA utilities and asset prices with either Normal or Log-normal return distributions. The first order derivative with respect to wealth of a HARA utility function of the form (6) is

\[ u'(w(r, T)) \propto \left( 1 + \left( \frac{\eta}{\lambda} \right) (w(r) - w(0)) \right)^{-1/\eta}. \]

Log-normal risky asset price dynamics are

\[ S(T) = S(0) \exp \left( \left( r + u_r T - \frac{1}{2} \sigma^2 T \right) \right), \]

with continuous logarithmic excess returns \( r \sim N(\mu, \Sigma) \) under personal probabilities and \( r \sim N(0, \Sigma) \) under implied risk-neutral probabilities. So, in the Log-normal, HARA utility case, condition (11) becomes:

\[ \left( 1 + \left( \frac{\eta}{\lambda} \right) (w(r, T) - w(0)) \right)^{-1/\eta} \propto \exp \left( -\mu^T \Sigma^{-1} r \right). \]

It follows that, setting \( w(0) = 1 \) and letting \( m = \eta \Sigma^{-1} \mu \), the PV of wealth at time horizon \( T \) is:

\[ w(r, T) = 1 + \frac{\lambda}{\eta} \left( k \exp (m^T r T) - 1 \right), \]
with constant $k$ so that $E_Q[w] = 1$. Therefore

$$k = \exp \left(-\frac{1}{2}m^T \Sigma m \right)$$

and

$$w(r, T) = \left(1 - \frac{\lambda}{\eta} \right) + \left( \frac{\lambda}{\eta} \right) \exp \left( m^T r T - \frac{1}{2} m^T \Sigma m T \right). \tag{14}$$

In this case, the PV of wealth is the sum of a minimum value, or floor, equal to $(1 - \lambda/\eta)$ and an excess value above the floor, or buffer, initially equal to $(\lambda/\eta)$ and varying over time proportionally to the product of powers of the asset prices.

Wealth function (14) can, in theory, be generated in a Black–Scholes economy by maintaining exposures to the risky assets at a constant multiple $m$ of the buffer size. That is a constant proportionality portfolio insurance strategy ($CPPI$) in continuous time. The dynamic allocation of risky assets is:

$$\omega(r, T) = \left( \lambda \Sigma^{-1} \mu \right) \exp \left( m^T r T - \frac{1}{2} m^T \Sigma m T \right). \tag{15}$$

Initially, it equals the Markowitz’s optimal portfolio allocation. The corresponding $CER^*$ is calculated in Appendix F and is shown in Table VII.

The curvature of the wealth function (14) with respect to the $i^{th}$ asset price is:

$$\frac{w''}{w} = \frac{(m_i - 1)}{S_i(T)}. \tag{16}$$

When $m_i > 1$, the payoff is convex and increasing with $S_i(T)$ (like a long call position). When $m_i = 1$, the payoff increases linearly with $S_i(T)$, and when $m_i < 1$, the payoff is increasing in $S_i(T)$ but concave (like a short put position). Except for the linear case, the absolute value of the curvature decreases when $S_i(T)$ increases. We infer that, for all $m_i \neq 1$, there should be more interest from investors in holding standard options at low strikes than at high strikes.

On average, $m$ cannot be too different from 1 because the net payoff curvature, or net volatility position, per risky asset is nil in the market, so investors with relatively lower $\eta$ than average should hold assets with larger $\mu/\sigma^2$ – typically, bonds – whereas investors with higher $\eta$ than average should hold portfolios with lower $\mu/\sigma^2$ – typically, equities. On average over time, for balanced equity/bond portfolios, historical values of $\mu/\sigma^2$ range from about 2 to 7. Thus, with $\eta\mu/\sigma^2$ close to 1 on average, $\eta$ should be in the range 0.15 to 0.50 on average across investors.

At the constant absolute risk tolerance limit ($\eta \downarrow 0$), that is, with an exponential utility function, the optimal wealth function (14) reduces to (12), but it is now a logarithmic function of the risky asset prices, thus increasing with $S(T)$ and concave:

$$w(S(T)) = 1 + \lambda \mu^T \Sigma^{-1} \left[ \log \left( \frac{S(T)}{S(0)} \right) - (u^T r - \frac{1}{2} \sigma \circ \sigma) T \right]. \tag{17}$$

This concavity contradicts the zero net curvature, or zero net volatility position, for all risky assets in the market. It would imply a net short volatility position or an excess of contrarian
TABLE VII: Optimal return profiles and their CER*s with HARA utilities, Normal and Log-normal return distributions

\( w = \text{PV of future wealth} \)

\( \omega = \text{vector of risky asset allocations} \)

\( CER^* = \text{maximum certain equivalent excess return} \)

<table>
<thead>
<tr>
<th>Utility Function</th>
<th>Managed Return Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential (( \eta = 0 ))</td>
<td>Static, linear profile with normally distributed price returns (single period)</td>
</tr>
<tr>
<td></td>
<td>Dynamic, logarithmic profile with log-normally distributed price returns (unrealistic because always concave)</td>
</tr>
<tr>
<td></td>
<td>( w (r, T) = 1 + \lambda \mu^T \Sigma^{-1} r )</td>
</tr>
<tr>
<td></td>
<td>( \omega = \lambda \Sigma^{-1} \mu )</td>
</tr>
<tr>
<td></td>
<td>( CER^* = \frac{1}{2} \lambda \mu^T \Sigma^{-1} \mu )</td>
</tr>
<tr>
<td>HARA (( \eta &gt; 0 ))</td>
<td>Dynamic, exponential profile with normally distributed price returns (unrealistic because always convex)</td>
</tr>
<tr>
<td></td>
<td>Dynamic, power profile with log-normally distributed price returns</td>
</tr>
<tr>
<td></td>
<td>( w (r, T) = \left(1 - \frac{1}{\eta}\right) + \frac{1}{\eta} \exp \left( m^T r T - \frac{1}{2} m^T \Sigma m T \right) )</td>
</tr>
<tr>
<td></td>
<td>( \omega (r, T) = \left(\lambda \Sigma^{-1} \mu\right) \exp \left( m^T r T - \frac{1}{2} m^T \Sigma m T \right) )</td>
</tr>
<tr>
<td></td>
<td>( CER^*.T = \log \left(1 + \frac{1}{2} \left( \exp \left( m^T \mu T - \frac{1}{27} m^T \Sigma m T \right) - 1 \right) \right) )</td>
</tr>
</tbody>
</table>

over trend following trading strategies. Therefore, the combination of negative exponential utilities and log-normally distributed asset returns is unrealistic.

Likewise, the combination of a HARA utility function with \( \eta > 0 \) and therefore \( m > 0 \) with multivariate normal excess return distributions \( r \sim N (\mu, \Sigma) \) under personal probabilities and \( r \sim N (0, \Sigma) \) under implied risk-neutral probabilities is not realistic because it would lead to a convex wealth function. Indeed, relation (14) would still hold, but it would now be an exponential function of the risky asset prices, thus increasing with \( S(T) \) but always convex:

\[
w (S (T)) = \left(1 - \frac{\lambda}{\eta}\right) + \left(\frac{\lambda}{\eta}\right) \exp \left[ m^T \left( S (T) - S (0) - (\mu + \omega^T r) T - \frac{1}{2} m^T \Sigma m T \right) \right], \quad (18)
\]

with constant curvature \( m > 0 \) in each asset. At the exponential utility function limit (\( \eta \downarrow 0 \)), it reduces to the linear payoff (12). Table VII summarizes the results.

When a benchmark portfolio is available and one applies a CAPM or other linear portfolio model, the attractiveness of the optimal return profiles could be evaluated in terms of maximum gains in \( CER \) or, equivalently, in terms of GIRs, as we did in Section IV.

X Skewness of Credit Risks and the Credit Spread Puzzle

For a risk-neutral investor, or if a credit risk is efficiently diversified, a bond credit spread – the difference between the yield to maturity of a defaultable bond and the yield to maturity of an identical or closely related bond issued by a high quality issuer (AAA rated) – should
compensate for the expected default loss, that is, should be approximately equal to the product of the default probability times the loss given default assigned to that bond. But numerous studies have shown that credit spread implied default probabilities are often several times larger than empirical estimates of default probabilities (or historical default frequencies). This is so for bonds rated from AA to CCC. Even after carefully accounting for sources of differences between defaultable bonds and near default-free bonds – liquidity, tax effects, etc. – credit spreads still imply a large risk premium. This has been called the credit spread puzzle (Elton, Gruber, Agrawal, and Mann (2002)). We address this puzzle by examining the uncertainty of return in a large and well-diversified portfolio of US corporate bonds and calculating the excess expected profit necessary to make this portfolio reasonably attractive to investors.

There are two sources of return uncertainty in a portfolio of credit sensitive instruments: changing credit spreads and defaults. In the absence of default, the value of a portfolio of corporate bonds or other credit sensitive instruments (e.g., credit default swaps, collateralized debt obligations, ...) changes with changes in the appreciation of credit risks. This uncertainty is skewed. Credit spreads are positive and mean reverting. In times of credit crises (early 1970s, early 1990s, early 2000s, and 2008), default frequencies are more than 5 times their long-term historical average; credit spreads tend to vary even more, reflecting a positive correlation between average loss given default and frequency of default as well as possible over-reaction of investors.

The second source of uncertainty is in default losses over the investment horizon given default probabilities and losses given default. Default probabilities for all US corporate bonds have been about 1.5% per year on average over the last 40 years. If default events for these bonds were independent and had all the same probability \( p \), the probability of observing \( n \) defaults in a portfolio of \( N \) bonds would be given by the binomial distribution with expectation \( pN \) and variance \( p(1-p)N \). For a very large portfolio, the fraction of bonds defaulting would converge towards \( p \) as its variance, equal to \( p(1-p)/N \) would converge towards zero. Likewise, there would be convergence towards an average probability of default if default events have different probabilities but are independent. However this does not match observations. Default frequencies of US corporate bonds vary considerably from year to year. The two main causes of variation are likely to be changes in default probabilities and correlations between default events.

These two sources of uncertainty are dependent as the credit spreads of bonds usually worsen before default. One popular approach to model dependencies among changes in credit spreads and default is to associate credit spreads with credit ratings and map the discrete credit ratings and default states to continuous variables that could be interpreted as proxies for the net asset values of issuer firms. Credit rating transitions and the default state occur when the corresponding proxy variable crosses predefined barriers set at levels corresponding to the probabilities of reaching these various states. These so-called firm-value models were first introduced by Vasicek (1987). In the simplest models, a covariance matrix of a multivariate Normal distribution is used to characterize the joint distribution of the proxy variables at some future date. This has been implemented in well-known commercial credit portfolio models such as CreditMetrics. Other models may use other distributional assumptions, copulas, etc. This forms also the conceptual basis for the estimation of minimum credit risk capital requirements in the Basel II & III capital adequacy regulations.
Table VIII: Transition probabilities and excess returns for a corporate 'BB' rated bond

<table>
<thead>
<tr>
<th>Credit Rating</th>
<th>Transition Probability (%)</th>
<th>Excess return (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A and above</td>
<td>0.7</td>
<td>+7.0</td>
</tr>
<tr>
<td>BBB</td>
<td>8</td>
<td>+5.0</td>
</tr>
<tr>
<td>BB</td>
<td>81</td>
<td>+0.72</td>
</tr>
<tr>
<td>B</td>
<td>9</td>
<td>-4.0</td>
</tr>
<tr>
<td>CCC and below</td>
<td>1</td>
<td>-15.0</td>
</tr>
<tr>
<td>Default</td>
<td>1.3</td>
<td>-50.0</td>
</tr>
</tbody>
</table>

In the hypothetical limiting case of an infinitely large portfolio where all bonds have the same default probability \( p \) and all pair-wise correlations between normally distributed proxy variables are equal to \( \rho \), the cumulative distribution for the fraction of defaults, \( X \), is

\[
\text{Prob} \left[ X \leq x \right] = N \left( \left( \sqrt{1/\rho} - 1 \right) N^{-1} \left( x \right) - \left( 1/\sqrt{\rho} \right) N^{-1} \left( p \right) \right). \tag{19}
\]

Despite its simplicity, this distribution has been shown by Schoenbucher (2003) to give a good approximation of the fraction of defaulting US corporate bonds over the years 1970-2000. His best fit parameter estimates are \( p = 1.31\% \) and \( \rho = 11.21\% \).

Basel II Foundation Internal Rating Based method (FIRB) adopts the same assumptions implicitly when it sets a capital requirement for credit losses based on a 99.9\% level of confidence that it will not be exceeded, and uses the formula:

\[
\text{Capital requirement} = (\text{loss given default}) \times N^{-1} (z) \times (\text{Maturity Adjustment}),
\]

where

\[
z = \left( 1/\sqrt{1 - \rho} \right) \left( N^{-1} (p) + \sqrt{\rho} N^{-1} (0.999) \right),
\]

and \((\text{Maturity Adjustment})\) is an increasing function of maturity equal to 1 when the maturity is 2.5 years. But there is no attempt in Basel II to aggregate default risks with credit spread risks. Default risks are estimated for instruments kept in a banking book where market risks - including credit spread risks - are ignored, whereas credit spread risks are estimated as specific market risks for instruments kept in a trading book where default risks are ignored. CreditMetrics aggregates default risks with credit spread risks using simulations. Analytical approximations for very large portfolios have been proposed by Lucas, Klaassen, Spreij, and Satetmans (2001). To illustrate, we generate a distribution of return aggregating default and credit spread uncertainties for a very large portfolio of similar corporate bonds. The parameters – transition probabilities and corresponding changes in excess returns – are as given in Table VIII. They are representative of 5-year BB rated corporate bonds in normal times. The bond default probabilities and the pair-wise correlations between the proxy variables are the best fit estimators from Schoenbucher (2003). The bonds are assumed to be priced to yield a nil expected excess return.

Figure 7 displays the yearly excess return distribution generated with these data. It has a

\[35\]Because capital requirements are designed to cover unexpected losses only, the Basel II FIRB formula is adjusted by taking away expected losses in as much as they may already be covered by provisions.

\[35\]
standard deviation of 1.16%, a skewness of -1.37 and an excess kurtosis of 3.87. These last two characteristics should make it less attractive to risk-averse HARA utility investors than a Normal distribution with same nil expected return and variance.

To obtain a GSR in the range normally required by investors for traditional assets, typically between 0.4 and 0.8, the credit spread needs to be raised to a level such that the risk-neutral implied default probability is two to three times larger than the estimated default probability. This is illustrated by the two curves in Figure 8. The top curve expresses the GSR(\(\mid F; N, E\)) assuming, falsely, that credit risks are normally distributed. The lower curve expresses the GSR(\(\mid F; p, E\)) based on the skewed return distribution produced by our large portfolio approximation. In the range of interest, GSR(\(\mid F; p, E\)) is about 15% lower than GSR(\(\mid F; N, E\)). Put another way, negative skewness and excess kurtosis account for about 15% of the credit risk premium in the range of interest.

XI Summary and Conclusions

Rationalization means both logical explanation and suggesting the need to abandon; we do both by explaining the conditions leading to some preference criteria and suggesting to abandon those criteria for which no clear conditions can be identified. A preference criterion is only as relevant as the information it relies upon. A CER* expresses, in terms of equivalent sure return on total wealth, how much a risky asset is potentially worth to an investor, based on her probabilistic return forecasts, her risk attitude, and her investment alternatives. Increasing functions of CER*s are equivalent criteria. Among them, generalized Sharpe ratios (GSRs), when the risk-free asset is the only alternative investment, and generalized information ratios (GIRs), when a benchmark portfolio is also available, are more suitable for public reporting because they are
hardly sensitive to local absolute risk aversion which varies widely among investors.

In contrast, preference criteria that are not equivalent to $CER^*$s do not satisfy the fundamental axioms of choice supporting utility theory and indeed may violate them. Such are criteria based on return quantiles, or centered lower partial moments, or satisfying ‘coherence’ properties (e.g., conditional value-at-risk), to name only a few. In particular, the coherence properties of super-additivity and concavity are not necessary properties to ensure risk aversion; and proportionality to size of investment is incompatible with smooth (twice-differentiable) concave utility functions.

The evaluation of a $CER^*$ requires an investment optimization, usually through a numerical procedure. Analytical approximations of $CER^*$s and $GSR$s based on the first four moments of a return distribution and the sensitivity of the investor’s risk tolerance to wealth avoid this problem and give insight into preferences for skewness and kurtosis. Skewness is valued positively and excess kurtosis negatively with $HARA$ utilities. Preferences for skewness and kurtosis increase with the sensitivity of the investor’s risk tolerance to wealth. Our analytical approximations converge towards the exact criteria when return distributions approach normality. Otherwise they may be inaccurate and violate basic axioms of rational behavior such as respect for stochastic dominance; moment-based analytical approximations should therefore be used with care.

Rational preference criteria have many applications. $CER^*$s and equivalent criteria are especially useful to assess the attractiveness of assets with non-normally distributed returns. We find that $CER^*$s differ between an asset and options on this asset, even when the investor
agrees with the risk-neutral pricing of the options. Thus optimal portfolios should generally contain payoffs that are non-linear in assets returns, but probably not standard options. Optimal pay-offs are more likely to be smooth power functions of the underlying asset price, as approximated by CPPI strategies. Only unusual views or risk attitudes would justify more complex payoffs. CER*’s also contribute to explaining the credit spread puzzle; they penalize the negatively skewed returns of credit risk portfolios and thus contribute to explaining why credit spread implied default probabilities are generally much larger than historical default frequencies.

Objectivity demands that public reporting of funds performance be based on historical returns, but investors are interested in future performance. Probabilistic return forecasts over a standard investment period (e.g., 1 year) must therefore be inferred from historical returns, using minimal standard assumptions. Associations of investment professionals should set the standard calculation rules. They should also reflect investors’ risk attitudes. Fortunately, an investor’s local coefficient of risk tolerance is immaterial to the reporting of preferences in terms of GSRs and GIRs; what counts is the sensitivity of risk tolerance to wealth; a value of the order of 0.25 should satisfy the majority of investors. Finally they should express performance of funds in optimal combinations with a few broad benchmark portfolios (e.g., a balanced ‘market’ portfolio and possibly separate broad bond and equity portfolios). An indication of standard statistical error in the performance estimate – calculated according to basic rules – would be a useful complement as statistical errors are likely to be large.

Rational preference criteria have many other applications, both for research (e.g., explanation of risk premia) and for management and policy making (e.g., design of executive incentives, risk management). Conversely, it is confusing and even unethical to use heuristic preference criteria or performance measures without understanding the conditions under which they may be justified.

Acknowledgment

The author wishes to express his gratitude for comments and assistance received from colleagues and PhD students, especially from Johanna Scheller.
A Maximum Certainty Equivalent Excess Return of a Risky Investment: The \((F; N, E)\) Case

In the \((F; N, E)\) case defined in Section II, an investor’s risk attitude is described by an exponential utility function on the PV of future wealth, \(u(w) = -\exp(-w/\lambda), \lambda > 0\). She has unlimited access (in both long or short quantities) to a risk-free asset \(F\) and a risky asset offering a normally distributed excess return \(r \sim N(\mu, \sigma^2)\) per unit of time. We assume, without loss of generality, that the investor’s initial wealth is one unit of wealth.

The investor allocates her initial wealth in proportions \(\omega\) to the risky asset and \((1 - \omega)\) to the risk-free asset. The PV of her wealth after one unit of time is therefore \(w = 1 + \omega r\) and is normally distributed \(w \sim \Phi(1 + \omega \mu, (\omega \sigma)^2)\). The expected utility of this allocation is therefore:

\[
E[u(w)] = -\exp\left[-(1 + \omega \mu)/\lambda + \frac{1}{2} (\omega \sigma/\lambda)^2\right].
\]

By definition, it is equal to the utility of the certainty equivalent excess return \(CER(\omega|F; N, E)\), therefore:

\[
1 + CER(\omega|F; N, E) \equiv u^{-1}\{E[u(w)]\} = -\lambda \log(-E[u(w)]) = 1 + \omega \mu - (\omega \sigma)^2/(2\lambda),
\]

so

\[
CER(\omega|F; N, E) = \omega \mu - (\omega \sigma)^2/(2\lambda).
\]

The optimal allocation to the risky asset maximizing \(CER(\omega|F; N, E)\) is \(\omega^* = \lambda \mu/\sigma^2\) and the minimum sure return in excess of the risk-free rate that the investor would require for giving up the opportunity of an optimal allocation between the risky and the risk-free asset is:

\[
CER^*(\omega^*|F; N, E) = \frac{1}{2}\lambda (\mu/\sigma)^2.
\]

B Homogeneity of Degree One and First Order Approximation of Risk Premium

A preference criterion \(C(.)\) is positive homogeneous of degree one if its value varies linearly with the size of the investment. That is, for any investment \(X\),

\[
C(aX) = a.C(X), \text{ for all } a \geq 0.
\]

A preference criterion \(C(.)\) is said to display strict risk aversion if for any non-zero investment \(X\) with expected value \(\mu\), the risk premium \(R(.)\), defined as the difference between the values of the criterion for \(\mu\) and for \(X\), is greater or equal to zero with equality if and only if \(X = \mu\), that is

\[
R(X) = C(\mu) - C(X) \geq 0 \text{ with } R(X) = 0 \text{ iff } X = \mu.
\]

Therefore, if a preference criterion is positive homogeneous of degree one, so is the corresponding risk premium.

---

\[36\text{The normality assumption applies to the period of time over which the rates are defined. It cannot be preserved with compounding over multiple periods.}\]
We now consider a certainty equivalent (CE) as a preference criterion calculated from a twice-differentiable utility function. By definition:

\[ u [CE (X)] \equiv E [u (X)] . \]

Taking a Taylor series expansion on both sides of this definition for deviations of \( X \) from its expected value \( \mu = E [X] \), we obtain approximately:

\[
u + u'. (CE (X) - \mu) \cong u + u' . E [(x - \mu)] + \frac{1}{2} u'' . E [(x - \mu)^2],
\]

with utility \( u \) and first and second order derivatives \( u' \) and \( u'' \) evaluated at \( \mu \), so that

\[
R (X) = \mu - CE (X) \cong -\frac{1}{2} \left( \frac{u''}{u'} \right) \sigma^2 .
\]

Therefore, since \( u' \) is always positive (non-satiation principle), \( u'' \) must be negative for the risk premium of small investments to be positive, that is, the utility function must be concave for risk-averse investors. The risk premium is approximately proportional to the variance of the risk. It is therefore at least degree 2 in investment size.

Thus, the CE of an investment calculated with a twice-differentiable utility function cannot be homogeneous of degree 1 in investment size. This property is not the same as homogeneity degree 1 in final wealth. With some utility functions, in particular HARA utilities (6), the CE is homogeneous degree 1 in final wealth. But that does not mean that the CE of an investment is proportional to the initial investment amount. It is central to the concept of risk aversion that the risk premium is not scalable with the size of an investment, as the expected return is. Indeed, with twice-differentiable, concave utility functions, there is always a small investment amount below which a risky investment with positive expected excess return is attractive (i.e., has a positive CE) and an investment amount beyond which it appears less attractive than an investment in a risk-free asset (i.e., has a negative CE).

C Maximum Certainty Equivalent Excess Return of a Risky Asset above a Benchmark Portfolio: The \((F, B; N, E)\) Case

We revisit the calculations in Appendix A, adding a benchmark portfolio with excess return distribution \( r_B \sim \Phi (\mu_B, \sigma_B^2) \) and correlation coefficient \( \rho \) with risky asset \( A \), the active portfolio. The investor’s unit wealth is allocated in proportions \( \omega \) to the active portfolio, \( \omega_B \) to the benchmark portfolio, and \((1 - \omega - \omega_B)\) to the risk-free asset. We seek the unconstrained optimal allocation maximizing the CER of an investor with exponential utility function \( u (w) = -\exp (-w/\lambda) \), \( \lambda > 0 \). The PV of the wealth of the investor after a unit investment period is:

\[
w = 1 + \omega r + \omega_B r_B ,
\]

with final distribution

\[
w \sim N \left( 1 + \omega \mu + \omega_B \mu_B , (\omega \sigma)^2 + 2 \rho \omega \sigma \omega_B \sigma_B + (\omega_B \sigma_B)^2 \right) ,
\]
and therefore its $CER$ is:

$$CER(\omega, \omega_B|F; N, E) = \omega \mu + \omega_B \mu_B - [(\omega \sigma)^2 + 2\rho \omega \sigma_B \sigma + (\omega_B \sigma_B)^2]/(2\lambda).$$

The optimal allocations maximizing $CER(\omega, \omega_B|F; N, E)$ are:

$$\omega^* = \frac{(\mu - \mu_B \rho \sigma/\sigma_B)}{\sigma^2 (1 - \rho^2)},$$

$$\omega_B^* = \frac{(\mu_B - \mu \rho \sigma_B/\sigma)}{\sigma_B^2 (1 - \rho^2)}.$$

They yield the maximum $CER$:

$$CER^* (\omega^*, \omega_B^*|F; N, E) = \frac{1}{2} \lambda \frac{[(\mu/\sigma)^2 - 2\rho (\mu/\sigma) (\mu_B/\sigma_B) + (\mu_B/\sigma_B)^2]}{(1 - \rho^2)}.$$

Subtracting the $CER^*$ with an optimal investment in the benchmark portfolio only, that is:

$$CER^* (B|F; N, E) = \frac{1}{2} \lambda (\mu_B/\sigma_B)^2,$$

yields the $CER^*$ contribution of the risky asset above the benchmark:

$$\Delta CER^* (A|F, B; N, E) = \frac{1}{2} \lambda \frac{[(\mu/\sigma)^2 - \rho (\mu_B/\sigma_B)]^2}{(1 - \rho^2)}.$$

To gain further insight in this result, assume that the benchmark portfolio is a highly diversified market portfolio, $M$, so that, according to a CAPM model

$$\mu = \alpha + \beta \mu_M,$$

and call $\sigma_\epsilon$ the specific risk of the risky asset with respect to the market portfolio. Then we obtain:

$$\omega^* = \frac{\lambda \alpha}{\sigma_\epsilon^2},$$

$$\Delta CER^* (.|F, M; N, E) = \frac{1}{2} \lambda (\alpha/\sigma_\epsilon)^2,$$

where $\alpha/\sigma_\epsilon$ is recognized as the information ratio. In the same way as we have defined a generalized Sharpe ratio in (8), we can therefore define a generalized information ratio as:

$$GIR (.|p, u) \equiv \sqrt{(2/L(0)) CER^* (.|F, M;p, u)},$$

where $L(0)$ is the current local coefficient of risk tolerance, $\lambda$ in the case of a negative exponential utility function.

Either $\Delta CER^* (.|F, M;p, u)$, or the equivalent generalized information ratio $GIR (.|p, u)$, calculated with forecast return distribution and personal utility function, are suitable preference criteria for ranking the attractiveness of risky assets when they are considered in combination with a benchmark portfolio, irrespective of any choice of equilibrium asset pricing model. In the
special case of bivariate Normal return distribution, CAPM and constant absolute risk tolerance, the \( GIR \) reduces to the well-known information ratio. Alternatively, calling \( M_F \) the optimal allocation to the market portfolio \( M \) in the \((M, F)\) portfolio, that is, with \( \omega^*_M = \lambda \mu_M / \sigma^2_M \), the optimal allocation to the active portfolio \( A \) in \((A, M_F, F)\) is:

\[
\omega^* = \frac{\lambda}{\sigma^2} (\mu - \beta \mu_M) = \frac{\lambda \alpha}{\sigma^2}
\]

The corresponding maximum increase in \( CER \) is:

\[
\Delta CER^* (\cdot | F, M_F; N, E) = \frac{1}{2} \lambda (\alpha / \sigma)^2.
\]

An equivalent criterion is therefore the ratio of the excess return on \( A \) relative to the market portfolio divided by the total risk of \( A \), which can be generalized as:

\[
GIR (\cdot | p, u) \equiv \sqrt{(2/L(0)) \Delta CER^* (\cdot | F, M_F; p, u)}.
\]

In this instance where the allocation to the market portfolio is held fixed, the active portfolio \( A \) contributes its total risk, whereas in the previous instance the allocation to the market portfolio was rebalanced so that the active portfolio contributed only its specific risk. We do not find any general conditions under which the well-known Treynor ratio \((\mu / \beta)\) and the \((\alpha / \beta)\) ratio would correspond to a \( CER \) maximization. Together with related criteria (e.g., Modigliani and Modigliani (1997) \( M^2 \), Scholz and Wilkens (2005) MRAP), they should be regarded as ad hoc and therefore difficult to interpret and best avoided.

**D Compatibility of Common Preference Criteria with Sensible Utility Functions**

We review several families of commonly used preference criteria and seek whether they are compatible with sensible utility functions, that is twice-differentiable with first order derivative \( u' \geq 0 \) and second order derivative \( u'' \leq 0 \).

**Proposition 1a:** Only utility functions that are in the class of power functions of a return below a threshold and nil above this threshold are compatible with preference criteria that are monotonically decreasing functions of a lower partial moment of degree \( n \geq 0 \) below the threshold and do not depend on any other characteristic of the return distribution (e.g., opposite of a lower tail probability, or opposite of a put value struck at the threshold, both under the physical distribution).

**Proof:** For any return distribution and threshold \( \tau \), we must have

\[
E_P [u (r)] = E_P [h ((\max (\tau - r), 0)^n)] ,
\]

where \( h(\cdot) \) is a monotonically decreasing function. This implies that \( u(r) = h ((\max (\tau - r), 0)^n) \) almost everywhere and, because the expectation operator is linear, \( h(\cdot) \) must also be linear. So, compatible utility functions are defined, within a positive linear transformation, as

\[
u (r) = -(\max (\tau - r), 0)^n , \text{ almost everywhere.}
\]
Utility functions in this family are concave when \( n \geq 1 \) and twice-differentiable when \( n \geq 2 \).

**Proposition 1b:** Only utility functions that are in the class defined in **Proposition 1a** plus a positive linear component are compatible with preference criteria determined by a lower partial moment and an unconditional expected value (e.g., call less two puts, all struck at the threshold).

**Proof:** Following the same argument as in **Proposition 1a**, compatible utility functions are defined, within a linear transformation, as

\[
u(r) = \alpha r - \max(\tau - r, 0)^n, \quad \alpha > 0, \text{ almost everywhere.}\]

Utility functions in this family are concave when \( n \geq 1 \) and twice-differentiable when \( n \geq 2 \).

**Proposition 2:** No utility function is compatible with preference criteria that are function of a normalized lower partial moment of order \( n \neq 1 \) only, or combined with an expected value (e.g., mean less \([LPM(2, \tau)]^{1/2}\)). The case \( n = 1 \) is covered in **Proposition 1a**.

**Proof:** The argument in **Proposition 1a** forces the criteria to be linear in \( LPM(n, \tau) \). Conversely, if the criteria were not linear in \( LPM(n, \tau) \), Jensen’s inequality would show that no utility function is compatible with the criteria.

**Proposition 3:** No utility function is compatible with preference criteria that are function of a normalized centered lower partial moment \([LPM(n, \mu)]^{1/n}\) only or combined with an expected value (e.g., ‘coherent’ preference criterion such as expected value less semi-standard deviation).

**Proof:** It follows from **Proposition 2** that only the case \( n = 1 \) could be compatible with a utility function and from **Propositions 1b** that this utility function would be of the form \( u(r) = \alpha r - \max(\mu - r, 0) \) for all \( \mu \). Therefore the utility function can only be linear and, contrary to the premise, cannot depend on the normalized centered lower partial moment.

**Proposition 4:** No utility function is compatible with preference criteria that are function of a quantile. In particular, no utility function is compatible with \( \text{VaR}(\alpha) \) and expected shortfall \( \text{ES}(\alpha) \) preference criteria.

**Proof:** All return distributions sharing the same preference criteria and the same \( \alpha \)-quantile, \( r_{\alpha} \), should have the same expected utility. The only possibility are constant utilities \( \alpha_- \) below \( r_{\alpha} \) and \( \alpha_+ \) above \( r_{\alpha} \). As this must be true for all possible values of \( r_{\alpha} \), we must have \( \alpha_- = \alpha_+ \), hence utilities must be constant, which contradicts the premise that the preference criteria are function of a quantile.

### E  Risk Aversion, Concavity, and Coherence

We stressed in Section II that a preference criterion should be able to reflect risk aversion. When the preference criterion is an expected utility, it follows directly from Jensen’s inequality that the utility function must be concave. The corresponding CER also displays risk aversion since it is an equivalent criterion.
But risk aversion should not be equated with concavity of the criterion itself. By definition, a preference criterion is a concave function of investment allocations if for any two risky investments $X$ and $Y$, regardless of their joint distribution, we have:

$$C(aX + (1 - a)Y) \geq aC(X) + (1 - a)C(Y),$$

for any $0 \leq a \leq 1$. For example, a $CE$ calculated with a concave utility function always displays risk aversion but is not necessarily concave as the following trivial, if contrived, example shows. Consider a piece-wise linear utility function $u(w) = 2aw$ for $w < 0$ and $u(w) = aw$ for $w \geq 0$. With $a > 0$, it is monotonically increasing and concave.

A risky investment $X$ produces either 0 or +1 with equal probabilities; a risky investment $Y$, independent of $X$, produces either 0 or -1 with equal probabilities. Consider now an investment half in $X$ and half in $Y$. The expected return is nil and the $CE$ is negative because of risk aversion (the exact value is -1/16). On the other hand, the certainty equivalents of $X$ and $Y$ taken separately are equal to their expected values, that is +0.5 and -0.5, respectively, since the utility function is linear over their respective domains. The equal mix of the two certainty $CE$s is therefore nil. This proves convexity (the opposite of concavity) of the $CE$ criterion in this case. Conversely, concavity of the preference criterion does not by itself imply risk aversion. For example, the expected value criterion is concave but not strictly risk-averse.

These results challenge an attempt to design preference criteria from an axiomatic design of risk metrics. The intent of a risk metrics is to quantify how ‘bad’ a risk is, whereas a preference criterion is a measure of ‘good’ the returns are relative to the risk taken. So the opposite (or the inverse) of a risk metric can be used to build a preference criterion. For example, quantiles or $RoVaR$s are preference criteria based on the popular $VaR$ metric. But there are some disadvantages to $VaR$ as a risk metric; among others, $VaR$ cannot be used safely for risk budgeting because it is not a sub-additive measure of risk. In other words, $VaR(a)$ for the sum of several activities is not always lower or equal to the sum of the $VaR(a)$ of the individual activities. Therefore $VaR(a)$ limits imposed on individual activities do not ensure a $VaR(a)$ limit for the combined activities.

The inadequacy of $VaR$ for risk budgeting has prompted a search for desirable properties of risk metrics. Artzner, Delbean, Eber, and Heath (1999) state four properties, including sub-additivity, that a ‘coherent’ risk metric should possess. We translate these properties into to define a coherent preference criterion, $Coh$, as follows:

37 Compared to properties for coherent risk metrics, inequalities in (P2) and (P4) have to be reversed and -b replaces +b on the r.h.s. of (P1).

38 Weak stochastic dominance (order 1) was originally stated by Artzner, Delbean, Eber, and Heath (1999).

(P1) Risk-free translation: $Coh(X + b) = Coh(X) + b$, for any scalar $b$.

(P2) Monotonicity: $Coh(X) \geq Coh(Y)$, if $X$ stochastically dominates $Y$.

(P3) Positive homogeneity: $Coh(aX) = aCoh(X)$, for any scalar $a \geq 0$.

(P4) Super-additivity: $Coh(X + Y) \geq Coh(X) + Coh(Y)$ for any $X, Y$.

Taken together (P3) and (P4) also ensure concavity. Concavity of a preference criterion favors diversification. It shows that the satisfaction gained from any weighted average of two investments is greater or equal to the corresponding weighted average of the satisfactions gained from
each investment. As an illustration, preference criteria of the form \( \mu - [LPM (n, \mu)]^{1/n} \) satisfy all four coherence axioms, whereas a quantile (or \(-VaR(\alpha)\)) satisfies only the first three and is therefore not a coherent preference criterion.

One cannot question the appropriateness of Artzner’s axioms of coherence without specifying more precisely the role of a risk metric. Indeed, one can doubt whether a single risk metric can be used in all applications. A great deal of information is lost when reducing a probability distribution to a single number regardless of the decision context and the risk attitude of the decision maker. At any rate, these axioms do not necessarily deliver preference criteria capable of displaying risk aversion. We have already shown that concavity is neither necessary nor sufficient for risk aversion.

To impose risk aversion Acerbi (2002) requires two additional properties:

(P5) **Estimability:** The preference criterion is fully determined by the return distribution

(P6) **Co-monotonicity:** \( Coh(X + Y) = Coh(X) + Coh(Y) \) if \( y \) is an increasing (deterministic) function of \( x \)

This leads to the class of preference criteria known as spectral indices, a simple representative of that class is the expected shortfall, \( ES_\alpha \), defined as:

\[
ES(\alpha) \equiv E[r|r < r_\alpha]/\alpha = r_\alpha - LPM(1, r_\alpha)/\alpha.
\]

But because spectral indices satisfy (P3), they produce a risk premium that is proportional to the size of a risky investment. This runs contrary to our initial observation (Section I) that a risk premium should increase faster than the size of the risky investment and to such an extent that for any risky investment there is a size beyond which the risk becomes unacceptable.

In brief, the properties defining coherent and spectral indices are not all necessary properties for preference criteria and the homogeneity property (P3) is undesirable. One can also argue that unless a risk metric reflects an investor’s risk attitude, it is simply a measure of uncertainty (or dispersion).

### F Certainty Equivalent Returns of Optimal Return Profiles

We calculate the \( CER^* \)'s of the optimal managed return profiles derived in Section IX for investors with \( HARA \) utilities when the underlying assets are either normally or log-normally distributed.

In the \((N, E)\) case the optimal managed return is given by (12), the forecast excess return distribution is \( r \sim N(\mu, \Sigma) \), and the utility function is \( u(w) = -\exp(-w/\lambda) \). To be consistent with the definition of period returns, the \( CER^* \) of a unit investment is defined so that the certainty equivalent of the PV of the final value of the investment is equal to \( 1 + CER^* \). So, by definition:

\[
\begin{align*}
    u(1 + CER^*) &= E_p[u(w)], \\
    -\exp(-CER^*/\lambda) &= E_p[-\exp(-\mu^T\Sigma^{-1}r)].
\end{align*}
\]
hence

\[ CER^* = \frac{1}{2} \lambda \mu^T \Sigma^{-1} \mu. \]  \hspace{1cm} \text{(F.2)}

In the Log-normal, HARA case, the optimal managed return is given by (14), the forecast excess log-return distribution is \( r \sim N(\mu, \Sigma) \), and the utility function is given by (6), so, after simplification,

\[
\begin{align*}
    u(w(r, T)) &= \text{sign}(\eta - 1) \exp \left( (1 - 1/\eta) \left( m^T r T - \frac{1}{2} m^T \Sigma m T \right) \right), \\
    E_P[u(w(r, T))] &= \text{sign}(\eta - 1) \exp \left( (1 - 1/\eta) \left( m^T \mu T - \frac{1}{2 \eta} m^T \Sigma m T \right) \right).
\end{align*}
\]

To be consistent with the definition of log-returns, the \( CER^* \) of a unit investment is defined so that the certainty equivalent of the PV of the value of the investment is equal to \( \exp (CER^* : T) \).

Equating the utility of the \( CER \) to the expected utility above, gives:

\[
1 + \left( \frac{\eta}{\lambda} \right) (\exp(CER^* : T) - 1) = \exp \left( m^T \mu T - \frac{1}{2 \eta} m^T \Sigma m T \right),
\]

\[
CER^* : T = \log \left( 1 + \left( \frac{\lambda}{\eta} \right) \left( \exp \left( m^T \mu T - \frac{1}{2 \eta} m^T \Sigma m T \right) - 1 \right) \right). \hspace{1cm} \text{(F.3)}
\]

Expression F.3 applies also to the Normal case when the risky asset returns are normally distributed. When \( CER^* : T \ll 1 \), \( CER^* \) in F.3 reduces to F.2; that is also the case at the exponential utility limit \( (\eta \downarrow 0) \).


Meucci, A., 2005, Risk and asset allocation (Berlin: Springer).


