Available online at www.sciencedirect.com
SCIENCE DIRECT
JOURNAL OF Econometrics
Journal of Econometrics 116 (2003) 365–386
www.elsevier.com/locate/econbase

Portfolio choice with endogenous utility: a large deviations approach

Michael Stutzer*

Burridge Center for Securities Analysis and Valuation, Leeds School of Business, University of Colorado, Boulder, CO 20309-0419, USA

Abstract

This paper provides an alternative behavioral foundation for an investor's use of power utility in the objective function and its particular risk aversion parameter. The foundation is grounded in an investor's desire to minimize the objective probability that the growth rate of invested wealth will not exceed an investor-selected target growth rate. Large deviations theory is used to show that this is equivalent to using power utility, with an argument that depends on the investor's target, and a risk aversion parameter determined by maximization. As a result, an investor's risk aversion parameter is not independent of the investment opportunity set, contrary to the standard model assumption.

JEL classification: C4; D8; G0

Keywords: Portfolio theory; Large deviations; Safety-first; Risk aversion

1. Introduction

What criterion function should be used to guide personal investment decisions? Perhaps the earliest contribution was Bernoulli's critique of expected wealth maximization, which led him to advocate maximization of the expected log wealth as a resolution of the St. Petersburg Paradox. This was resurrected as a long-term investment strategy in the 1950s, and is now synonymously described as either the log optimal, growth-maximal, geometric mean, or Kelly investment strategy. As also noted in their

* Tel.: +1-319-335-1239.
E-mail address: michael-stutzer@colorado.edu (M. Stutzer).

0304-4076/03/$ - see front matter
doi:10.1016/S0304-4076(03)00112-X

excellent survey on this portfolio selection rule, Hakansson and Ziemba (1995, pp. 65–70) argue that "…the power and durability of the model is due to a remarkable set of properties", e.g. that it "*almost surely* leads to more capital in the long run than any other investment policy which does not converge to it".[1]

But even as a long-term investment strategy, the log optimal portfolio is problematic. It often invests very heavily in risky assets, which has led several researchers to highlight the possibilities that invested wealth will fall short of investor goals, even over the multi-decade horizons typical of young workers saving for retirement. For example, MacLean et al. (1992, p. 1564) note that "the Kelly strategy never risks ruin, but in general it entails a considerable risk of losing a substantial portion of wealth". Findings like these motivated Browne (1995, 1999a) to develop a variety of alternative, shortfall probability-based criteria, in specific continuous-time portfolio choice problems. Browne (1999b) considers these ideas in the context of the simplest possible investment decision, which will also be utilized to illustrate the criterion developed herein. Further discussion of his work is included in Section 2.3. Another similarly motivated criterion for continuous time portfolio choice is developed in Bielecki et al. (2000), which will be discussed further in Section 2.2.

The problem is exacerbated when investors have specific, short to medium term values for their respective investment horizons. If so, some criteria will lead to horizon-dependent optimal asset allocations, but others will not. For example, Samuelson (1969) proposes the criterion of intertemporal maximization of expected discounted, time-additive constant relative risk aversion (CRRA) power utility of consumption. He proves that when asset returns are IID, portfolio weights are independent of the horizon length. So in that case, long horizon investors should not invest more heavily in stocks than do short horizon investors. Samuelson (1994) provided caveats to this investor advice, citing six modifications of this specification that will result in horizon dependencies.[2]

But an investment advisor, hired to help an investor formulate asset allocation advice, may have difficulty determining a specific value for the investor's horizon. The advisor may be unable to determine an investor's exact horizon length when it exists, while other investors may not have a specific investment horizon length at all. A considerable simplification results when an infinite horizon is assumed, as has also been done when deriving many, but not all, consumption-based asset pricing models.[3] An exception to the infinite horizon formulations is found in Detemple and Zapatero (1991). Of course, the cost of this simplification is the inability to model horizon dependencies.

While the time horizon parameter is irrelevant for Samuelson's intertemporal power utility investor with IID returns, the optimal asset allocation is still very sensitive to the specific risk aversion parameter adopted, so an advisor would have to determine it with

[1] See the analysis of Algoet and Cover (1988) and the lucid exposition of Cover and Thomas (1991, Chapter 15) for more information on the growth maximal portfolio problem. For a spirited normative defense of the growth maximal portfolio criterion, see Thorp (1975).

[2] However not all of these modifications would support the oft-repeated advice to invest more heavily in stocks when the investor's horizon is longer.

[3] For a survey, see Kocherlakota (1996).

precision. An even more basic consideration is specification of the utility functional form and its argument. Should it be a power function, or an exponential function, or perhaps some function outside the HARA class? Should the argument be a function of current wealth, current consumption, or some function of the consumption path (as in habit formation models)? As a first step toward answering these questions, Section 2 of this paper develops a new criterion of investor behavior. It starts from the observation that the realized growth rate of investor wealth is a random variable, dependent on the returns to invested wealth and the time that it is left invested (i.e. the investment horizon). To obviate the need to specify a value for the latter, first assume that an investor acts as-if she wants to ensure that the (horizon-dependent) realized growth rate of her invested wealth will exceed a numerical target that she has, e.g. 8% per year. By choosing a portfolio that results in a higher expected growth rate of wealth than the target rate, the investor can ensure that the probability of not exceeding the target growth rate decays to zero asymptotically, as the time horizon $T \to \infty$. But the probability that the realized growth rate of wealth at finite time T will not exceed the target might vary from portfolio to portfolio. Which portfolio should be chosen? Without adopting a specific value of T, a sensible strategy is to choose a portfolio that makes this probability decay to zero as fast as possible as $T \to \infty$. This will ensure that the probability will be minimized for all but the relatively small values of T. In other words, the decay rate maximizing portfolio will maximize the probability that the realized growth rate will exceed the target growth rate at time T, for all but relatively small values of T. In fact, this turns out to be true for all values of T in the special IID cases studied in Sections 2.1 and 3.

Calculation of the decay rate maximizing portfolio is enabled by use of a simply stated, yet powerful result from large deviations theory, known as the Gärtner–Ellis Theorem. Straightforward application of it in Section 2.2 provides an expected power utility formulation of the decay rate criterion. But there are two important differences between this formulation and the standard expected power utility problem. First, the argument of the utility function is the ratio of invested wealth to a level of wealth growing at the constant target rate. Second, the value of the power, i.e. the risk aversion parameter, is also determined by maximization. As a result, a decay rate maximizing investor's degree of relative risk aversion will depend on the investment opportunity set, an effect absent in extant uses of power utility.

Because this endogenous degree of risk aversion is greater than 1, the third derivative of the utility is positive, so there is also an endogenous degree of skewness preference. This is fortunate, as some have argued that skewness preference helps explain expected asset returns. To see why, note that in the standard CAPM, investor aversion to variance makes an asset return's covariance with the market return a risk factor, so it is positively related to an asset's expected return. Kraus and Litzenberger (1976) argue that investor preference for positively skewed wealth distributions (ceteris paribus) should make market coskewness an additional factor, that should be negatively related to an asset's expected return. They thus generalized the standard CAPM to incorporate a market coskewness factor. The estimated model supports this implication of investor skewness preference. Harvey and Siddique (2000) extended this approach by incorporating conditional coskewness, concluding that "a model incorporating coskewness is helpful in explaining the cross-sectional variation of asset returns”. [4]

The decay rate maximization criterion also nests Bernoulli’s expected log maximization (a.k.a. growth optimal) criterion. An investor who has a target growth rate suitably close to the maximum feasible expected growth rate has an endogenous degree of risk aversion slightly greater than 1. As a result, the associated decay rate maximizing portfolio approaches the expected log maximizing portfolio. If the investor’s target growth rate is lower, the investor uses a higher degree of risk aversion, and the associated decay rate maximizing portfolio is more conservative, with a lower expected growth rate, but a higher decay rate for the probability of underperforming that target growth rate (and hence a higher probability of realizing a growth rate of wealth in excess of that target). The (perhaps unlikely) presence of an unconditionally riskless asset, i.e. one with an intertemporally constant return, provides a floor on the attainable target growth rates. When the target growth rate is sufficiently near that floor, the investor’s risk aversion will be quite high, and the associated decay rate maximizing portfolio will be close to full investment in the unconditionally riskless asset. The relationship between the target growth rate and the associated (maximum) decay rate of the probability that it will not be exceeded quantifies the tradeoff between growth and shortfall risk that has concerned analysts studying the expected log criterion.

Exact calculation of the decay rate (or equivalently, the expected power utility) requires the exact portfolio return process. In practice, the distribution is not exactly known. Even if its functional form is known, its parameters must still be estimated. To cope with this lack of exact knowledge, Section 3 adopts the assumption that portfolio log returns are IID with an unknown distribution, and follows Kroll et al. (1984) in estimating expected utility by substitution of a time average for the expectation operator. The estimated optimal portfolio and endogenous risk aversion parameter are those that jointly maximize the estimated expected power utility. An illustrative application of this estimator is included, contrasting decay rate maximization to both Sharpe Ratio and expected log maximization when allocating funds among domestic industry sectors. In it, decay rate maximization selects portfolios with higher skewness than Sharpe ratio maximization does. The IID assumption that underlies the estimator also permits the use of both a relative entropy minimizing, Esscher transformed log return distribution and a cumulant expansion to help interpret the empirical findings.

Section 4 summarizes the most important results, and concludes with some good topics for future research.

2. Porfolio analysis

Following Hakansson and Ziemba (1995, p. 68), the wealth at time T resulting from investment in a portfolio is $W_T = W_0 \prod_{t=1}^{T} R_{pt}$, where R_{pt} is the gross (hence positive)

[4] Hence, it is possible that an asset pricing model incorporating decay rate maximizing investors could outperform the CAPM, which incorporates Sharpe ratio maximizing investors. This topic is left for another paper.

rate of return between times $t-1$ and t from a portfolio p. Note that W_T does not depend on the length of the time interval between return measurements, but only on the product of the returns between those intervals. Dividing by W_0, taking the log of both sides, multiplying and dividing the right-hand side by T and exponentiating both sides produces the alterative expression

$$W_T = W_0 \left[\mathrm{e}^{\sum_{t=1}^{T} \log R_{pt}/T} \right]^{\mathrm{T}} = W_0 [\mathrm{e}^{\overline{\log R_p}}]^{\mathrm{T}}. \tag{1}$$

From (1), we see that W_T is a monotone increasing function of the realized time average of the log gross return, denoted $\overline{\log R_p}$, which is the realized growth rate of wealth through time T. When the log return process is ergodic in the mean, this will converge to a number denoted $\mathrm{E}[\log R_p]$, as $T \to \infty$. Accordingly, there was early (and still continuing) interest in the portfolio choice that maximizes this expected growth rate, i.e. selects the portfolio $\arg\max_p \mathrm{E}[\log R_p]$, also known as the "growth optimal" or "Kelly" criterion. As noted by Hakansson and Ziemba (1995, p. 65) "…the power and durability of the model is due to a remarkable set of properties", e.g. that it "*almost surely* leads to more capital in the long run than any other investment policy which does not converge to it". [5]

But maximizing the expected log return often invests very heavily in assets with volatile returns, which has led several researchers to highlight its substantial downside performance risks. Specifically, we will now examine the probability of the event that the realized growth rate of wealth $\overline{\log R_p}$ will not exceed a target growth rate $\log r$ specified by the investor or analyst. This is an event that will cause W_T in (1) to fail to exceed an amount equal to that earned by an account growing at a constant rate $\log r$. The following subsection uses a simple and widely analyzed portfolio problem to calculate this downside performance risk for the growth optimal portfolio and a portfolio chosen to minimize it.

2.1. The normal case

A simple portfolio choice problem, used in Browne (1999b), requires choice of a proportion of wealth p to invest in single stock, whose price is lognormally distributed at all times, with the rest invested in a riskless asset with continuously compounded constant return i. In this case, $\log R_{pt} \sim IID\ \mathcal{N}(\mathrm{E}[\log R_p], Var[\log R_p])$. We now compute the probability that $\overline{\log R_p} \leqslant \log r$. Because the returns are independent, $\overline{\log R_p} \sim \mathcal{N}(\mathrm{E}[\log R_p], Var[\log R_p]/T)$. The elementary transformation to the standard normal variate Z shows that the desired probability is

$$Prob[\overline{\log R_p} \leqslant \log r] = Prob\left[Z \leqslant \frac{\log r - \mathrm{E}[\log R_p]}{\sqrt{Var[R_p]/T}} \right]. \tag{2}$$

[5] See the analysis of Algoet and Cover (1988) and the lucid exposition of Cover and Thomas (1991, Chapter 15) for more information on the expected log criterion. For a spirited normative defense of this criterion, see Thorp (1975).

In order to minimize (2), i.e. to maximize the complementary probability that $\overline{\log R_p} > \log r$, one must choose the proportion of wealth p to minimize the expression on the right-hand side of (2). This is equivalent to maximizing -1 times this expression. Independent of the specific value of T, this portfolio stock weight is

$$\arg\max_p \frac{\mathrm{E}[\log R_p] - \log r}{\sqrt{Var[\log R_p]}}. \tag{3}$$

Portfolio (3) will differ considerably from the following growth optimal portfolio

$$\arg\max_p \mathrm{E}[\log R_p] \tag{4}$$

because of the presence of the target $\log r$ in the numerator of (3) and the standard deviation of the log portfolio return in its denominator. Portfolio (3) will also differ from the following Sharpe Ratio maximizing portfolio:

$$\arg\max_p \frac{E[R_p] - i}{\sqrt{Var[R_p]}} \tag{5}$$

because of the presence of the presence of the target $\log r$ in (3) in place of the riskless rate i in (5), and because of the presence of log gross returns in (3) in place of the net returns used in (5).

It will soon prove useful to reformulate the rule (3) in the following way. Note that $Prob[\overline{\log R_p} \leqslant \log r]$ will not decay asymptotically to zero unless the numerator of (3) is positive, so we need only consider portfolios p for which

$$\mathrm{E}[\log R_p] > \log r, \tag{6}$$

in which case the objective in problem (3) can be equivalently reformulated by squaring, and dividing by 2. The result is the following criterion:

$$\arg\max_p D_p(\log r) \equiv \arg\max_p \frac{1}{2}\left(\frac{\mathrm{E}[\log R_p] - \log r}{\sqrt{Var[\log R_p]}}\right)^2. \tag{7}$$

In order to quantitatively compare criteria (4), (5), and (7), it is useful to follow Browne (1999b) in using a parametric stochastic stock price process that results in the stock price being lognormally distributed at all times t, so that $\log R_{pt} \sim IID$ $\mathcal{N}(\mathrm{E}[\log R_p], Var[\log R_p])$ as assumed above. Specifically, the stock price S follows the geometric brownian motion with drift $\mathrm{d}S/S = m\,\mathrm{d}t + v\,\mathrm{d}W$, where m denotes the instantaneous mean parameter, v denotes the instantaneous volatility parameter, and W denotes a standard Wiener process. The bond price B follows $\mathrm{d}B/B = i\,\mathrm{d}t$. Denoting the period length between times t and $t+1$ by Δt, Hull (1993, p. 210) shows that

$$\mathrm{E}[\log R_p] = (pm + (1-p)i - p^2v^2/2)\Delta t, \tag{8}$$

$$Var[\log R_p] = p^2v^2\Delta t. \tag{9}$$

Now substitute Eqs. (8) and (9) into (7), and write down the first-order condition for the maximizing stock weight p. You can verify by substitution that the following

p solves it:

$$\arg\max_p D_p(\log r) = \sqrt{\frac{2(\log r - i)}{v^2}}. \tag{10}$$

Using (8), the growth optimal criterion (4) yields the portfolio

$$\arg\max_p E[\log R_p] = \frac{(m - i)}{v^2}. \tag{11}$$

Using Browne's (1999b, p. 77) parameter values $m = 15\%$, $v = 30\%$, $i = 7\%$, and a target growth rate $\log r = 8\%$, the outperformance probability maximizing rule (10) advocates investing a constant $p = 47\%$ of wealth in the stock, while the growth optimal rule (11) advocates $p = 89\%$. Of course, (10)'s $p = 47\%$ minimizes the probability that the realized growth rate $\overline{\log R_p} \leqslant 8\%$. Fig. 1 illustrates the phenomena, by graphing $Prob[\overline{\log R_p} \leqslant 8\%]$ for the two portfolios, and a third portfolio with just 33% invested in the stock. It shows that $Prob[\overline{\log R_p} \leqslant 8\%]$ decays to zero for all three portfolios, but decays at the fastest rate when (10)'s $p = 47\%$ is used. Section 2.2 will show that the rate of probability decay rate in Fig. 1 is $D_p(\log r)$ *in* (7). Fig. 1 also shows that even though investors can invest in a riskless asset earning 7%, and can try to beat the modest 8% target growth rate by also investing in a stock with an instantaneous expected return of 15%, there is still almost a 20% probability that the investor's realized growth rate of wealth after 50 years will be less than 8%!

Table 1 contrasts performance statistics for the outperformance probability maximizing portfolios and the growth optimal portfolio $p = 89\%$ over the feasible range of target growth rates $\log r$. Because the riskless rate of interest is only 7%, the probability of earning more than a target rate $\log r > 7\%$ is always less than one. If the target rate $\log r \leqslant 7\%$, the investor could always ensure outperforming that rate by investing solely in the riskless asset. Hence the lower limit of the feasible target growth rates is the 7% riskless rate.[6] Line 1 in Table 1 shows that in order to maximize the probability of outperforming a target growth rate one basis point higher than this, i.e. $\log r = 7.01\%$, the investor need invest only $p = 5\%$ of wealth in the stock. As a result of this conservative portfolio, this investor will have a relatively low probability of not exceeding this target; the decay rate of the underperformance probability is $\max_p D_p(7.01) = 3.19\%$. But by investing 89% of wealth in the stock, the growth optimal investor will have a higher probability of not exceeding this 7.01% target, because its associated decay rate is just 0.88%. This occurs despite its much higher expected growth rate $E[\log R_p]$ (10.6% vs. 7.4%) and higher expected net return $\mu = pm + (1 - p)i$ (14.1% vs. 7.4%). Of course, the major reason for this is its higher volatility $\sigma = pv$ (26.7% vs. 1.5%), which increases the probability that a bad series of returns will drive the growth optimal portfolio's realized growth rate below $\log r = 7.01\%$. Also note in line 1 that in order to maximize the probability of outperforming the 7.01% target, the investor must choose a portfolio with a higher expected growth rate (7.4%) than the target, as explained earlier.

[6] Of course, if a riskless rate does not exist, it would not provide a floor on the feasible target rates.

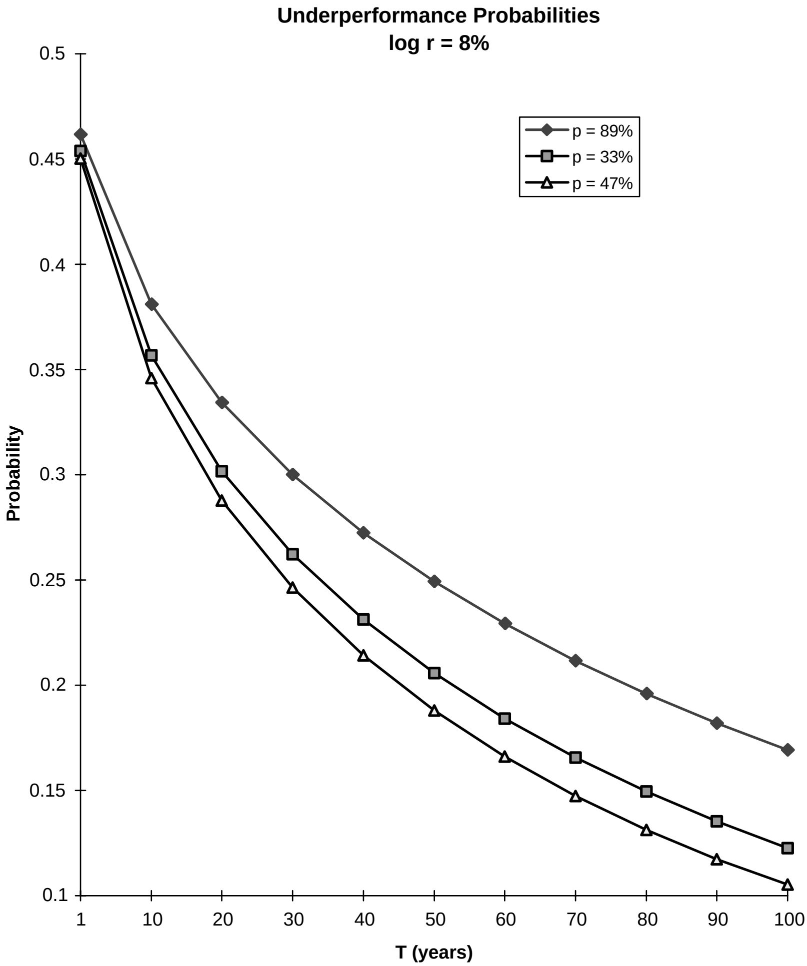

Fig. 1. The probability of not exceeding the 8% target growth rate approaches zero at a portfolio dependent rate of decay. The rate of decay is highest for the portfolio with $p = 47\%$ invested in the stock.

There is an important tradeoff present in Table 1. Note from columns 1 and 3 that investors with successively higher growth targets $\log r$ have successively lower underperformance probability decay rates $\max_p D_p(\log r)$. This implies that investors with higher targets will be exposed to a higher probability of realizing growth rates of wealth that do not exceed their respective targets. This occurs despite the fact that they did the best they could to minimize the probability of that happening. This is a consequence of the successively more aggressive portfolios needed to ensure asymptotic outperformance of their successively higher targets.

This tradeoff is analogous to the tradeoff between mean and standard deviation associated with the efficiency criterion that selects the portfolio with the smallest standard

Table 1
Performance statistics for the maximum expected log portfolio and maximum decay rate portfolios associated with feasible target growth rates $\log r$, when portfolios are formed from a lognormally distributed stock with $m=15\%$ instantaneous mean return and $v=30\%$ instantaneous volatility, and a riskless asset with instantaneous riskless rate $i = 7\%$

	Performance of portfolio (11) vs. portfolio (10)				
$\log r\%$	Stock weight $p\%$	$D_p(\log r)\%$	$E[\log R_p]\%$	$\mu\%$	$\sigma\%$
7.01	89 (5)	0.88 (3.19)	10.6 (7.4)	14.1 (7.4)	26.7 (1.5)
7.5	89 (33)	0.66 (1.39)	10.6 (9.2)	14.1 (9.6)	26.7 (9.9)
8.0	89 (47)	0.46 (0.78)	10.6 (9.8)	14.1 (10.8)	26.7 (14.1)
8.5	89 (58)	0.30 (0.44)	10.6 (10.1)	14.1 (11.6)	26.7 (17.4)
9.0	89 (67)	0.17 (0.22)	10.6 (10.3)	14.1 (12.4)	26.7 (20.1)
9.5	89 (75)	0.08 (0.09)	10.6 (10.5)	14.1 (13.0)	26.7 (22.5)
10.0	89 (82)	0.02 (0.02)	10.6 (10.5)	14.1 (13.6)	26.7 (24.6)
10.6	89 (89)	0.00 (0.00)	10.6 (10.6)	14.1 (14.1)	26.7 (26.7)

deviation of return, once the investor fixes a mean return. Here, the criterion selects the portfolio with the highest underperformance probability decay rate, once the investor fixes a target growth rate. In this way, the tradeoff between $\log r$ and $\max_p D_p(\log r)$ can be thought of as an alternative efficiency frontier, which yields the growth optimal portfolio on one extreme and full investment in the constant interest rate (when it exists) on the other. The efficiency frontier is graphed in Fig. 2, which shows it to be a convex curve in this example. In Section 2.2, we will see that this is true more generally, i.e. with multiple risky assets, whose log returns are not necessarily normal nor IID.

Finally, there is just one risky asset used to form the optimal portfolios in Table 1, so the mean-standard deviation efficiency frontier is just swept out by varying the stock weight and calculating the mean μ and standard deviation σ of the net returns. For comparison purposes, this is reported in the last two columns of Table 1. Reading down the last three columns of the table, note that the difference between μ and the expected growth rate $E[\log R_p]$ of wealth grows wider as the standard deviation of portfolio returns σ gets larger. The mean return increasingly overstates the expected growth rate of wealth as portfolio volatility increases. This is due to (1); as Hakansson and Ziemba (1995, p. 69) note, "...capital growth (positive or negative) is a multiplicative, not an additive process".[7] Here, due to lognormality, there is a precise relationship between the two: $E[\log R_p] = \mu - \sigma^2/2$ (see Hull, 1993, p. 212).

The following section will show that a simple, yet powerful result from large deviations theory permits us to rigorously characterize $D_p(\log r)$ in Table 1 as the decay rate of the portfolios' underperformance probabilities graphed in Fig. 1. More importantly, the result also shows how to correctly calculate the decay rate and associated decay rate maximizing portfolios when portfolio returns are not lognormally distributed.

[7] In this regard, see Stutzer (2000) for a simpler model of fund managers who use arithmetic average net returns rather than average log gross returns, under the assumption that net returns are IID.

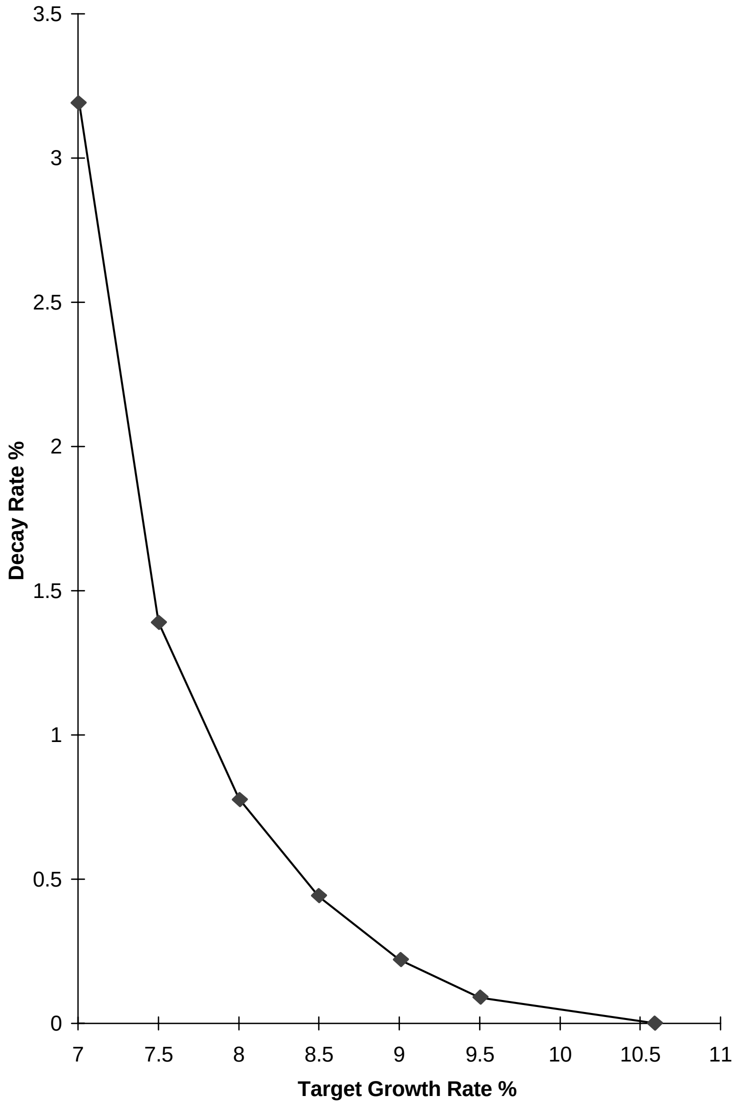

Fig. 2. The convex tradeoff between the target growth and underperformance probability decay rates for the optimal portfolios in Table 1. The convexity is generic.

2.2. The general case

As shown in the last section, when a portfolio's log returns were IID normally distributed, exact underperformance probabilities of the realized growth rate could be easily calculated using (2). But it is widely accepted that stock returns are often skewed and leptokurtotic. Even if they weren't, the skewed returns of derivative securities like stock options are inherently non-normally distributed. Hence there is an important need to rank portfolios according to their underperformance probabilities $Prob[\overline{\log R_p} \leqslant \log r]$ in non-IID, non-normal circumstances. It is now shown how to calculate the decay rate of this probability in more general cases. We will then apply the general result to prove that $D_p(\log r)$ in (7) is indeed the correct decay rate for the IID normal case.

As in the previous section, we seek to rank portfolios p for which the underperformance probability $Prob[\overline{\log R_p} \leqslant \log r] \to 0$ as $T \to \infty$. Calculation of this probability's decay rate $D_p(\log r)$ is facilitated by use of the powerful, yet simple to apply, Gärtner–Ellis Large Deviations Theorem, e.g. see Bucklew (1990, pp. 14–20). For a log portfolio return process with random log return $\log R_{pt}$ at time t, consider the following time average of the partial sums' log moment generating functions, i.e.:

$$\phi(\theta) \equiv \lim_{T\to\infty} \frac{1}{T} \log \mathrm{E}[\mathrm{e}^{\theta \sum_{t=1}^{T} \log R_{pt}}] = \lim_{T\to\infty} \frac{1}{T} \log \mathrm{E}\left[\left(\frac{W_T}{W_0}\right)^{\theta}\right], \tag{12}$$

where the last expression is found by using (1) to compute $W_T/W_0 = \prod_t R_{pt}$, substituting $\log(W_T/W_0)$ for the sum of the logs in (12), and simplifying. Hence (12) depends on the value of the random W_T, and so does not depend on the particular discrete time intervals between the log returns $\log R_{pt}$. We maintain the assumptions that the limit in (12) exists for all θ, possibly as the extended real number $+\infty$, and is differentiable at any θ yielding a finite limit. From the last expression in (12), these assumptions must apply to the asymptotic growth rate of the expected power of W_T/W_0. Some well-analyzed log return processes will satisfy these hypotheses, as will be demonstrated shortly by example. However, these assumptions do rule out some proposed stock return processes, e.g. the stable Levy processes with characteristic exponent $\alpha < 2$ and hence infinite variance, used in Fama and Miller (1972, pp. 261–274).

The calculation of the decay rate $D_p(\log r)$ is the following Legendre–Fenchel transform of (12):

$$D_p(\log r) \equiv \max_{\theta} \theta \log r - \phi(\theta). \tag{13}$$

When log returns are independent, but not identically distributed, (12) specializes to

$$\phi(\theta) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \log \mathrm{E}[\mathrm{e}^{\theta \log R_{pt}}] = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \log \mathrm{E}[R_{pt}^{\theta}]. \tag{14}$$

When log returns are additionally identically distributed (IID), (12) simplifies to

$$\phi(\theta) = \log \mathrm{E}[\mathrm{e}^{\theta \log R_p}] = \log \mathrm{E}[R_p^{\theta}], \tag{15}$$

which when substituted into (13) yields the decay rate calculation for the IID case. This result will form the basis for the empirical application in Section 3. It is known as *Cramer's Theorem* (Bucklew, 1990, pp. 7–9).

To illustrate these calculations, let us return to the widely analyzed case where the log portfolio return $\log R_{pt}$ is a covariance-stationary normal process with absolutely summable autocovariances. Then the partial sum of log returns in (12) is also normally distributed. The mean and variance of it can be easily calculated by adapting Hamilton's (1994, p. 279) calculations for the distribution of the sample mean, i.e. the partial sum divided by T. One immediately obtains

$$\log(W_T/W_0) \equiv \sum_{t=1}^{T} \log R_{pt} \sim \mathcal{N}\left(T\mathrm{E}[\log R_p],\ T\,Cov_0 + \sum_{\tau=1}^{T-1}(T-\tau)(Cov_\tau + Cov_{-\tau})\right), \tag{16}$$

where $\mathrm{E}[\log R_p]$ denotes the log return process' common mean and Cov_τ denotes its τ-lagged autocovariance. Formula (12) is the limiting time average of the log moment generating functions of these normal distributions. Now remember from elementary statistics that a normal distribution's log moment generating function is linear in its mean and quadratic in its variance. As a result, use (12) to calculate

$$\phi(\theta) = \lim_{T\to\infty} \frac{1}{T}\left(T\mathrm{E}[\log R_p]\theta + \left(T\,Cov_0 + \sum_{\tau=0}^{T-1}(T-\tau)(Cov_\tau + Cov_{-\tau})\right)\theta^2/2\right)$$
$$= \mathrm{E}[\log R_p]\theta + \sum_{\tau=-\infty}^{\tau=+\infty} Cov_\tau \theta^2/2. \tag{17}$$

Now substitute (17) into (13) and set its first derivative with respect to θ equal to zero to find that the maximum in (13) is attained by the following maximizer:

$$\theta_{\max} = (\log r - \mathrm{E}[\log R_p]) \Bigg/ \sum_{\tau=-\infty}^{\tau=+\infty} Cov_\tau\ . \tag{18}$$

Substituting (18) back into (17) and rearranging yields the underperformance probability decay rate

$$D_p(\log r) = \frac{1}{2}\left(\frac{\mathrm{E}[\log R_p] - \log r}{\sqrt{\sum_{\tau=-\infty}^{\tau=+\infty} Cov_\tau}}\right)^2. \tag{19}$$

Note that maximization of the decay rate (19) rewards portfolios with a high expected growth rate $E[\log R_p]$ (in its numerator) and a low asymptotic variance $\sum_{\tau=-\infty}^{\tau=+\infty} Cov_\tau \equiv \lim_{T\to\infty} Var[\log(W_T/W_0)]/T$ (in its denominator). This differs from the criterion in Bielecki et al. (2000), which is approximately the asymptotic expected growth rate minus a multiple of the asymptotic variance. For the IID case used in Section 2.1, all covariance terms in (19) are zero except $Cov_0 \equiv Var[\log R_p]$, so the decay rate function (19) reduces to the expression (7) used in Section 2.1 and Table 1. Fig. 3 depicts this decay rate function over a range of $\log r$, for each of the three portfolios whose underperformance probabilities are graphed in Fig. 1. There, we see that the portfolio $p = 47\%$ from (10) does indeed have the highest decay rate when $\log r = 8\%$.

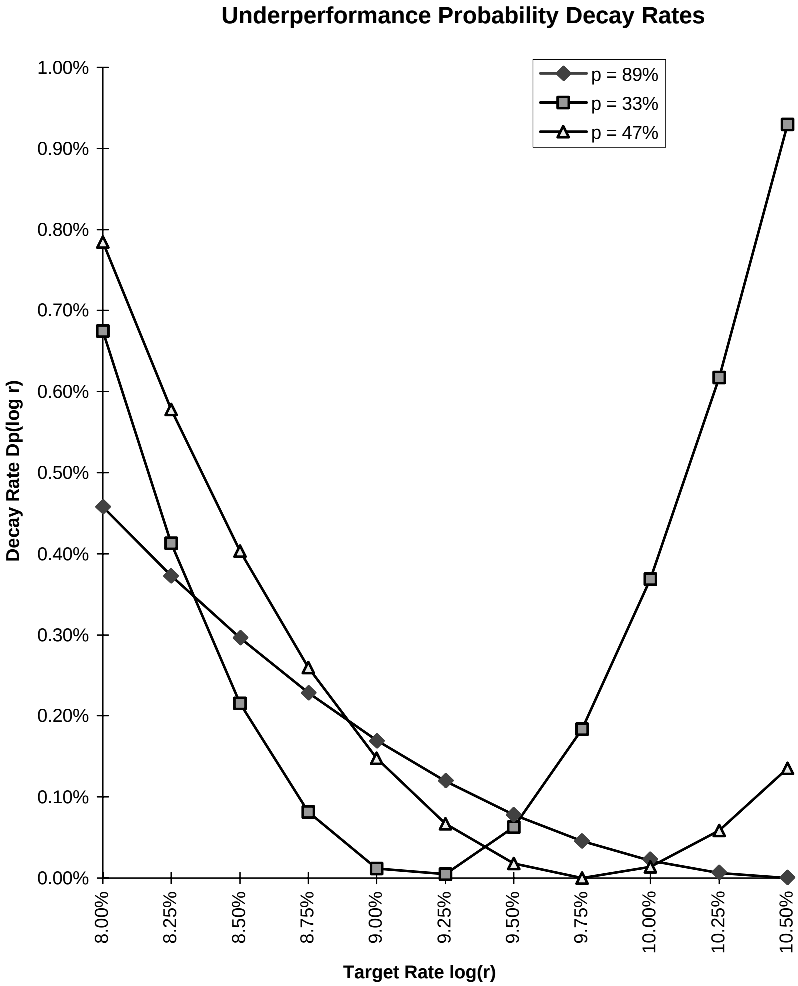

Fig. 3. The decay rate function $D_p(\log r)$ is convex, with a minimum at $\log r = E[\log R_p]$. The portfolio $p = 47\%$ attains the highest decay rate when $\log r = 8\%$.

Note that the decay rate function $D_p(\log r)$ in (19) for a covariance stationary Gaussian portfolio log return process is non-negative, and is a strictly convex function of $\log r$, achieving its global minimum of zero at the value $\log r = \mathrm{E}[\log R_p]$. These properties are true for more general processes (for a discussion, see Bucklew, 1990). As a result, remember from (6) that the decay rate criterion ranks portfolios with $\mathrm{E}[\log R_p] > \log r$, and apply the envelope theorem to the general rate function (13) to yield

$$\frac{\mathrm{d}D_p(\log r)}{\mathrm{d}\log r} = \frac{\partial}{\partial \log r} \max_\theta \theta \log r - \phi(\theta) = \theta_{\max} < 0 \tag{20}$$

as seen in the special case (18). Now differentiate (20) to find

$$\frac{\mathrm{d}D_p^2(\log r)}{\mathrm{d}\log r^2} = \frac{\mathrm{d}\theta_{\max}}{\mathrm{d}\log r} > 0 \tag{21}$$

due to convexity of $D_p(\log r)$. Again due to the envelope theorem, (20) and (21) continue to hold for $\max_p D_p(\log r)$ as well. Fig. 2 depicts the convexity of $\max_p D_p(\log r)$ over the relevant range of $\log r$ in the example of Table 1.

2.3. Analogy with power utility

The general decay rate criterion is a generalization of the expected power utility criterion. To uncover the generalization, substitute the right-hand side of (12) into (13), to derive

$$\begin{aligned}\max_p D_p(\log r) &\equiv \max_p \max_\theta \theta \log r - \lim_{T\to\infty} \frac{1}{T} \log \mathrm{E}\left[\left(\frac{W_T}{W_0}\right)^\theta\right] \\ &= \max_p \max_\theta \log r^\theta - \lim_{T\to\infty} \frac{1}{T} \log \mathrm{E}\left[\left(\frac{W_T}{W_0}\right)^\theta\right] \\ &= \max_p \max_\theta - \lim_{T\to\infty} \frac{1}{T} \log \mathrm{E}\left[\left(\frac{W_T}{W_0 r^T}\right)^\theta\right],\end{aligned} \tag{22}$$

which yields the following large T approximation:

$$-\,\mathrm{e}^{-\max_p D_p(\log r)T} \approx \max_p \mathrm{E}\left[-\left(\frac{W_T}{W_0 r^T}\right)^{\theta_{\max}(p)}\right], \tag{23}$$

where we write $\theta_{\max}(p)$ in (23) to stress dependence (through the joint maximization (22)) of θ on the portfolio p. The left-hand side of (23) increases with D_p, so a large T approximation of the portfolio ranking is produced by use of the expected power utility on the right-hand side of (23).

There are both similarities and differences between the right-hand side of (23) and a conventional expected power utility $\mathrm{E}[-(W_T)^\theta]$. From (20), $\theta_{\max}(p) < 0$. Evaluating it at the investor's decay rate maximizing portfolio p, note that the power function in (23) with the form $U = -(\bullet)^{\theta_{\max}(p)}$ increases toward zero as its argument grows to infinity, is strictly concave, and has a constant degree of relative risk aversion $\gamma \equiv 1 - \theta_{\max}(p) > 1$. Furthermore, $\theta_{\max}(p) < 0$ implies that the third derivative of U is positive, so the criterion exhibits positive skewness preference. But there are two important differences between the concepts. First, the argument of the power function in (23) is altered; it is the *ratio* of invested wealth to a "benchmark" level of wealth accruing in an account that grows at the geometric rate r. While absent from traditional criteria, this ratio is also present in other non-standard criteria, such as Browne's (1999a, p. 276) criterion to "maximize the probability of beating the benchmark by some predetermined percentage, before going below it by some other predetermined percentage". Browne (1999a, p. 277) notes that "...the relevant state variable is the *ratio* of the investor's wealth to the benchmark".[8] Second, conventional portfolio theory assumes that the risk aversion parameter θ is a preference parameter that is independent of the investment opportunity set. But in (23), $\theta = \theta_{\max}(p)$ is determined by maximization, and hence is *not* independent of the investment opportunity set. Investors could utilize different investment opportunity sets, either because of differential regulatory constraints, such as hedge funds' greater ability to short sell, or because of different opinions about the parameters of portfolios' log return processes. *When this happens, investors will have different decay rate maximizing portfolios* p, *and different degrees of risk aversion* $\gamma = 1 - \theta_{\max}(p)$, *even if they have the same target growth rate* $\log r$.

Assuming that asset returns are generated by a continuous time, correlated geometric Brownian process, Browne (1999a, p. 290) compares the formula for the optimal portfolio weights resulting from his criterion, to the formula resulting from conventional maximization of expected power utility at a fixed terminal time T. In this special case, he finds that the two formulae are isomorphic, i.e. there is a mapping between the models' parameters that equates the two formulae. He concludes that "there is a connection between maximizing the expected utility of terminal wealth for a power utility function, and the objective criteria of maximizing the probability of reaching a goal, or maximizing or minimizing the expected discounted reward of reaching certain goals". Connection (23) between decay rate maximization and expected power utility is quite specific, yet does not depend on a specific parametric model of the assets' joint return process.

Critics such as Bodie (1995, p. 19) have argued that "the probability of a shortfall is a flawed measure of risk because it completely ignores how large the potential shortfall might be". It is possible that this is a fair assessment of expected power utility maximization of wealth at a fixed horizon date T, subject to a "Value-At-Risk" (VaR) constraint that fixes a low probability for the event that terminal wealth could fall below a fixed floor. This problem was intensively studied by Basak and Shapiro (2001, p. 385), who concluded that "The shortcomings...stem from the fact that the VaR agent is concerned with controlling the probability of a loss rather than its magnitude". They proposed replacing the VaR constraint with an ad hoc expected loss constraint, resulting in fewer shortcomings. The investor's target growth rate serves a similar function in the horizon-free, unconstrained criterion (22).

3. Non-parametric implementation

In the IID case, there is a simple, distribution-free way to estimate $D_p(\log r)$ for a portfolio p. Following the comparative portfolio study of Kroll et al. (1984), we replace the expectation operator in (15) by an historical time average operator, substitute into (13), and numerically maximize that.[9] This estimator eliminates the need for prior

[8] While Browne considers a stochastic benchmark, the constant growth benchmark here can be modified to consider an arbitrary stochastic benchmark, at the cost of fewer concrete expository results.

[9] It is important to remember that the log moment generating function of the log return distribution necessarily has to exist near $\theta_{\max}$ in order for this technique to work here.

knowledge of the log return distribution's functional form and parameters. Specifically, let $R_p(t) = \sum_{j=0}^{n} p_j R_j(t)$ denote the historical return at time t of a portfolio comprised of $n+1$ assets with respective returns $R_j(t)$, with constantly rebalanced portfolio weights $\sum_j p_j = 1$. The estimator is

$$\hat{D_p}(\log r) = \max_{\theta} \theta \log r - \log \left[\frac{1}{T} \sum_{t=1}^{T} \left(\sum_{j=0}^{n} p_j R_j(t) \right)^{\theta} \right] \tag{24}$$

and the optimal portfolio weights are estimated to be

$$\hat{p} = \arg \max_{p_1,\ldots,p_n} \max_{\theta} \theta \log r - \log \left[\frac{1}{T} \sum_{t=1}^{T} \left(\sum_{j=1}^{n} p_j R_j(t) + \left(1 - \sum_{j=1}^{n} p_j \right) R_0(t) \right)^{\theta} \right]. \tag{25}$$

The maximum expected log portfolio was similarly estimated, by numerically finding the weights that maximize the time average of $\log R_p(t)$.

Let us now contrast the estimated decay rate maximizing portfolio (25) to both the expected log and Sharpe ratio maximizing, constantly rebalanced portfolios formed from Fama and French's 10 domestic industry, value-weighted assets,[10] whose annual returns run from 1927 through 2000. The sample cross-correlations of the 10 industries' gross returns range from 0.32 to 0.86, suggesting that diversified portfolios of them will provide significant investor benefits. The sample covariance matrix is invertible, permitting estimation of the Sharpe ratio maximizing "tangency" portfolio, by multiplying this inverse by the vector of sample mean excess returns over a riskless rate, and then normalizing the result. We assume that it was possible to costlessly store money between 1927–2000, with no positive *constant* nominal rate riskless asset available.[11] Hence we assume a zero constant riskless rate when computing the Sharpe ratio maximizing tangency portfolio of the 10 industry assets.

The results are seen in Table 2.

The performance statistics in Table 2 show that the Sharpe ratio maximizing portfolio has almost no skewness. But the decay rate maximizing portfolios all have a skewness of about 1, as does the expected log maximizing portfolio. This reflects the skewness preference inherent in the generalized expected power utilities with degrees of risk aversion greater than (in the log case, equal to) one.[12] In fact, these investors prefer all odd order moments and are averse to all even order moments. To see this, note that (15) is the cumulant generating function for the (assumed) IID log portfolio return

[10] The data are currently available for download from a website maintained by Kenneth French at MIT.

[11] Treasury Bills are not a *constant* rate riskless asset, like the one used to form portfolios in Section 2.1. A fixed percentage of wealth invested in Treasury Bills is just like any other risky asset.

[12] See Kraus and Litzenberger (1976) and Harvey and Siddique (2000) for evidence that investors prefer skewness.

Table 2
Comparison of estimated Sharpe ratio, expected log, and decay rate maximizing portfolios from Fama-French 10 industry indices, 1927–2000

Industries	Asset moments			Portfolio weights				
	μ	σ	Skewness	Max Sharpe	$\log r$ 5%	$\log r$ 10%	$\log r$ 15%	Max Log
NoDur	0.130	0.198	−0.12	0.80	0.92	1.0	1.11	1.15
Durbl	0.166	0.328	0.86	−0.01	0.27	0.52	0.97	1.22
Oil	0.137	0.220	0.01	0.75	0.77	0.96	1.24	1.36
Chems	0.146	0.225	0.63	0.14	0.35	0.52	0.89	1.15
Manuf	0.136	0.254	0.21	0.03	−0.10	0	0.11	0.20
Telcm	0.123	0.200	0.07	0.35	0.48	0.38	0.30	0.28
Utils	0.118	0.225	0.25	0.05	−0.20	−0.34	−0.61	−0.76
Shops	0.141	0.256	−0.25	−0.13	−0.44	−0.60	−0.96	−1.2
Money	0.142	0.245	−0.43	−0.23	−0.20	0.07	0.48	0.70
Other	0.106	0.242	−0.04	−0.76	−0.86	−1.5	−2.52	−3.09
Performance statistics								
Mean				0.148	0.162	0.195	0.248	0.278
Std. dev.				0.153	0.181	0.240	0.368	0.446
Skewness				−0.02	1.05	1.07	1.06	1.05
Decay rate $D_p(\log r)$					0.18	0.04	0.004	0
Risk aversion $1-\theta_{\max}(p)$					5.3	2.5	1.3	1

distribution. Substituting it into (13) and evaluating it at $\theta_{\max}(p)$ yields the following cumulant expansion:

$$D_p(\log r) = (\log r - \mathrm{E}[\log R_p])\theta_{\max}(p) - \frac{Var[\log R_p]}{2}\theta_{\max}(p)^2 - \sum_{i=3}^{\infty}\frac{\kappa_i}{i!}\theta_{\max}(p)^i, \tag{26}$$

which uses the facts that $\mathrm{E}[\log R_p]$ is the first cumulant of the log return distribution and that $Var[\log R_p]$ is its second cumulant, while κ_i denotes its ith order cumulant. Because $\theta_{\max}(p) < 0$, we see that the decay rate increases in odd-order cumulants and decreases in even-order cumulants. With normally distributed log returns, all the cumulants in the infinite sum are zero. But with non-normally distributed returns, increased skewness will increase the decay rate (due to κ_3). The relative weighting of the mean, variance and skewness in (26) is determined by their sizes, the sizes of the higher order cumulants, the target growth rate $\log r$, and the value of $\theta_{\max}(p) < 0$ associated with $\log r$.

The top panel of Table 2 contains the 10 industry weights in each portfolio. As is typical of estimated Sharpe ratio maximizing portfolios with more than a few assets, it is heavily long in just three industries (Non-durables, Oil, and Telecommunications). The decay rate maximizing portfolio for the target growth rate $\log r = 0.10$ is also heavily invested in these industries, but in addition it has considerable long positions in the two most positively skewed industries (Durables and Chemicals). The Sharpe ratio maximizing portfolio is heavily short in one industry (Other). The decay rate maximizing portfolios are heavily short in both this industry and as well as two others (Shops and Utilities). The differences between Sharpe ratio and decay rate maximizing portfolios are due to the presence of the target growth rate in decay rate maximization, its use of log gross returns rather than net returns when calculating portfolio means and variances, and the presence of higher order moments. It is difficult to assess the impact of higher order moments on the differences in portfolio weights. Bekaert et al. (1998, p. 113) were able to produce only a two percentage point difference in an asset weight, when simulating the effects of its return's skewness over the range -1 to 2.0, on the portfolio chosen by an expected power utility maximizing agent whose degree of risk aversion was close to 10. This suggests that the use of a target growth rate, and the use of log gross returns rather than arithmetic net returns, account for most of the differences between the decay rate and Sharpe ratio maximizing portfolios' weights.

The convergence of decay rate maximizing portfolios to the expected log maximizing portfolio is seen when reading across the last four columns of Table 2. The last two rows in the bottom panel of Table 2 show the relationship between the target growth rates, their respective efficient portfolios' maximum decay rates, and their respective endogenous degrees of risk aversion. Despite the fact that $\theta_{\max}(p)$ is determined by maximization in (25), we see that the degree of risk aversion $1-\theta_{\max}(p)$ is not unusually large in any of the decay rate maximizing portfolios tabled,[13] and converges toward 1 as $\log r \to \max_p \mathrm{E}[\log R_p]$. An alternative interpretation of this is enabled by computing the first order condition for $\theta_{\max}(p)$ in the IID case. To do so, substitute (15) into (13) and differentiate to find

$$\mathrm{E}\left[\log R_p \frac{\mathrm{d}Q}{\mathrm{d}P}\right] = \log r, \tag{27}$$

where the *Esscher* transformed probability density

$$\frac{\mathrm{d}Q}{\mathrm{d}P} = \frac{R_p^{\theta_{\max}(p)}}{\mathrm{E}[R_p^{\theta_{\max}(p)}]} \tag{28}$$

is used to compute the expected log return (i.e. growth rate) in (27).[14] Furthermore, a result known as Kullback's Lemma (1990) shows that the Esscher transformed density (28) is the solution to the following constrained minimization of relative entropy, whose minimized value is the decay rate, i.e.

$$D_p(\log r) = \min \mathrm{E}\left[\frac{\mathrm{d}Q}{\mathrm{d}P} \log \frac{\mathrm{d}Q}{\mathrm{d}P}\right] \text{ s.t. (27).} \tag{29}$$

From (27), an efficient portfolio has the highest decay rate among those with a fixed *transformed* expected growth rate equal to $\log r$. As $\log r \to \max_p \mathrm{E}[\log R_p]$,

[13] Of course, it can get unusually large when the target growth rate is unusually low, i.e. when the investor is unusually conservative.

[14] See Gerber and Shiu (1994) for option pricing formula derivations that use the Esscher transform to calculate the risk-neutral density required for option pricing.

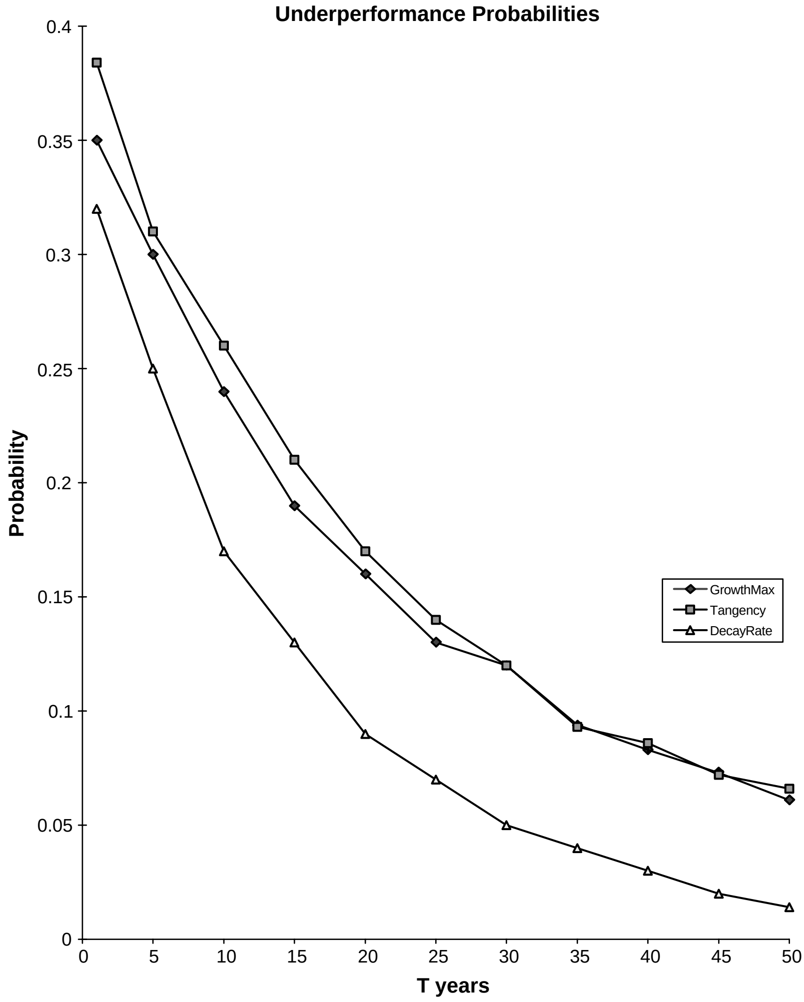

Fig. 4. Bootstrap estimated underperformance probabilities for portfolios in Table 2, when $\log r = 10\%$.

$\theta_{\max}(p) \to 0$, density (28) concentrates at unity and the minimal relative entropy in (29) approaches zero, i.e. the transformed probabilities approach the actual probabilities. As a result, the transformed expected log return in (27) approaches the actual expected log return, so constraint (27) collapses the portfolio constraint set onto the log optimal portfolio.

In order to determine if a decay rate maximizing portfolio in Table 2 will have lower underperformance probabilities than the Sharpe ratio and expected log maximizing portfolios do, the probabilities were estimated by resampling the portfolios' log returns 5000 times for each investment horizon length T, and then tabulating the empirical frequency of underperformance for each T. The results for the target decay rate

$\log r = 10\%$ are graphed in Fig. 4. Fig. 4 shows that the estimated decay rate maximizing portfolio in Table 2 had lower underperformance probabilities for all values of T.

3.1. More general estimators

The empirical estimates above were made under the assumption of IID returns. There is little evidence of serially correlated log returns in many equity portfolios, and what evidence there is finds low serial correlation. Hence there is little benefit in using an efficient estimator for the covariance stationary Gaussian rate function (20), e.g. using a Newey–West estimator of its denominator. But the presence of significant GARCH (perhaps with multiple components) effects (see Bollerslev, 1986) in log returns, as described in Tauchen (2001, p. 58), motivates the need for additional research into efficient estimation of (12) and (13) under specific parametric process assumptions. Alternatively, it may be possible to find an efficient nonparametric estimator for (12) and (13) by utilizing the smoothing technique exposited in Kitamura and Stutzer (1997, 2002) to estimate the expectation in (12).

4. Conclusions and future directions

A simple large deviations result was used to show that an investor desiring to maximize the probability of realizing invested wealth that grows faster than a target growth rate should choose a portfolio that makes the complimentary probability, i.e. of wealth growing no faster than the target rate, decay to zero at the maximum possible rate. A simple result in large deviations theory was used to show that this *decay rate maximization* criterion is equivalent to maximizing an expected power utility of the ratio of invested wealth to a "benchmark" wealth accruing at the target growth rate. The risk aversion parameter that determines the required power utility, and the investor's degree of risk aversion, is also determined by maximization and is hence endogenously dependent on the investment opportunity set. Yet it was not seen to be unusually large in the applications developed here.

The highest feasible target growth rate of wealth is that attained by the portfolio maximizing the expected log utility, i.e. that with the maximum expected growth rate of wealth. Investors with lower target growth rates choose decay rate maximizing portfolios that are more conservative, corresponding to degrees of risk aversion that exceed 1. As the target growth rate falls, it is easier to exceed it, so the decay rate of the probability of underperforming it goes up. The relationship between possible target growth rates and their corresponding maximal decay rates form an efficiency frontier that replaces the familiar mean-variance frontier. An investor's specific target growth rate determines the specific decay rate maximizing portfolio chosen by her. A decay rate maximizing investor does not choose a portfolio attaining an expected growth rate of wealth equal to her target growth rate (instead it is higher than her target). But there is an *Esscher* transformation of probabilities, under which the *transformed* expected growth rate of wealth *is* the target growth rate.

Researchers choosing to work in this area may select from several interesting topics. First, it is easy to generalize the analysis to incorporate a stochastic benchmark. This would be helpful in modelling an investor who wants to rank the probabilities that a group of similarly styled mutual funds will outperform their common style benchmark. Second, one could calculate the theoretical decay rate function using a multivariate GARCH model for the asset return processes, and then estimate the resulting function. Third, one could extend the decay rate maximizing investment problem to the joint consumption/portfolio choice problem, enabling the derivation of consumption-based asset pricing model with a decay rate maximizing representative agent. If it is possible to construct a model like this, the representative agent's degree of risk aversion will depend on the investment opportunity set—an effect heretofore unconsidered in the equity premium puzzle.

Acknowledgements

Thanks are extended to Eric Jacquier, the editors, and other participants at the Duke University Conference on Risk Neutral and Objective Probability Measures, to seminar participants at NYU, Tulane University, University of Illinois-Chicago, Chicago Loyola University, Georgia State University, University of Alberta, Bachelier Finance Conference, Eurandom Institute, Morningstar, Inc., and Goldman Sachs Asset Management, and to Edward O. Thorp, David Bates, Ashish Tiwari, Georgios Skoulakis, Paul Kaplan, and John Cochrane for their timely and useful comments on the analysis.

References

Algoet, P., Cover, T.M., 1988. Asymptotic optimality and asymptotic equipartition property of log-optimal investment. Annals of Probability 16, 876–898.

Basak, S., Shapiro, A., 2001. Value-at-risk based risk management: optimal policies and asset prices. Review of Financial Studies 14, 371–405.

Bekaert, G., Erb, C., Harvey, C., Viskanta, T., 1998. Distributional characteristics of emerging market returns and asset allocation. Journal of Portfolio Management 24, 102–116.

Bielecki, T.R., Pliska, S.R., Sherris, M., 2000. Risk sensitive asset allocation. Journal of Economic Dynamics and Control 24, 1145–1177.

Bodie, Z., 1995. On the risk of stocks in the long run. Financial Analysts Journal 51, 18–22.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Browne, S., 1995. Optimal investment policies for a firm with a random risk process: exponential utility and minimizing the probability of ruin. Mathematics of Operations Research 20, 937–958.

Browne, S., 1999a. Beating a moving target: optimal portfolio strategies for outperforming a stochastic benchmark. Finance and Stochastics 3, 275–294.

Browne, S., 1999b. The risk and rewards of minimizing shortfall probability. Journal of Portfolio Management 25, 76–85.

Bucklew, J.A., 1990. Large Deviation Techniques in Decision, Simulation, and Estimation. Wiley, New York.

Cover, T.M., Thomas, J.A., 1991. Elements of Information Theory. Wiley, New York.

Detemple, J., Zapatero, F., 1991. Asset prices in an exchange economy with habit formation. Econometrica 59, 1633–1657.

Fama, E., Miller, M., 1972. The Theory of Finance. Holt, Rhinehart and Whinston, New York.

Gerber, H., Shiu, E., 1994. Option pricing by Esscher Transforms. Transactions of the Society of Actuaries 46, 99–140.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

Hakansson, N.H., Ziemba, W.T., 1995. Capital growth theory. In: Jarrow, R.A., Maksimovic, V., Ziemba, W.T. (Eds.), Handbooks in Operations Research and Management Science: Finance, Vol. 9. North-Holland, Amsterdam.

Harvey, C., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 40, 1263–1293.

Hull, J., 1993. Options, Futures, and Other Derivative Securties. Prentice-Hall, Englewood Cliffs, NJ.

Kitamura, Y., Stutzer, M., 1997. An information-theoretic alternative to generalized method of moments estimation. Econometrica 65, 861–874.

Kitamura, Y., Stutzer, M., 2002. Connections between entropic and linear projections in asset pricing estimation. Journal of Econometrics 107, 159–174.

Kocherlakota, N.R., 1996. The equity premium: Its still a puzzle. Journal of Economic Literature 34, 42–71.

Kraus, A., Litzenberger, R.H., 1976. Skewness preference and the valuation of risk assets. Journal of Finance 31, 1085–1100.

Kroll, Y., Levy, H., Markowitz, H., 1984. Mean-variance versus direct utility maximization. Journal of Finance 39, 47–61.

MacLean, L.C., Ziemba, W.T., Blazenko, G., 1992. Growth versus security in dynamic investment analysis. Management Science 38, 1562–1585.

Samuelson, P.A., 1969. Lifetime portfolio selection by dynamic programming. Review of Economics and Statistics 51, 239–246.

Samuelson, P.A., 1994. The long-term case for equities. Journal of Portfolio Management 21, 15–24.

Stutzer, M., 2000. A portfolio performance index. Financial Analysts Journal 56, 52–61.

Tauchen, G., 2001. Notes on financial economics. Journal of Econometrics 100, 57–64.

Thorp, E.O., 1975. Portfolio choice and the Kelly criterion. In: Ziemba, W.T., Vickson, R.G. (Eds.), Stochastic Optimization Models in Finance. Academic Press, New York.