SAN DIEGO STATE UNIVERSITY

Department of Mathematics and Statistics

On Optimal Investment in the Long Run: Rank Dependent Expected Utility as a "Bridge" between the Maximum-Expected-Log and Maximum- Expected-Utility Criteria

Vladimir Rotar

WORKING PAPER

December, 2002, the new version: December 24, 2003

Keywords and phrases: portfolio optimization, investment policy, the maximum expected utility, the maximum expected logarithm.

Supported in part by the Russian Foundation of Basic Research, grant # 00-01-00194.

ABSTRACT

This paper concerns the well known paradox of inconsistency of the maximum-expected-utility (MEU) and the maximum-expected-log (MEL) criteria in investment dynamic models. The goal of the paper is to consider this phenomenon at the level of premises, and to suggest a generalized criterion, namely the rank dependent expected utility (RDEU) approach which allows to "bridge the gap" between the MEU and MEL criteria. The preference order in the RDEU approach is preserved by the functional

$$U(F) = \int_0^\infty u(x) d\Psi(F(x)),$$

where F is a probability distribution, u is a utility function, and Ψ is a transforming or weighting function: the subject "transforms" the real distribution function $F(x)$ into another one, $\Psi(F(x))$, assigning different weights to different probabilities.

One of main goals of the paper is to establish conditions on the tail of Ψ, and on the utility function u, under which the optimal investment in the long run corresponds to the MEL policy.

1 Introduction and an example.

1.1 Background and motivations

1.1.1 The MEU and MEL criteria.

This paper considers optimal investment in time, and concerns the long-known fact that the maximum-expected-utility (MEU) and the maximum-expected-log (MEL) criteria prove to be inconsistent even for large time horizons. This is, in a certain sense, a paradox since both criteria have reasonable justifications based on assumptions which, though maybe are restrictive (as, say, the independence axiom), but are natural at least in the first approximation. Below we try to explain the inconsistency mentioned at the level of premises, and apply to this problem some relatively recent achievements of the modern utility theory. Specifically we make use of the rank dependent expected utility (RDEU) approach which, as will be seen, allows to "bridge the gap" between the MEU and MEL criteria, and in a sense to "reconcile" the results based on the two approaches.

The history of the question will be considered later in Section 1.1.4 after we state the problem.

Let the initial wealth of an investor $W_0 = 1$, and the wealth after T periods of time

$$W_T = W_{T\pi} = \prod_{t=1}^{T}(1+R_t(\pi)),$$

where $R_t(\pi)$ is the return in period t under an investment policy π. In the standard framework π is a portfolio-vector, and $R_t(\pi) = \pi \cdot \mathbf{R}_t$, where $\mathbf{R}_t$ is the vector of the returns of assets in the market. Below we do not need such a specification and may view π as a policy of a general nature.

Suppose that for each π the random returns $\{R_i(\pi)\}$ are independent and identically distributed (i.i.d.), and at each period t the investor chooses the same policy not depending on the previous history (a "buy-and-hold" policy). In the asymptotic analysis the last assumption does not, in essence, restrict generality; we discuss this issue in more detail in Section 1.1.3. Nevertheless strategies π may depend on the horizon T.

Let $m(\pi) := E\{\ln(1+R_1(\pi))\}$, and there exist a policy $\hat{\pi} = Arg\max_{\pi} m(\pi)$. If for a policy π and some $c > 0$ it is true that $m(\hat{\pi}) - m(\pi) \geq c$, then (see references in Section 1.1.4)

$$(W_{T\hat{\pi}}/W_{T\pi}) \to \infty, \text{ as } T \to \infty, \text{ almost surely.} \tag{1.1.1}$$

The property (1.1.1) itself may serve as a desirable, though not necessary, requirement to an asymptotically optimal portfolio. In view of (1.1.1) it might seem that in the long run any "reasonable" policy π should be close in a sense to $\hat{\pi}$. However, as is well known, it is not the case for the MEU criterion

$$E\{u(W_{T\pi})\} = \int_0^\infty u(x)dF_{T\pi}(x), \tag{1.1.2}$$

where $u(x)$ is the *utility function* of the investor, and $F_{T\pi}$ is the distribution function of the random variable (r.v.) $W_{T\pi}$. Say, if $u(x) = x^\alpha$, then

$$E\{u(W_{T\pi})\} = \left(E\{(1+R_1(\pi))^\alpha\}\right)^T,$$

and the maximum is attained under a policy π' which maximizes $E\{(1+R_1(\pi))^\alpha\}$. Clearly, π' is not close to $\hat{\pi}$ in general, and cannot be close asymptotically since π' does not depend on T at all. For more complicated u, the analysis is more complex, but the conclusion is the same; see, e.g., Merton and Samuelson (1974), Markowitz (1976), references therein and in Section 1.1.4. In Appendix 1 we give a couple of simple but illustrative examples including that with a bounded utility function.

There has been a great deal of discussion on which criterion - MEU or MEL - is more preferable or realistic; see again Section 1.1.4. In this paper we are not concerned about which approach is better, but rather what makes them different, and the answer to this question is simple. Any integral criterion of the type (1.1.2) takes into account possibilities of "very large" values occurring with "very small" probabilities, while the property (1.1.1) has to do only with probabilities, and in a certain sense eliminates events with negligible probabilities.

If one deals with fixed, not growing variables as, say, in a one-time fixed investment, the difference between the two criteria can be not significant. However, once we deal with growing variables (as $W_{T\pi}$ in dynamics; for example, in a long period investment into a retirement fund), the difference mentioned may be dramatic.

The question is whether it is possible, maintaining at least some features of the classical MEU criterion, to make the decision process more flexible with respect to large deviations. One of possible answers is in making use of the *Rank Dependent Expected Utility (RDEU)* approach.

1.1.2 The RDEU criterion.

On a space of probability distributions F on $[0,\infty)$ consider a preference ordering preserved by the functional

$$U(F) = \int_0^\infty u(x)d\Psi(F(x)), \tag{1.1.3}$$

where u is a *utility* function, and the function Ψ is assumed to be non-decreasing, $\Psi(0)=0$, $\Psi(1)=1$.

The "*transformation*" function Ψ reflects the attitude of the subject to different probabilities: the subject "transforms" the real distribution function $F(x)$ into another one, $\Psi(F(x))$, assigning different weights to different probabilities.

For binary gambles such an approach was suggested by Kahneman and Tversky (1979) in their Prospect Theory, the full model was considered in Quiggin (1982), though some earlier Quigin's papers contained some relevant ideas. For convenience of reading a brief survey of the RDEU theory and some historical remarks are given in Appendix 2. A rather full history of the RDEU approach and a rich bibliography may be found in monographs Wakker (1989), Quiggin (1993), and Luce (2000).

A simple example is $\Psi(p)=p^\beta$. If $\beta=1$, the subject perceives F as it is, and hence deals with the "usual" expected utility (1.1.2). If $\beta<1$, the investor overestimates the probability for the wealth to be less that a fixed value: the investor is "security-minded". In the case $\beta>1$, the investor underestimates the probability mentioned, being "potentially-minded" [see, e.g., Lopes (1987, 1990)].

A limiting example is a truncation: if for a fixed $q\in[0,1]$

$$\Psi(p) = \begin{cases} p & \text{if } p<1-q, \\ 1 & \text{if } p\geq 1-q, \end{cases}$$

then

$$U(F) = qu(\gamma_q(F)) + \int_0^{\gamma_q(F)} u(x)dF(x), \tag{1.1.4}$$

where $\gamma_q(F)$ is the $(1-q)$-quantile of F. In this case the investor does not distinguish values greater than $\gamma_q(F)$ (viewed as too large) and occurring with a probability of q (viewed as too small). One may consider it as the existence of a perception threshold. The functional (1.1.4) is not linear and should be distinguished from the naive criterion when truncation is carried out at a fixed, perhaps, big value not depending on F.

If F is a distribution of a r.v. taking only two values, say, a and $b > a$ with probabilities p and $1-p$, respectively, then

$$U(F) = u(a)\Psi(p) + u(b)[1-\Psi(p)], \tag{1.1.5}$$

and $\Psi(p)$ "transforms" the probability p.

Let an investor having, say, the utility function $u(x) = \sqrt{x}$, can choose from two future retirement plans: either the annual pension will be equal to $X = \$100,000$, or to $Y = \$50,000$ or $\$200,000$ with equal probabilities. (We do not consider here annuities in dynamics.) For the numbers above the expected utility criterion leads to a slight preference for the latter plan ($E\{u(X)\} \approx 316$ and $E\{u(Y)\} \approx 335$), which for a large number of real people would hardly reflect their preferences. (At least the author would choose X.) On the other hand, under the criterion (1.1.5), as is easy to calculate, the investor would prefer X if $\Psi(1/2) > 0.59 > 1/2$, which means that such an investor would slightly overestimate the probability of the unlucky event to get $\$50,000$. So, one can expect $\Psi(p)$ to be concave for large p's. Certainly, the above primitive example is given merely for illustration.

1.1.3 The goal of the paper.

In this paper we establish conditions on Ψ, under which the optimal policy converges to the MEL-policy as $T \to \infty$. Roughly these conditions require the tail $1-\Psi(p)$ as $p \to 0$, or/and $\Psi(p)$ as $p \to 0$, to vanish sufficiently fast. Though the conditions mentioned are non-necessary, as will be seen in Section 1.2, they are close in a certain sense to minimal. The asymptotic optimality of the MEL-policy is understood as follows.

Let $\{\pi_T\}$ denote a sequence of policies where the integer $T \to \infty$. Suppose that for such a sequence, and some $c > 0$

$$m(\hat{\pi}) - m(\pi_T) \geq c$$

at least for large T, that is, π_T is not optimal w.r.t. the MEL-criterion at least for large T. Then the conditions established in the next sections would imply that there exists T_0 such that for all $T > T_0$

$$U(F_{T\pi_T}) < U(F_{T\hat{\pi}}), \tag{1.1.6}$$

that is, π_T is not optimal w.r.t. the RDEU-criterion (1.1.3) too.

If the space of policies under consideration is endowed by a metric $\|\cdot\|$ - in the standard framework it may be Euclidean, and if the policy $\hat{\pi}$ is, in a certain sense, unique with respect to this metric (see Section 2.3 for detail), then for the policy $\tilde{\pi}_T$ maximizing $U(F_{T\pi})$, (1.1.6) would imply that $\|\tilde{\pi}_T - \hat{\pi}\| \to 0$, as $T \to \infty$.

Probably it makes sense to say that such a result *is not to be considered* a *justification of the RDEU criterion in portfolio optimization*: it should be an object of a separate project. The goal of this paper is just to show that once we accept the RDEU framework, then the optimal policies may be close, under certain conditions, to the MEL-policy. In my opinion, this fact itself is an argument in favor of the RDEU approach, but nevertheless the problem to what extent RDEU reflects *real* possible investors' preferences, is open.

The last but important remark concerns the fact that we deal here only with policies not depending on the previous history. This is certainly a restriction but not as serious as it might look: under the i.i.d. assumption and some rather mild additional conditions optimal strategies are asymptotically, for large time horizons, close to stationary strategies of the type mentioned. However, the rigorous proof of this fact (especially when we apply the RDEU criterion) would take long calculations and make the framework and the result much less explicit. For this reason, to make the exposition clearer we start with stationary buy-and-hold strategies from the very beginning*: in the asymptotic analysis* it does not, in essence, restrict generality.

In this connection it worthwhile to note that in practically all papers where the difference between the MEL and MEU criteria has been discussed (for example, in basic papers by Merton and Samuelson (1974), and Markowitz (1976)) the choice of policies under consideration was the same and, I believe, for the same reason. The goal of this paper is in the introduction and analysis of a new criterion. So, it makes sense probably, at least in the first stage, to do this in the framework of the same model as has been considered before.

1.1.4 Historical remarks.

The MEL criterion itself was considered, for example, in Markowitz (1959), Latane (1959) and Breiman (1961). For properties of the MEL-portfolio as applied to bounded utilities see also Goldman (1974). A rather general model was investigated later in Algoet and Cover (1988), see also references therein.

It is worth noting that the maximization of the expected logarithm appears also in the analysis of stochastic analogues of the von Neumann-Gale model. In particular, the only natural stochastic analogue of the von Neumann ray is the balanced path that maximizes the expected logarithm of the growth rate. See, e.g., Arnold, Evstigneev, Gundalach (1994), Evstigneev and Taksar (2001), and references therein.

Different aspects of the application of the MEU criterion to portfolio optimization were considered in a great many of papers; see, e.g., Samuelson (1969), Hakansson (1971), Merton (1973), Breeden (1979); these papers also contain substantial reference lists.

The comparison of the two criteria, and a deep sophisticated discussion may be found in Samuelson (1969, 1971), Goldman (1974), Merton and Samuelson (1974), Markowitz (1976), Ophir (1978, 1979), Latane (1979), Samuelson (1979), and also in Markowitz's remarks following Samuelson (1988).

One can find in the literature some remarks on the relevancy of large deviations to the inconsistency of the MEU and MEL criteria (see, e.g., Latane (1959), Ophir (1978, 1979); Samuelson (1979)), though all these remarks are implicit. To my knowledge, the only paper where the inconsistency of the MEL and MEU criteria has been explicitly connected with large deviations, is the unpublished working paper by Kim, Omberg and Russell (1993) [15].

In [15] the authors suggest to divide the whole space of elementary events in two groups: "non-extreme events" which correspond to "moderate" values of the wealth, and extreme events "to be remaining events farther out in the tails". The authors suggest to consider the expected utility only over the "non-extreme outcomes", that is, to truncate integration in (1.1.2) by a "large" number depending on T (the same for all F). Some examples in [15] show that it may lead to the MEL policy.

I do not know papers where the RDEU approach was applied to portfolio optimization, though in general the idea of using RDEU in Economics is not new [see, e.g., Simonsen & Sérgio (1991), Dow and Sérgio (1992), Epstein and Tan Wang (1994), Mukerji & Tallon (1998), Tallon (1998), which does not exhaust the whole possible references.

The rest of the paper is organized as follows. In Section 1.2 we consider a particular example with the power transforming function in the standard geometrical Brownian motion framework. This example shows to what extent we should narrow the class of Ψ's.

The rest results concern the discrete time model as more difficult for analysis. The reader will easily see that similar results are true for the continuous time model too. General results are given in Section 2. Section 3 concerns a truncation criterion; see also comments in the end of Section 2. Proofs are given in Section 4. In Appendix 1 we consider two particular examples of the inconsistency of the MEU and MEL criteria. Appendix 2 contains a survey concerning the RDEU approach.

1.2 An example with power transformation functions for a simple continuous time model.

Next we consider a simple example when Ψ is a power function. In the context of the modern utility theory this case is viewed as too simple to be "realistic" (see Appendix 2 for comments and references), but it can serve as a good preliminary illustration of *what one can expect* in the RDEU framework.

To make an example simpler, we consider here just a power utility function and the standard continuous time scheme with a risk free and one risky securities governed, respectively, by the equations

$$dB_t = rB_t dt, \quad \text{and} \quad dS_t = S_t(mdt + \sigma dZ_t),$$

where Z_t is a Wiener process.

Let W_t be the total wealth at time t, $W_0 = 1$, and θ be the share of the wealth invested into the risky security. We assume θ to be a constant perhaps depending on the time horizon T. As is well known, in this case

$$W_t = \exp\{\mu(\theta)t + \theta\sigma Z_t\}, \tag{1.2.1}$$

where $\mu(\theta) = r + (m-r)\theta - \theta^2\sigma^2/2$. From (1.2.1) one gets the well known optimal θ under the MEL-criterion: $\theta_{\text{MEL}} = (m-r)/\sigma^2$.

Next we apply the RDEU criterion. We will see that the result should depend on the type of the utility function. Let first $u(x) = x^a,\ 0 < \alpha < 1$.

In this case large values of W_t matter, so the main property of $\Psi(p)$ should concern its behavior for p close to one. Set $\Psi(p) = 1-(1-p)^\beta,\ \beta \geq 1$. Then for the distribution F_T of W_T

$$U(F_T) = -\int_{-\infty}^{\infty} \exp\left\{\alpha\left(\mu(\theta)T + \theta\sigma\sqrt{T}z\right)\right\} d[1-\Phi(z)]^\beta, \tag{1.2.2}$$

where Φ is the standard normal distribution function. Calculations show that the maximizer of (1.2.2) is

$$\theta = \theta(T) = \frac{m-r}{(1-\alpha/\beta)\sigma^2}\left(1 + (\beta-1)O\left(\frac{1}{T}\right)\right). \tag{1.2.3}$$

For $\beta = 1$ we naturally get the optimal policy in the Merton's MEU model: $\theta_{\mathrm{MEU}} = (m - r)\big/[(1-\alpha)\sigma^2]$ (see, e.g., Duffie [6], Merton [28]). For $\beta > 1$ and large T the value θ shifts to θ_{MEL}, and

$$\bar{\theta} := \lim_{T\to\infty} \theta(T) = \frac{m-r}{(1-\alpha/\beta)\sigma^2}.$$

The greater the value of β, the closer $\bar{\theta}$ is to θ_{MEL}, and farther from θ_{MEU}. In the limiting case $\beta \to \infty$ one has $\bar{\theta} \to \theta_{\mathrm{MEL}}$. [If $T \to \infty$, and $\beta \to \infty$ simultaneously, the picture is more complicated, depending on which characteristic grows faster.]

Consider now $u(x) = -x^{-\alpha}$, $\alpha > 0$. Unlike in the previous case, here small values of W_t matter. Hence now the asymptotics of $\Psi(p)$ for $p \to 0$ is important. Set $\Psi(p) = p^{\beta}$, $\beta \geq 1$. It is not difficult to calculate that in this case the maximizer of $U(F_T)$ is

$$\theta(T) = \frac{m-r}{(1+\alpha/\beta)\sigma^2}\left(1+(\beta-1)O\left(\frac{1}{T}\right)\right), \tag{1.2.4}$$

to which similar comments apply.

Thus, though the power transforming function causes a shift towards the MEL policy, for the optimal policy to converge to the MEL strategy, as $T \to \infty$, the tail of $\Psi(p)$ should vanish faster than a power function. It is lucky that this is the case that - for other reasons - attracted the attention of a number of researchers last years. We use one of these results in the next section.

2 A discrete time model: general results.

2.1 Conditions on Ψ.

Note first that the criterion (1.1.3) is often specified in the literature in terms of the function $w(p) = 1 - \Psi(1-p)$ which is referred to as a *weighting function*. If $u(0) = 0$ (which does not restricts generality if $u(x)$ is bounded from below),

$$U(F) = \int_0^{\infty} w(1-F(x))du(x).$$

In some calculations the last representation proves to be more convenient.

We make here use of the relatively recent but already well known result by Prelec (1998), who established axioms under which the weighting function admits the explicit representation

$$w(p) = w_{\beta\eta}(p) = \exp\left\{-[-\beta\ln(p)]^{\eta}\right\}, \tag{2.1.1}$$

where β, η are positive parameters. The corresponding transformation function $\Psi_{\beta\eta}(p) = 1 - w_{\beta\eta}(1-p)$.

Prelec's axioms and some properties of (2.1.1) are discussed in Appendix 2. Now we just note that if $\eta = 1$, then (2.1.1) directly leads to the power function p^{β}, while for $\eta > 1$ the function $w_{\beta\eta}$ is S-shaped, and $w_{\beta\eta}(p) \to 0$ as $p \to 0$ (and hence $\Psi_{\beta\eta}(p) \to 1$ as $p \to 1$) faster than any power function. In this case the subject is "strongly security-minded", and in a rather small degree takes into account possibilities of "lucky" large deviations.

Below we use Prelec's representation, however since we are concerned only with large deviations, we will *not need to specify the precise shape of* $\Psi(p)$ *but only its asymptotic behavior* as $p \to 1$, *and/or* $p \to 0$. More specifically, we impose in general

Condition (Ψ): *For some* $\beta > 0$ *and* $\eta > 1$

$$\Psi(p) \le w_{\beta\eta}(p) \qquad \text{for } p \le 1/2, \tag{2.1.2}$$

$$\Psi(p) \ge 1 - w_{\beta\eta}(1-p) \qquad \text{for } p > 1/2. \tag{2.1.3}$$

If, as a matter of fact, (2.1.2) and (2.1.3) hold separately for different parameters β and η, we can choose the "worst", that is, the smallest, β and η. Since taking a smaller positive β we still have a property (Ψ), we can choose β small enough for $w_{\beta\eta}(1/2) \ge 1/2$, which makes (2.1.2) and (2.1.3) consistent.

As will be seen, if the utility function u is bounded from below, (2.1.2) may be replaced by the weaker condition

$$\Psi(p) \le Mp^a \qquad \text{for } p \le 1/2, \tag{2.1.4}$$

and some positive M, a. Clearly, for (2.1.4) and (2.1.3) to be consistent, one should have $M2^{-a} \ge 1 - w_{\beta\eta}(1/2)$. Note also that, as is easy to check, Prelec's function $\Psi_{\beta\eta}$ itself satisfies (2.1.4).

If u is bounded from above, (2.1.3) may be replaced by the condition

$$\Psi(p) \ge 1 - M(1-p)^a \qquad \text{for } p > 1/2. \tag{2.1.5}$$

and some $M, a > 0$. The similar remarks on consistency of (2.1.5) and (2.1.2), and the function $\Psi_{\beta\eta}$ apply to this case too.

2.2 Conditions on u.

The goal here is to make these conditions rather weak. This leads to a bit complicated formulations, so it is worth emphasizing that conditions below are very mild in the context of the utility maximization, and exclude just "very bad" utility functions. In particular, these conditions hold for "classical" functions $u(x) = \frac{x^{1-\gamma}}{1-\gamma}, \quad \gamma \ne 1,$ or $u(x) = \ln x$, or more generally for functions

$$u(x) = \frac{h(x)x^{1-\gamma}}{1-\gamma}, \quad \gamma \ne 1, \tag{2.2.1}$$

where h is a slowly varying function (s.v.f.)[1], or for s.v.f.'s themselves from a very large class of those. More detailed comments are given below.

We could avoid some conditions, if we have restricted ourselves to regularly

varying functions but it would have narrowed the setup of the problem too much, and not only from a mathematical point of view: the behavior of human beings is not always regular.

It proves also to be convenient to consider the case $u(\infty) < \infty$ separately, which we do in Section 2.4. Nevertheless the scheme of this section includes some bounded from above functions too.

[1]A function h is s.v. as $x \to \infty$ if $[h(kx)/h(x)] \to 1$, as $x \to \infty$, for any $k > 0$. Say, $\ln x, (\ln x)^s$ for any s, $\ln\ln x$, etc., are s.v.f.'s. Certainly, if $h(x) \to const$, it is s.v. If h is a s.v.f., then $[h(x)/x^{\varepsilon}] \to 0$, as $x \to \infty$ for any $\varepsilon > 0$. Similarly one deifines a s.v.f. as $x \to 0$. For details, see, e.g., Seneta (1976) [43], also Ibragimov and Linnik (1971), Rotar (1998).

Clearly, we can consider only non-decreasing and taking at least two values functions u. Then without loss of generality we can assume that $u(x) > 0$ for sufficiently large x. It will make conditions simpler. Below, when saying that something is true for large x we mean, as usual, that it is true for all x greater than a fixed x_0.

Condition $(u1)$ *There exist non-negative constants s and C_1 such that*

(a) for all $y \geq 1$ and sufficiently large x

$$u(xy) \leq C_1 u(x) y^s; \tag{2.2.2}$$

(b) for all $y \leq 1$ and sufficiently small (that is, close to zero) $x > 0$

$$|u(xy)| \leq C_1 |u(x)| \big/ y^s\,. \tag{2.2.3}$$

If (2.2.2) and (2.2.3) hold separately with different C_1 and s, we can choose the "worst" from these parameters. Next we discuss (2.2.2) and (2.2.3).

1. If $u(\infty) < \infty$, (2.2.2) holds automatically. The same concerns (2.2.3) if $u(0) > -\infty$.

2. If $u(x) = h(x)x^\alpha$ where $\alpha \geq 0$, and h is s.v. as $x \to \infty$, then (2.2.2) holds. For (2.2.3) to hold it suffices that $u(x) = \bar{h}(x)x^{-\alpha}$ where $\alpha \geq 0$, and $\bar{h}$ is s.v. as $x \to 0$.

3. If $u(0) = 0$ and $u(x)$ is concave, then (2.2.2) is true for all x with $C = s = 1$, while (2.2.3) holds automatically.

4. It is easy to see that, if (2.2.2) holds, $u(x) \leq Cx^s$ for some C and large x, that is, $u(x)$ grows, as $x \to \infty$, not faster than a power function. For instance, the exponential utility function e^x does not satisfy (2.2.2). Similarly, if (2.2.3) holds, $u(x) \geq -Cx^{-s}$ for sufficiently small x.

Conditions (2.2.2) and (2.2.3) require u to grow not too fast and not too irregularly. The next condition requires u to grow not too slow.

Condition $(u2)$. *For any $d > 1$ and any $l > 0$ there exists a constant $C(d,l) > 0$ such that for large x*

$$\frac{u(x)}{u(x^d)} \leq 1 - \frac{C(d,l)}{x^l}. \tag{2.2.4}$$

This condition covers a large class of utility functions but also requires some remarks.

1. Clearly (2.2.4) is true if

$$\limsup_{x\to\infty} \left[u(x) \big/ u(x^d) \right] < 1 \tag{2.2.5}$$

   As to (2.2.5), it obviously holds, say, for all $u(x) = h(x)x^\alpha$ where $\alpha > 0$ and h is s.v., or if, for example, $u(x) \sim \ln^\alpha(1+x)$, $\alpha > 0$, as $x \to \infty$. On the other hand, (2.2.5) excludes u growing "too slowly", say, as $\ln\ln(e+x)$. Since for such functions the theorem below, as a matter of fact, is true, we impose the weaker condition (2.2.4) which covers such functions.

2. Condition (2.2.4) holds for some bounded functions too, as, say, $u(x) = 1 - 1/\ln(e+x)$, but for example for $u(x) = 1 - 1/(1+x)^\alpha$, $\alpha > 0$, (2.2.4) is not true for all l. We consider such functions in Section 2.4 in a bit different terms.

3. Functions that grow as power ones at least for sequences of x's, but do not satisfy (2.2.4), certainly exist but look exotic. Let, say, $u(x_k) = \sqrt{x_k}$ for $x_k = \exp\{4^k\}$, $k = 1, 2, \dots$, and $u(x)$ is constant on $[x_k, x_{k+1})$. Then $\left[u(x_k)\,/u(x_k^2)\right] = 1$ for all k, that is, (2.2.4) is not true for $d = 2$.

4. It is not true however that (2.2.4) holds for any s.v.f. tending to infinity, but examples are rather exotic and we skip them here.

2.3 The main theorem.

First we need some conditions on r.v.'s $R_t(\pi)$.

Conditions (R**).**

(1) There exists a policy $\hat{\pi}$ *maximizing* $m(\pi)$, *and* $\hat{m} := m(\hat{\pi}) > 0$.

(2) Suppose
$$\sup_{\pi} \sigma(\pi) < \infty, \tag{2.3.1}$$
where $\sigma^2(\pi) := Var\{\ln(1 + R_1(\pi))\}$.

(3) *For a positive* c_0, *and all* π
$$|\ln(1 + R_1(\pi)) - m(\pi)| \le c_0 \sigma(\pi), \tag{2.3.2}$$
that is, the normalized r.v. $[\ln(1 + R_1(\pi)) - m(\pi)]/\sigma(\pi)$ is bounded uniformly in π.

The third requirement is imposed to avoid complicated conditions on the tails of the distributions of $R_t(\pi)$. It means, in particular, that the r.v. $\ln(1 + R_1(\pi))$ is bounded, and it excludes policies π for which the variance $\sigma^2(\pi)$ is "very small" while the r.v. $\ln(1 + R_1(\pi))$ itself may take "not small" values with "very small" probabilities.

Certainly it excludes, say, log-normally distributed r.v.'s , but for the discrete time model and from a "realistic" point of view such a restriction is not too strong.

Next we consider the first result. When comparing the policy $\hat{\pi}$ with a policy π, we presuppose that $\pi = \pi_T$, that is, perhaps depends on T. This is the case, for example, if we consider a policy maximizing $U(F_{T\pi})$. When considering a sequence of policies $\{\pi_T\}$ we assume that the integer parameter $T \to \infty$ but perhaps takes not all sequential natural values, that is, $\{\pi_T\} = \{\pi_{T_1}, \pi_{T_2}, \dots\}$ where $\{T_k\}$ is an increasing sequence of integers, and $T_k \to \infty$, as $k \to \infty$.

Theorem 1 *Assume that conditions (*Ψ*), (u1), (u2) and (R) hold. Suppose that for a sequence* $\{\pi_T\}$ *and a fixed* $c > 0$ *it is true that* $m(\hat{\pi}) - m(\pi_T) \ge c$ *at least for large* T. *Then there exists* T_0 *perhaps depending on the sequence* $\{\pi_T\}$ *and such that for all* $T > T_0$
$$U(F_{T\pi_T}) < U(F_{T\hat{\pi}}). \tag{2.3.3}$$

If $u(0) > -\infty$, *condition (2.2.3) in (u1) holds automatically, and instead of (2.1.2) in (*Ψ*) it suffices to assume (2.1.4) to be true for some positive M and a.*

Let now the space of all possible policies be endowed by a metric $\|\cdot\|$, and $\hat{\pi}$ is unique w.r.t. this metric in the following usual sense.

Condition $(\hat{\pi})$: *For each $\delta > 0$ there exists a positive ε depending only on δ and such that, if $\|\hat{\pi} - \pi\| \geq \delta$, then $m(\hat{\pi}) - m(\pi) \geq \varepsilon$.*

Clearly, Theorem 1 implies

Corollary 2 *Suppose that all conditions of Theorem 1 plus condition $(\hat{\pi})$ hold, and for each T there exists a policy $\tilde{\pi}_T$ maximizing $U(F_{T\pi})$. Then $\|\tilde{\pi}_T - \hat{\pi}\| \to 0$, as $T \to \infty$.*

2.4 The case of u bounded from above.

In this case we slightly change conditions. First without loss of generality we assume $u(\infty) = 0$. Next, instead of both conditions $(u1)$ we consider

Condition $(u1+)$**.** *There exist non-negative s and C_1 such that (2.2.3) hold for all x and all $y \leq 1$.*

The sense of *(2.2.3)* for small x's has been already discussed. Regarding large x's note that

1. Condition (2.2.3) is true for all x and $y \leq 1$ for strictly negative $u(x) = h(x)/x^{\alpha}$, where $\alpha \geq 0$, and h is s.v.f. for $x \to 0$, and $x \to \infty$, as well.
2. As is easy to see, (2.2.3) implies $u(x) \geq -C/x^s$ for a constant C $(= u(1)/C_1)$, and $x > 1$, that is, $u(x) \to 0$, as $x \to \infty$, not faster than a power function.

The last remark means that the theorem below does not cover bounded functions $u(x)$ converging to $u(\infty)$ too fast; say, the exponential utility function $-e^{-\alpha x}, \alpha > 0$. An analysis of proofs below allows to conjecture that this reflects the essence of the matter, and the corresponding result cannot be proved in the rank-dependent-utility framework.

Next we formulate the following counterpart of (2.2.4).

Condition $(u2+)$. *For any $d > 1$ and any l there exists a constant $C(d,l)$ such that for large x*

$$\left|\frac{u(x^d)}{u(x)}\right| \leq 1 - \frac{C(d,l)}{x^l}. \tag{2.4.1}$$

Note that in (2.4.1) we deal with $u(x) \to 0$, as $x \to \infty$. The remarks similar to those following (2.2.4) may apply to this case too.

Theorem 3 *Assume that conditions $(u1+)$, $(u2+)$ and (R) hold. Regarding the function Ψ suppose that (2.1.2) for some $\beta > 0$ and $\eta > 1$, and (2.1.5) for some positive M and a, are true. Let $m(\hat{\pi}) - m(\pi_T) \geq c$ for a sequence $\{\pi_T\}$ and a fixed $c > 0$ at least for large T. Then there exists T_0 such that $U(F_{T\pi_T}) < U(F_{T\hat{\pi}})$ for all $T > T_0$. If in addition condition $(\hat{\pi})$ holds, and there exists a policy $\tilde{\pi}_T$ maximizing $U(F_{T\pi})$, then $\|\tilde{\pi}_T - \hat{\pi}\| \to 0$, as $T \to \infty$.*

The above general results should be viewed rather as qualitative. In concrete situations, even if we accept the hypothesis that the investor assigns different weights to different probabilities, it could be still a problem to figure out the particular weighting function of the investor. One can hope that in future some models for a “typical” investor will be elaborated - and Prelec’s representation is a step in this direction, but for now this topic is not so developed. In the light of this, the simple but explicit truncation criterion could occur to be useful. In the next section we consider it in detail. It will allow to remove or essentially weaken some conditions, and to consider a quantitative estimate for the time T_0.

3 On the truncation criterion

To include into consideration a larger class of utility function we consider truncation from both sides. More specifically, we fix $q \in [0,1]$, and consider the criterion

$$U(F) = q[u(\gamma_{q+}(F)) + u(\gamma_{q-}(F))] + \int_{\gamma_{q-}(F)}^{\gamma_{q+}(F)} u(x)dF(x), \tag{3.1}$$

where $\gamma_{q-}(F)$ and $\gamma_{q+}(F)$ are q- and $(1-q)$- quantiles of F, respectively. As was noted in Section 1.1.2 this is a limiting case of the rank dependent utility: the investor does not distinguish "too large" values occurring with a small probability of q, as well as "too small values" occurring with the same probability. Truncation in the areas of large and small values should be interpreted differently. If an investor does not distinguish too large values of gains, she exhibits a threshold of perception in the zone of "lucky" events, demonstrating a cautious behavior. On the other hand, if the investor does not distinguish too small values of the wealth, that is, too large values of losses, she has a threshold of perception in the area of ruin. It reflects the well known phenomenon when people exhibit different types of behavior depending on whether they deal with gains or losses (see, e.g., Luce (2000)).

By the sense of the problem, q is small, so we can assume $q \leq 1/\dot{4}$.

In view of (2.3.1)

$$W_{T\pi} = \exp\left\{m(\pi)T + \sigma(\pi)\sqrt{T} \cdot \xi_{T\pi}\right\}, \tag{3.2}$$

where

$$\xi_{T\pi} = \frac{\ln W_{T\pi} - m(\pi)T}{\sigma(\pi)\sqrt{T}} = \frac{\sum_{t=1}^{T}[\ln(1+R_t(\pi)) - m(\pi)]}{\sigma(\pi)\sqrt{T}}$$

is asymptotically normal for each π such that $\sigma(\pi) \neq 0$. If $\sigma(\pi) = 0$, we can define $\xi_{T\pi}$ as a standard normal r.v.; it will not cause a misunderstanding below.

For each pair of distributions F and G, set $\|F-G\|_\infty = \sup_x |F(x) - G(x)|$. Let, as before, $F_{T\pi}(x) = P(W_{T\pi} \leq x)$, $F^*_{T\pi}(x) = P(\xi_T \leq x)$ and $\Delta_T(\pi) = \|F^*_{T\pi} - \Phi\|_\infty$. Clearly, $\Delta_T(\pi) \to 0$, as $T \to \infty$, for each π.

Condition (UNA: *Uniform Normal Approximation):*

$$\sup_\pi \Delta_T(\pi) \to 0, \quad \text{as} \quad T \to \infty. \tag{3.3}$$

It is a rather weak condition. For example, since by the Berry-Esseen theorem (see, e.g., [9]) if $\sigma(\pi) \neq 0$, then

$$\Delta_T(\pi) \leq \frac{\beta(\pi)}{\sigma^3(\pi)\sqrt{T}}, \tag{3.4}$$

where $\beta(\pi) = E\{|\ln R_1(\pi) - m(\pi)|^3$, condition (3.3) holds, say, if the supremum of the Lyapunov ratios,

$$L = \sup_{\pi:\sigma(\pi)\neq 0} (\beta(\pi)/\sigma^3(\pi)) < \infty. \tag{3.5}$$

Theorem 4 *Let the policy $\hat{\pi}$ exist and $m(\hat{\pi}) > \dot{0}$. Let $m(\hat{\pi}) - m(\pi_T) \geq c$ for a sequence $\{\pi_T\}$ and some fixed $c > 0$. Then, in the case of the criterion (3.1), $U(F_{T\pi_T}) < U(F_{T\hat{\pi}})$ for $T \geq$ some T_0 if the following conditions hold: (2.3.1), UNA, and either (2.2.4) if $u(\infty) = \infty$, or (2.4.1) if $u(\infty) = 0$.*

If we deal just with (3.1) we can estimate the above threshold time T_0. To make the expressions below simpler we consider, instead of (3.3), condition (3.5) - the latter implies the former, and instead of (2.2.4) or (2.4.1) - a simpler though slightly stronger

Condition $(u <)$. *There exist a positive k_u such that for $k > k_u$*

$$\tilde{u}(k) := \sup_{x\geq 1} \frac{u(x)}{u(kx)} < \frac{1}{4}, \quad \text{if } u(\infty) = \infty, \quad \text{and} \quad \tilde{u}(k) := \sup_{x\geq 1} \left| \frac{u(kx)}{u(x)} \right| < \frac{1}{4}, \quad \text{if } u(\infty) = 0. \tag{3.6}$$

We write $1/4$ just for simplicity; any number less than $1/2$ can be chosen. Clearly, any function (2.2.1) satisfies (3.6).

Let $\varphi(x)$ be the standard normal density, and

$$C_q = 1/\varphi(\gamma_q(\Phi) + 1) = \sqrt{2\pi} e^{(\gamma_q(\Phi)+1)^2/2}. \tag{3.7}$$

Theorem 5 *Consider the criterion (3.1), and assume conditions (σ), (3.5), and ($u <$) to hold. Then $U(F_{T\pi}) \leq U(F_{T\hat{\pi}})$ for all*

$$T > T_0 = \Gamma_1 + \frac{\Gamma_2}{m(\hat{\pi})} + \frac{\Gamma_3}{m(\hat{\pi}) - m(\pi)} + \frac{\Gamma_4}{(m(\hat{\pi}) - m(\pi))^2}, \tag{3.8}$$

where

$$\Gamma_1 = 4L^2(C_q^2 + 4),\ \Gamma_2 = \ln k_u,\ \Gamma_3 = 2\ln k_u,\ \Gamma_4 = 4(\gamma_q(\Phi) + 1)^2 \bar{\sigma}^2,\ \textit{and}\ \bar{\sigma} = \sup_{\pi} \sigma(\pi).$$

The above values of Γ's are just rough estimates. More accurate estimation may lead to much more precise though more cumbersome expressions for Γ's.

4 Proofs

4.1 Proof of Theorem 1

Below we will consider only sequences of policies $\{\pi_T\}$ for which $\inf_T m(\pi_T) > 0$ and $\inf_T \sigma(\pi_T) > 0$; otherwise calculations are much simpler.

We make use of the exponential bounds for large deviations (see, e.g., [9], [18], [37]). For the r.v. $\xi_{T\pi}$ under consideration it may be formulated as follows: for all π

$$P(\xi_{T\pi} > x) \leq \exp\left\{ -\frac{x^2}{2} \left(1 - \frac{xc_0}{2\sqrt{T}} \right) \right\} \quad \text{for} \quad 0 \leq x \leq \sqrt{T}/c_0,$$

$$P(\xi_{T\pi} > x) \leq \exp\left\{-\frac{x\sqrt{T}}{4c_0}\right\} \quad \text{for} \quad x \geq \sqrt{T}/c_0,$$

where c_0 is the constant from (2.3.2). The same bounds are true for $P(\xi_T < -x),\ x > 0$.

We simplify this as

$$P(\xi_{T\pi} > x) \leq g_T(x), \quad P(\xi_{T\pi} < -x) \leq g_T(x), \tag{4.1.1}$$

where x is arbitrary, and

$$g_T(x) = \begin{cases} 1 & \text{for } x < 0, \\ \exp\{-x^2/4\} & \text{for } 0 \leq x \leq \sqrt{T}/c_0, \\ \exp\{-x\sqrt{T}/4c_0\} & \text{for } x \geq \sqrt{T}/c_0. \end{cases} \tag{4.1.2}$$

Set, as before, $F_{T\pi}(x) = P(W_{T\pi} \leq x)$, and $F^*_{T\pi}(x) = P(\xi_{T\pi} \leq x)$. If it cannot cause a misunderstanding, we omit sometimes the index T in π_T, and the index π in $\xi_{T\pi}$, $F_{T\pi}$ and $F^*_{T\pi}$, and write m and σ instead of $m(\pi)$ and $\sigma(\pi)$.

As usual, saying below that something, for example, an inequality, is true for large T we mean that it is true for all T greater or equal than some fixed T_1.Once it has been said, for the rest of the proof we consider only $T \geq T_1$. It means, in particular, that if, say, another inequality is true for $T \geq$ some T_2, than we consider $T \geq \max\{T_1, T_2\}$. We will not repeat it each time. The numbers T_1, T_2, etc., mentioned depend perhaps on parameters of the problem: c_0, C_1, M, a, etc.

Let c be a continuity point of $\Psi(F^*_T(x))$. Then,

$$\begin{aligned} U(F_T) &= \int_0^\infty u(x)d\Psi(F_T(x)) = \int_{-\infty}^\infty u(e^{mT}e^{\sigma\sqrt{T}y})d\Psi(F^*_T(y)) \\ &= \int_{-\infty}^{c-0} u(e^{mT}e^{\sigma\sqrt{T}y})d\Psi(F^*_{T\pi}(y)) - \int_c^\infty u(e^{mT}e^{\sigma\sqrt{T}y})d[1-\Psi(F^*_T(y))] \\ &= \int_{-\infty}^{c} u(e^{mT}e^{\sigma\sqrt{T}y})d\Psi(F^*_{T\pi}(y)) - \int_c^\infty u(e^{mT}e^{\sigma\sqrt{T}y})d[1-\Psi(F^*_T(y))] \\ &= u(e^{mT}e^{\sigma\sqrt{T}c}) + \int_c^\infty [1-\Psi(F^*_T(y))]du(e^{mT}e^{\sigma\sqrt{T}y}) - \int_{-\infty}^c \Psi(F^*_T(y))du(e^{mT}e^{\sigma\sqrt{T}y}). \end{aligned} \tag{4.1.3}$$

Hence, by (4.1.1) and since u is non-decreasing, for $c \geq 0$,

$$U(F_T) \leq u(e^{mT}e^{\sigma\sqrt{T}c}) + \int_c^\infty [1-\Psi(1-g_T(y))]du(e^{mT}e^{\sigma\sqrt{T}y}). \tag{4.1.4}$$

Let $\tilde{\Psi}(p)$ be any non-decreasing and continuous at zero function such that $\tilde{\Psi}(0) = 0, \tilde{\Psi}(1) = 1$, and $\tilde{\Psi}(p) \leq \Psi(p)$ for all p. Then (4.1.4) implies that

$$\begin{aligned} U(F_T) &\leq u(e^{mT}e^{\sigma\sqrt{T}c}) + \int_c^\infty [1-\tilde{\Psi}(1-g_T(y))]du(e^{mT}e^{\sigma\sqrt{T}y}) \\ &= u(e^{mT}e^{\sigma\sqrt{T}c})\tilde{\Psi}(1-g_T(c)) + \int_c^\infty u(e^{mT}e^{\sigma\sqrt{T}y})d\tilde{\Psi}(1-g_T(y)). \end{aligned}$$

Since c can be chosen arbitrary close to zero and $\tilde{\Psi}(p)$ is continuous at zero, eventually

$$U(F_T) \leq \int_0^\infty u(e^{mT}e^{\sigma\sqrt{T}y})d\tilde{\Psi}(1-g_T(y)). \tag{4.1.5}$$

Recall that in the case of Theorem 1 we assume that $u(x) > 0$ for sufficiently large x. Consequently, since we consider only the case $\inf_T m(\pi_T) > 0$, for sufficiently large T all values of $u(e^{mT}e^{\sigma\sqrt{T}y})$ in the r.-h.s. of (4.1.5) are positive for all y. Set

$$k_1 = k_1(T) = 2\beta^{-1/2}T^{1/2\eta}. \tag{4.1.6}$$

Since $\eta > 1$, there exists $T_1 = T_1(\beta,\eta)$ such that for $T \geq T_1$ one has $k_1(T) \leq \sqrt{T}/c_0$. So, for $T \geq \max\{T_1, (\beta \ln 2)^{\eta}\}$

$$g_T(k_1) = \exp\{-k_1^2/4\} = \exp\{-T^{1/\eta}/\beta\} \leq 1/2. \tag{4.1.7}$$

Then we can use (2.1.3) and set $\tilde{\Psi}(p) = 1 - w_{\beta\eta}(1-p)$ for $p \geq 1/2$, and $\tilde{\Psi}(p) = \min\{\Psi(p), 1 - w_{\beta\eta}(1-p)\}$ for $p < 1/2$. The $\tilde{\Psi}$ so chosen is continuous at zero. By (4.1.5)

$$U(F_T) \leq \int_0^{\infty} u(e^{mT}e^{\sigma\sqrt{T}y})d\tilde{\Psi}(1-g_T(y)) = \int_0^{k_1} + \int_{k_1}^{\infty} = J_1 + J_2. \tag{4.1.8}$$

For J_1 it suffices to write

$$J_1 \leq u(\exp\{mT + k_1\sigma\sqrt{T}\}) = u(\exp\{mT + 2\sigma\beta^{-1/2}T^{(1+\eta)/2\eta}\}), \tag{4.1.9}$$

where the r.-h.s. is positive for large T.

For large T we have also $1 - g_T(k_1(T)) > 1/2$, and hence $\tilde{\Psi}(1-g_T(y)) = 1 - w_{\beta\eta}(g_T(y))$ for $y \geq k_1(T)$. Next we use (2.2.2) which is true for $x >$ some x_1. Since $\exp\{mT + k_1(T)\sigma\sqrt{T}\} > x_1$ for large T, we can use (2.2.2) when estimating J_2. Consequently, for sufficiently large T

$$J_2 = \int_{k_1}^{\infty} u(e^{mT}e^{\sigma\sqrt{T}y})d[-w_{\beta\eta}(g_T(y))] \leq C_1 u(e^{mT})I, \tag{4.1.10}$$

where

$$I = I(T) = \int_{k_1(T)}^{\infty} \exp\{s\sigma\sqrt{T}y\}d[-w_{\beta\eta}(g_T(y))]. \tag{4.1.11}$$

As was told already, $k_1(T) \leq \sqrt{T}/c_0$ for large T, and hence we can write

$$I = \int_{k_1}^{\infty} = \int_{k_1}^{\sqrt{T}/c_0} + \int_{\sqrt{T}/c_0}^{\infty} = I_1 + I_2.$$

Furthermore,

$$\begin{aligned}
I_1 &= I_1(T) = \int_{k_1(T)}^{\sqrt{T}/c_0} \exp\{s\sigma\sqrt{T}y\}d[-w_{\beta\eta}(g_T(y))] \\
&= \int_{k_1(T)}^{\sqrt{T}/c_0} \exp\left\{s\sigma\sqrt{T}y\right\} d\left[-\exp\left\{-\left[\beta y^2/4\right]^{\eta}\right\}\right] = \int_{\sqrt{\beta}k_1(T)/2}^{\sqrt{\beta}\sqrt{T}/2c_0} \exp\left\{2\beta^{-1/2}s\sigma\sqrt{T}y\right\} d\left[-\exp\left\{-y^{2\eta}\right\}\right] \\
&\leq \int_{\sqrt{\beta}k_1(T)/2}^{\infty} \exp\left\{2\beta^{-1/2}s\sigma\sqrt{T}y\right\} d\left[-\exp\left\{-y^{2\eta}\right\}\right].
\end{aligned} \tag{4.1.12}$$

Thus, we need an estimate for integrals of the type

$$\int_R^\infty x^l \exp\{tx - x^{2\eta}\} dx \le \int_R^\infty x^l \exp\{tx - x^2 R^{2\eta-2}\} dx$$

$$= \frac{1}{R^{(\eta-1)(l+1)}} \int_{R^\eta}^\infty y^l \exp\{(t/R^{\eta-1})y - y^2\} dx \le \frac{C(l)}{R^{(\eta-1)(l+1)}} \int_{R^\eta}^\infty \exp\{(t/R^{\eta-1})y - y^2/2\} dx$$

$$= \frac{C(l)}{R^{(\eta-1)(l+1)}} \exp\left\{\frac{1}{2}\left(t/R^{\eta-1}\right)^2\right\} \int_{R^\eta}^\infty \exp\{-(y-(t/R^{\eta-1}))^2/2\} dx$$

$$\le \frac{C(l)}{R^{(\eta-1)(l+1)}} \exp\left\{\frac{1}{2}\left(t/R^{\eta-1}\right)^2\right\} [1 - \Phi(R^\eta - (t/R^{\eta-1}))].$$

where $C(\cdot)$ as usual, denotes a constant depending only on the argument in $(\cdot)$, and which may be *different in different formulas*. Applying the last inequality to (4.1.12), we have

$$\begin{aligned} I_1(T) \quad &\le \quad \frac{C(\eta)(2\beta^{-1/2})^{(\eta-1)2\eta}}{k_1^{(\eta-1)2\eta}(T)} \exp\left\{\frac{1}{2}\left((2/\sqrt{\beta})^\eta s\sigma\sqrt{T}/k_1^{\eta-1}\right)^2\right\} \\ &\quad \times \left[1 - \Phi\left((\sqrt{\beta}k_1/2)^\eta - (2/\sqrt{\beta})^\eta s\sigma\sqrt{T}/k_1^{\eta-1}\right)\right] \\ &\le \quad \frac{C(\eta)}{T^{(\eta-1)}} \exp\left\{2\beta^{-1}s^2\sigma^2 T^{1/\eta}\right\} \left[1 - \Phi\left(\sqrt{T} - 2\beta^{-1/2} s\sigma T^{1/2\eta}\right)\right], \end{aligned} \tag{4.1.13}$$

It suffices now to use that

$$\left(\frac{1}{x} - \frac{1}{x^3}\right)\varphi(x) \le 1 - \Phi(x) = \Phi(-x) \le \frac{1}{x}\varphi(x), \quad \text{for} \quad x > 0, \tag{4.1.14}$$

where $\varphi(x) = \Phi'(x)$ (see, e.g., [9], [18], [37]). Set $\bar{\sigma} = \sup_\pi \sigma(\pi) < \infty$ in view of (2.3.1). It is straightforward to derive from (4.1.13) and (4.1.14) that, if $\eta > 1$, then there exists $T_2 = T_2(\bar{\sigma}, C_1, \beta, \eta, s)$ such that for $T \ge T_2$

$$I_1(T) \le e^{-T/4}. \tag{4.1.15}$$

Furthermore,

$$\begin{aligned} I_2(T) \quad &= \quad \int_{\sqrt{T}/c_0}^\infty \exp\{s\sigma\sqrt{T}y\} d[-w_{\beta\eta}(\exp\{-y\sqrt{T}/4c_0\})] \\ &= \quad \int_{\sqrt{T}/c_0}^\infty \exp\left\{s\sigma\sqrt{T}y\right\} d\left[-\exp\left\{-\left[\frac{\beta y\sqrt{T}}{4c_0}\right]^\eta\right\}\right] \\ &= \quad \int_{\beta T/4c_0^2}^\infty \exp\left\{(4c_0\beta^{-1}s\sigma)z\right\} d\left[-\exp\left\{-z^\eta\right\}\right]. \end{aligned} \tag{4.1.16}$$

We see that there exists $T_3 = T_3(c_0, C_1, \beta, \eta, s)$ such that for all $T \ge T_3$

$$I_2 \le e^{-T}. \tag{4.1.17}$$

From (4.1.15) and (4.1.17) it follows that for large T

$$I(T) \leq e^{-T/4} + e^{-T} \leq 2e^{-T/4}, \tag{4.1.18}$$

and hence

$$J_2 \leq C_1 u(e^{mT})(e^{-T/4} + e^{-T}) \leq u(e^{mT})e^{-T/5} \tag{4.1.19}$$

for sufficiently large T.

Collecting (4.1.8), (4.1.9), and (4.1.19), we obtain that for sufficiently large T

$$\begin{aligned} U(F_T) &\leq u(\exp\{mT + 2\bar{\sigma}\beta^{-1/2}T^{(1+\eta)/2\eta}\}) + u(e^{mT})e^{-T/5} \\ &\leq u(\exp\{mT + 2\bar{\sigma}\beta^{-1/2}T^{(1+\eta)/2\eta}\})[1 + e^{-T/5}], \end{aligned} \tag{4.1.20}$$

where the r.-h.s. is positive.

We turn to lower bounds for $U(F_T)$. Let c be a continuity point of $\Psi(F_T^*(y))$, and $c \leq 0$. Then, by (4.1.3),

$$U(F_T) \geq u(e^{mT}e^{\sigma\sqrt{T}c}) - \int_{-\infty}^{c} \Psi(F_T^*(y))du(e^{mT}e^{\sigma\sqrt{T}y}) \geq u(e^{mT}e^{\sigma\sqrt{T}c}) - \int_{-\infty}^{c} \Psi(g_T(-y))du(e^{mT}e^{\sigma\sqrt{T}y}). \tag{4.1.21}$$

Let now $\tilde{\Psi}(p)$ be any non-decreasing and continuous at $p = 1$ function such that $\tilde{\Psi}(0) = 0, \tilde{\Psi}(1) = 1$, and $\tilde{\Psi}(p) \geq \Psi(p)$ for all p. Then it follows from (4.1.21) that

$$\begin{aligned} U(F_T) &\geq u(e^{mT}e^{\sigma\sqrt{T}c}) - \int_{-\infty}^{c} \tilde{\Psi}(g_T(-y))du(e^{mT}e^{\sigma\sqrt{T}y}) \\ &= u(e^{mT}e^{\sigma\sqrt{T}c})(1 - \tilde{\Psi}(g_T(-c))) + \int_{-\infty}^{c} u(e^{mT}e^{\sigma\sqrt{T}y})d\tilde{\Psi}(g_T(-y)) \\ &= u(e^{mT}e^{\sigma\sqrt{T}c})(1 - \tilde{\Psi}(g_T(-c))) + \int_{|c|}^{\infty} u(e^{mT}e^{-\sigma\sqrt{T}y})d[-\tilde{\Psi}(g_T(y))]. \end{aligned}$$

Since c can be chosen arbitrary close to zero and $\tilde{\Psi}(p)$ is continuous at one,

$$U(F_T) \geq \int_0^{\infty} u(e^{mT}e^{-\sigma\sqrt{T}y})d[-\tilde{\Psi}(g_T(y))] \tag{4.1.22}$$

Let now x_0 be the number such that $u(x) \leq 0$ for $x < x_0$, and $u(x) > 0$ for $x > x_0$. By condition $(u1)$ there exists $x_2 \leq x_0$ such that (2.2.3) holds for $x \leq x_2$. Let $y_0 = y_0(T) = m\sqrt{T}\sigma^{-1} - (\ln x_0)(\sigma\sqrt{T})^{-1}$, $y_2 = y_2(T) = m\sqrt{T}\sigma^{-1} - (\ln x_2)(\sigma\sqrt{T})^{-1}$. Clearly, $k_1(T) < y_0(T) \leq y_2(T)$ for large T.

Set $\tilde{\Psi}(p) = w_{\beta\eta}(p)$ for $p \leq 1/2$, and $\tilde{\Psi}(p) = \max\{w_{\beta\eta}(p), \Psi(p)\}$ otherwise. Note that $\tilde{\Psi}(p)$ is continuous at $p = 1$. Furthermore

$$U(F_T) \geq \int_0^{\infty} u(e^{mT}e^{-\sigma\sqrt{T}y})d[-\tilde{\Psi}(g_T(y))] = \int_0^{k_1} + \int_{k_1}^{y_0} + \int_{y_0}^{y_2} + \int_{y_2}^{\infty} = J_3 + J_4 + J_5 + J_6. \tag{4.1.23}$$

Considering T large enough for $g_T(k_1(T)) \leq 1/2$, we get that

$$
\begin{aligned}
J_3 &\geq u(\exp\{mT - k_1(T)\sigma\sqrt{T}\})[1-\tilde{\Psi}(g_T(k_1(T)))] = u(\exp\{mT-k_2(T)\sigma\sqrt{T}\})[1-w_{\beta\eta}(g_T(k_1(T)))] \\
&= u(\exp\{mT - 2\sigma\beta^{-1/2}T^{(1+\eta)/2\eta}\})[1-e^{-T}], \qquad (4.1.24)
\end{aligned}
$$

where the r.-h.s. is positive for large T.

It suffices to write $J_4 \geq 0$. The next integral

$$
J_5 \geq -|u(x_2)|\tilde{\Psi}(g_T(y_0)) \geq -|u(x_2)|w_{\beta\eta}(g_T(y_0)) \geq -|u(x_2)|w_{\beta\eta}(g_T(k_1)) = -|u(x_2)|e^{-T}. \quad (4.1.25)
$$

Furthermore, for $y \geq y_2$ we can use (2.2.3) in the following way

$$
\begin{aligned}
J_6 &= \int_{y_2}^{\infty} u(e^{mT}e^{-\sigma\sqrt{T}y})d[-w_{\beta\eta}(g_T(y))] = \int_{y_2}^{\infty} u(x_2\exp\{-\sigma\sqrt{T}(y-y_2)\})d[-w_{\beta\eta}(g_T(y))] \\
&\geq \int_{y_2}^{\infty} u(x_2\exp\{-\sigma\sqrt{T}y\})d[-w_{\beta\eta}(g_T(y))] \geq C_1 u(x_2)\int_{y_2}^{\infty}\exp\{s\sigma\sqrt{T}y\})d[-w_{\beta\eta}(g_T(y))] \\
&\geq -C_1|u(x_2)|\int_{k_1}^{\infty}\exp\{s\sigma\sqrt{T}y\})d[-w_{\beta\eta}(g_T(y))] \geq -2C_1|u(x_2)|e^{-T/4} \qquad (4.1.26)
\end{aligned}
$$

for large T in view of (4.1.18).

Since there exists a constant $C_2 > 0$, perhaps depending on parameters of the problem, such that $u(mT - 2\sigma\beta^{-1/2}T^{(1+\eta)/2\eta}) \geq C_2$ for large T, (4.1.23), (4.1.24), (4.1.25), and (4.1.26) imply for large T that

$$
\begin{aligned}
U(F_T) &\geq u(\exp\{mT - 2\sigma\beta^{-1/2}T^{(1+\eta)/2\eta}\})[1-e^{-T} - C_2^{-1}|u(x_2)|(e^{-T}+2C_1e^{-T/4})] \\
&\geq u(\exp\{mT - 2\sigma\beta^{-1/2}T^{(1+\eta)/2\eta}\})[1-e^{-T/5}], \qquad (4.1.27)
\end{aligned}
$$

where the r.-h.s. is positive.

Turning to the last step of the proof, we fix a policy π and write F_T, m, σ for $F_{T\pi}, m(\pi), \sigma(\pi)$, respectively. Symbols $\hat{F}_T, \hat{m}, \hat{\sigma}$ will correspond to the policy $\hat{\pi}$.

Let $\partial m := \hat{m} - m \geq c > 0$. Combining (4.1.20) and (4.1.27), taking into account that $\eta > 1$, and making use of (2.2.4), we have for large T

$$
\begin{aligned}
\frac{U(F_T)}{U(\hat{F}_T)} &\leq \frac{u(\exp\{mT + 2\bar{\sigma}\beta^{-1/2}T^{(1+\eta)/2\eta}\})[1+e^{-T/5}]}{u(\exp\{\hat{m}T - 2\hat{\sigma}\beta^{-1/2}T^{(1+\eta)/2\eta}\})[1-e^{-T/5}]} \\
&\leq \frac{u(\exp\{(m+\partial m/4)T)\})[1+e^{-T/5}]}{u(\exp\{(\hat{m}-\partial m/4)T)\})[1-e^{-T/5}]} \\
&\leq [1-C(l,d)\exp\{-l(m+\partial m/4)T\}]\cdot\left[1+3e^{-T/5}\right], \qquad (4.1.28)
\end{aligned}
$$

where $d = \frac{\hat{m}-\partial m/4}{m+\partial m/4} \geq \frac{4\hat{m}-c}{4\hat{m}-3c} > 1$. (The denominator is positive since $\hat{m} > m \geq 0$.) Since the positive l can be arbitrary small, we can set, say, $l = \frac{1}{6(m+\partial m/4)} \geq \frac{1}{6(\hat{m}+c/4)} > 0$. Note also that without loss of generality we can assume $C(l,d)$ increasing in both arguments. Similarly, larger l and/or d, larger the set of x's for which (2.2.4) is true. Hence, as is now easy to see, the r.-h. side of (4.1.28) is less than one for large T.

It remains to consider the case when $u(0) > -\infty$, which means that we can set $u(0) = 0$. In this case we keep the bound (4.1.20) as it is, and for a lower bound we can appeal to condition (2.1.4). Set $k_2 = k_2(T) = 2\sqrt{(\nu/a)T}$, where a is a parameter from (2.1.4), and a positive fixed $\nu \leq a/4c_0^2$ will be specified later. Set $\tilde{\Psi}(p) = Mp^a$ for $p \leq 1/2$, and $\tilde{\Psi}(p) = 1$ otherwise. (Note that in (2.1.4) we can certainly consider $M2^{-a} \leq 1$.)

Then

$$\begin{aligned} U(F_T) &\geq \int_0^\infty u(e^{mT}e^{-\sigma\sqrt{T}y})d[-\tilde{\Psi}(g_T(y))] \geq \int_0^{k_2} u(e^{mT}e^{-\sigma\sqrt{T}y})d[-\tilde{\Psi}(g_T(y))] \\ &\geq u(\exp\{mT - 2(\nu/a)^{1/2}\sigma T\})[1-\tilde{\Psi}(g_T(k_2))]. \end{aligned}$$

Since $k_2(T) \leq \sqrt{T}/c_0$, and $g_T(k_2(T)) \leq 1/2$ for large T, by (2.1.4) and (4.1.2) $\tilde{\Psi}(g_T(k_2(T))) = M\exp\{-\nu T\}$. So, for large T

$$U(F_T) \geq u(\exp\{(m - 2(\nu/a)^{1/2}\sigma)T\})[1 - M\exp\{-\nu T\}], \tag{4.1.29}$$

where the r.-h.s. is positive. Thus, (4.1.20) and (4.1.29) imply that

$$\frac{U(F_T)}{U(\hat{F}_T)} \leq \frac{u(\exp\{mT + 2\bar{\sigma}\beta^{-1/2}T^{(1+\eta)/2\eta}\})[1+e^{-T/5}]}{u(\exp\{(\hat{m} - 2(\nu/a)^{1/2}\hat{\sigma})T\})[1 - M\exp\{-\nu T\}]}. \tag{4.1.30}$$

Let $\nu = \min\{(a/4c_0^2), a(\partial m)^2/64\hat{\sigma}^2\}$. Then, since $\eta > 1$, for large T

$$\frac{U(F_T)}{U(\hat{F}_T)} \leq \frac{u(\exp\{(m + \partial m/4)T\})}{u(\exp\{(\hat{m} - \partial m/4)T\})} \cdot \frac{1+e^{-T/5}}{1 - M\exp\{-\nu T\}}. \tag{4.1.31}$$

Proceeding similar to (4.1.28) and applying (2.2.4) it is easy to show now that the r.-h.s. of (4.1.31) is less than one for large T.

The proof is complete.

4.2 Proof of Theorem 3.

The proof repeats the previous proof with the following exceptions. First, we set $u(\infty) = 0$. Second, the upper bound for $U(F_T)$ can be provided in the following way. Let $k_2 = k_2(T)$ be defined as above, and $\tilde{\Psi} = 1 - M(1-p)^a$ for $p > 1/2$, and $\Psi(p) = 0$ otherwise. Then, since $u(x) \leq 0$,

$$\begin{aligned} U(F_T) &\leq \int_0^\infty u(e^{mT}e^{\sigma\sqrt{T}y})d\tilde{\Psi}(1-g_T(y)) \leq \int_0^{k_2} u(e^{mT}e^{\sigma\sqrt{T}y})d\tilde{\Psi}(1-g_T(y)) \\ &\leq u(\exp\{mT + 2(\nu/a)^{1/2}\sigma T\})\tilde{\Psi}(1 - g_T(k_2(T))). \end{aligned}$$

Similar to what we did before we get that $\tilde{\Psi}(1 - g_T(k_2(T))) = 1 - M\exp\{-\nu T\}$ for large T and

$$U(F_T) \leq u(\exp\{(m + 2(\nu/a)^{1/2}\sigma)T\})[1 - M\exp\{-\nu T\}], \tag{4.2.1}$$

where the r.-h.s. is negative.

Consider a lower bound. In this case we set $\tilde{\Psi}(p) = w_{\beta\eta}(p)$ for $p \le 1/2$, and $\tilde{\Psi}(p) = \max\{w_{\beta\eta}(p), \Psi(p)\}$. For the same $k_1 = k_1(T)$ as above

$$U(F_T) \ge \int_0^\infty u(e^{mT} e^{-\sigma\sqrt{T}y}) d[-\tilde{\Psi}(g_T(y))] = \int_0^{k_1} + \int_{k_1}^{\infty} = J_7 + J_8. \tag{4.2.2}$$

First,

$$J_7 \ge u(\exp\{mT - k_1(T)\sigma\sqrt{T}\}) = u(\exp\{mT - 2\sigma\beta^{-1/2}T^{(1+\eta)/2\eta}\}), \tag{4.2.3}$$

where the l.-h.s. is negative.

For J_8 we have to use a condition on $u(x)$ of the type (2.2.3) for all x. Taking into account (4.1.7), we have for large T

$$\begin{aligned} J_8 &= \int_{k_1}^\infty u(e^{mT} e^{s\sigma\sqrt{T}y}) d[-w_{\beta\eta}(g_T(y))] \ge -C_1|u(e^{mT})| \int_{k_1}^\infty e^{s\sigma\sqrt{T}y} d[-w_{\beta\eta}(g_T(y))] \\ &= -C_1|u(e^{mT})|I, \end{aligned}$$

where I is the integral (4.1.11). So, in virtue of (4.1.18), for sufficiently large T

$$U(F_T) \ge u(\exp\{mT - 2\sigma\beta^{-1/2}T^{(1+\eta)/2\eta}\})[1 + e^{-T/5}], \tag{4.2.4}$$

where the r.-h.s. is negative.

The rest of the proof is similar to what has been done before. We should only take into account that $U(F_T)$ is now negative.

4.3 Proof of Theorem 5

We skip the proof of Theorem 4 since it follows the outline of the proof above and the proof of Theorem 5 below.

4.3.1 On quantiles.

Lemma 6 *Let C_q be defined as in (3.7), and for a distribution F*

$$\|F - \Phi\|_\infty \le 1/C_q. \tag{4.3.1}$$

Then

$$|\gamma_q(F) - \gamma_q(\Phi)| \le C_q\|F - \Phi\|_\infty, \tag{4.3.2}$$

where $\gamma_q(F)$ is the $(1-q)$-quantile of F.

PROOF. Set $\bar{q} = 1 - q$, and $\|F - \Phi\|_\infty = \Delta$. We have

$$\begin{aligned} 0 &= F(\gamma(F)) - \Phi(\gamma(\Phi)) = F(\gamma(F)) - \Phi(\gamma(F)) + \Phi(\gamma(F)) - \Phi(\gamma(\Phi)) \\ &\ge -\Delta + \Phi(\gamma(F)) - \bar{q}, \end{aligned}$$

and

$$\gamma(F) - \gamma(\Phi) \le \Phi^{-1}(\bar{q} + \Delta) - \Phi^{1}(\bar{q}). \tag{4.3.3}$$

(If $\bar{q}+\Delta \geq 1$, we assume $\Phi^{-1}(\bar{q}+\Delta)=\infty$.) Note that if the r.-h.s. of (4.3.3) were greater than one, it would imply that $\Delta > \min_{0\leq x\leq 1}\varphi(\gamma(\Phi)+x)\cdot 1$, which contradicts (4.3.1). Thus, from (4.3.3) it follows that

$$\gamma(F)-\gamma(\Phi) \leq [\min_{0\leq x\leq 1}\varphi(\gamma(\Phi)+x)]^{-1}\Delta.$$

The lower bound for $\gamma(F)-\gamma(\Phi)$ may be obtained similarly. ■

4.3.2 Proof of Theorem 5

For brevity we consider only u of Type 1; the general proof is similar. In the case of Type 1, it suffices to consider only truncation in the area of large values, that is, the criterion

$$U(F)=qu(\gamma_q(F))+\int_0^{\gamma_q(F)} u(x)dF(x),$$

where $\gamma_q(F)$ is the $(1-q)$-quantile of F.

Below we again suppress π as an index or an argument. First note that $\gamma_q(F_T)=\exp\{mT+\sigma\sqrt{T}\gamma_q(F_T^*)\}$, and

$$\begin{aligned} U(F_T) &= qu(\gamma_q(F_T))+\int_0^{\gamma_q(F_T)} u(x)dF_T(x) \\ &= qu(\exp\{mT+\sigma\sqrt{T}\gamma_q(F_T^*)\})+\int_0^{\gamma_q(F_T)} u(x)dP(e^{mT+\sigma\sqrt{T}\xi_T}\leq x) \\ &= qu(\exp\{mT+\sigma\sqrt{T}\gamma_q(F_T^*)\})+\int_{-\infty}^{\gamma_q(F_T^*)} u(e^{mT+\sigma\sqrt{T}x})dF_T^*(x). \end{aligned}$$

Let $\Delta_T=\|F_T^*-\Phi\|_\infty$. In view of (3.4) and (3.5), $\Delta_T\leq 1/C_q$,if

$$T\geq C_q^2L^2. \tag{4.3.4}$$

Then in view of (4.3.2), if (4.3.4) holds,

$$|\gamma_q(F_T^*)-\gamma_q(\Phi)|\leq C_q\Delta_T\leq 1. \tag{4.3.5}$$

Thus, if (4.3.4) is true, by Lemma 6,

$$\begin{aligned} U(F_T) &\leq u(\exp\{mT+\sigma\sqrt{T}\gamma_q(F_T^*)\}) \\ &\leq u(\exp\{mT+(\gamma_q(\Phi)+C_q\Delta_T)\sigma\sqrt{T}\} \\ &\leq u(\exp\{mT+(\gamma_q(\Phi)+1)\sigma\sqrt{T}\}. \end{aligned} \tag{4.3.6}$$

For a lower bound, if (4.3.4) is true, we write

$$\begin{aligned} U(F_T) &\geq qu(\exp\{mT+\sigma\sqrt{T}\gamma_q(F_T^*)\}) + \int_0^{\gamma(F_T^*)} u(e^{mT+\sigma\sqrt{T}x})dF_T^*(x) \\ &\geq qu(\exp\{mT+(\gamma_q(\Phi)-C_q\Delta_T)\sigma\sqrt{T}\}) + u(\exp\{mT\}P(0\leq\xi_T\leq\gamma_q(F_T^*)) \\ &\geq qu(\exp\{mT+(\gamma_{1/4}(\Phi)-C_q\Delta_T)\sigma\sqrt{T}\}) + u(\exp\{mT\}(1-q-F_T^*(0)). \end{aligned}$$

It is easy to see that, if $T \geq T_1 = 4(LC_q)^2$, then $C_q\Delta_T \leq (1/2) < \gamma_{1/4}(\Phi)$, and consequently,

$$\begin{aligned} U(F_T) &\geq qu(\exp\{mT\} + u(\exp\{mT\})\,[1-q-F_T^*(0)] \geq \\ &\geq u(\exp\{mT\})\,[(1/2)-\Delta_T] \geq u(\exp\{mT\})\left[(1/2)-(L/\sqrt{T})\right]. \end{aligned} \qquad (4.3.7)$$

Comparing T_1 with (4.3.4), we see that it suffices to consider below $T > T_1$.

We denote as above by $\hat{F}_T$ the distribution of W_T under the policy $\hat{\pi}$. Let $\hat{\Delta}_T = \|\hat{F}_T^* - \Phi\|_\infty$, and again $\partial m = \hat{m} - m > 0$, $\bar{\sigma} = \sup_\pi \sigma(\pi) < \infty$.

From (4.3.6), (4.3.7), and (3.4), it follows that

$$\frac{U(F_T)}{U(\hat{F}_T)} \leq \frac{u(\exp\{mT+(\gamma_q(\Phi)+1)\bar{\sigma}\sqrt{T}\})}{u(\exp\{\hat{m}T\})\left[(1/2)-LT^{-1/2}\right]} \leq \frac{u(\exp\{mT+(\partial m)T/2\})}{u(\exp\{\hat{m}T\})\,[1/4]}, \qquad (4.3.8)$$

if $T \geq T_4 = \max\{T_1, T_2, T_3\}$, where $T_2 = 4(\gamma_q(\Phi)+1)^2\bar{\sigma}^2/(\partial m)^2$, and $T_3 = 16L^2$.

Let first $m + (\partial m)/2 > 0$. Then, by (3.6), for $T \geq T_4$

$$\frac{U(F_T)}{U(\hat{F}_T)} \leq 4\tilde{u}\,(\exp\{(\partial m)T/2\}) < 1, \qquad (4.3.9)$$

if $T > T_5 = 2(\ln k_u)/(\partial m)$.

Thus, $[U(F_T)/U(\hat{F}_T) < 1$, for $T > \max\{T_4, T_5\}$.

If $m + (\partial m)/2 \leq 0$, we write

$$\frac{U(F_T)}{U(\hat{F}_T)} \leq \frac{4u(1)}{u(\exp\{\hat{m}T\})} \leq 4\tilde{u}\,(\exp\{\hat{m}T\}) < 1,$$

if $T > T_6 = \ln k_u/\hat{m}$. It remains to collect all T_i's ■

5 Appendix 1. Two simple examples of inconsistency of MEL and MEU criteria.

5.1 Lognormally distributed returns.

Comparison of the MEU and MEL criteria becomes clearer if we consider variances $\sigma^2(\pi) = Var\{\ln(1+R_1(\pi))\}$, and assume that

$$\sup_{\pi} \sigma(\pi) < \infty. \tag{5.1}$$

In this case one can write

$$W_{T\pi} = \exp\left\{ m(\pi)T + \sigma(\pi)\sqrt{T}\cdot\xi_{T\pi} \right\}, \tag{5.2}$$

where the r.v.

$$\xi_{T\pi} = \frac{\ln W_{T\pi} - m(\pi)T}{\sigma(\pi)\sqrt{T}} = \frac{\sum_{t=1}^{T}[\ln(1+R_t(\pi)) - m(\pi)]}{\sigma(\pi)\sqrt{T}} \tag{5.3}$$

is asymptotically normal, that is, for all π

$$P(\xi_{T\pi} \le x) \to \Phi(x) \quad \text{for all } \ x, \ \text{ as } \quad T \to \infty,$$

where Φ is the standard normal distribution function.

In view of (5.1) and (5.3) the term $m(\pi)T$ in (5.2) is dominant, so the policy $\hat{\pi}$ is asymptotically optimal in the sense of (1.1.1). For comparison we apply the MEU criterion with $u(x) = x^{\alpha}$. Assume for simplicity that $\xi_{T\pi}$ equals in distribution a standard normal r.v. ξ, which corresponds to log-normal returns $1+R_t$ (as, say, in the geometric-Brownian-motion scheme). Then (see also, e.g., Merton and Samuelson (1974))

$$\begin{aligned} E\{u(W_{T\pi})\} &= E\{u(\exp\{m(\pi)T + \sigma(\pi)\sqrt{T}\xi_{T\pi}\})\} = \exp\{\alpha m(\pi)T\}E\left\{\exp\{\alpha\sigma(\pi)\sqrt{T}\xi\right\} \\ &= \exp\{(\alpha m(\pi) + \alpha^2\sigma^2(\pi)/2)T\}. \end{aligned} \tag{5.4}$$

We see that the Laplace transform above "transforms $\sqrt{T}$ into T", and the influence of the non-random component $m(\pi)T$ and that of the random component $\sigma(\pi)\sqrt{T}\xi_{T\pi}$ prove to be of the same order.

Obviously, maximization of (5.4) in π may lead not to $\hat{\pi}$, and as a rule, if a policy $\tilde{\pi}_T$ (perhaps depending on T) is a maximizer of $E\{u(W_{T\pi})\}$, then $[E\{u(W_{T\tilde{\pi}_T})\}/E\{u(W_{T\hat{\pi}})\}] \to \infty$, as $T \to \infty$.

If ξ_T is normal only asymptotically, then the behavior of $Eu(W_{T\pi})$ for large T may not coincide with that when $\xi_T = \xi$: passing the limit inside the Laplace transform is improper in general, which was thoroughly investigated in Merton and Samuelson (1974). However for demonstration of the inconsistency of the two criteria under discussion the example when $\xi_T = \xi$, is enough.

5.2 A bounded utility function.

It may seem that for a bounded u, say if $u(\infty) < \infty$, the maximization of $Eu(W_{T\pi})$ can lead asymptotically, as $T \to \infty$, to $\hat{\pi}$ (since $W_{T\pi} \to \infty$, and hence $Eu(W_{T\pi}) \to u(\infty)$ for all "reasonable" π's). However it is also not true in general, though to make it explicit one should find a "good way" to compare two strategies for such a case. Merton and Samuelson (1974) considered to this end "the additional initial wealth the investor would require to be indifferent to giving up a program π' for a program π'' ". As another approach one can proceed from the rate with which the expected utility converges to the maximal possible, that is, the rate of convergence in $Eu(W_{T\pi}) \to u(\infty)$.

Consider the simplest case when again $\xi_T = \xi$, and $u(x) = x$, if $x \le 1$, and $= 1$, if $x > 1$. For a policy π set $m = m(\pi)$, $\sigma = \sigma(\pi)$. Straightforward calculations lead in this case to

$$1 - Eu(W_{T\pi}) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{T}}\left(\frac{\sigma}{m} - \frac{\sigma}{m+\sigma^2}\right)\exp\left\{-\frac{m^2}{2\sigma^2}T\right\}\left(1 + O(T^{-3/2})\right).$$

Hence if for another policy $\tilde{\pi}$ with $\tilde{m} = m(\tilde{\pi})$, $\tilde{\sigma} = \sigma(\tilde{\pi})$, it is true that $(m/\sigma) > (\tilde{m}/\tilde{\sigma})$, then

$$[1 - Eu(W_{T\pi})] \big/ [1 - Eu(W_{T\tilde{\pi}})] \to 0, \text{ as } T \to \infty.$$

For properties of the MEL-portfolio as applied to bounded utilities see also Goldman (1974).

6 Appendix 2. Rank dependent expected utility: a short survey.

6.1 A general framework.

We consider below only positive random variables (r.v.'s) and, respectively, probability distributions on $[0, \infty)$. As before, consider a preference ordering preserved by the functional

$$U(F) = \int_0^\infty u(x) d\Psi(F(x)), \tag{6.1}$$

where u is a *utility* function, and the function Ψ is assumed to be non-decreasing, $\Psi(0) = 0$, $\Psi(1) = 1$. Note that such a Ψ is itself a distribution function on $[0,1]$, and the function $\Psi(F(x))$ is also a distribution function on $[0, \infty)$. The domain of $U(F)$ is the set of F's for which $U(F)$ - for a given fixed u - is defined.

If $\Psi(p) = p$ (the subject perceives F as it is, not transforming it), we deal with the "usual" expected utility. If for a fixed $q \in [0,1]$,

$$\Psi(p) = \begin{cases} p & \text{if } p < 1 - q, \\ 1 & \text{if } p \geq 1 - q, \end{cases} \tag{6.2}$$

we come to (1.1.4).

The general case (6.1) is referred to as *Rank Dependent Expected Utility (RDEU).*

As was already mentioned, quite often the criterion (6.1) is specified in terms of not the transformation function Ψ but the function $w(p) = 1 - \Psi(1 - p)$which is referred to as a *weighting function*, though the same term is often applied to Ψ too. If $u(0) = 0$ one can rewrite (6.1) as

$$U(F) = \int_0^\infty w(1 - F(x)) du(x).$$

If a distribution F is that of a r.v. taking only two values, say, a and $b < a$ with probabilities p and $1 - p$, respectively, then

$$U(F) = u(a)w(p) + u(b)[1 - w(p)], \tag{6.3}$$

and $w(p)$ "transforms" the probability p.

The full RDEU model, including a set of axioms leading to (6.1), was first proposed in Quiggin (1982), though some earlier works of Quiggin had already contained some relevant ideas; see references in Quiggin (1993). Some models including weighting functions, say, in the case of binary gambles were considered earlier in the prospect theory of Kahneman and Tversky (1979). A special case of RDEU was independently considered in the "dual model" of Yaari (1987) developed further by Roell (1987). To my knowledge, an axiomatic system for the most general case including continuous distributions was considered in Wakker (1993, 1994). A rather full history of the question and a rich bibliography may be found in monographs Wakker (1989), Quiggin (1993), and Luce (2000).

It is worth noting that the general modern RDEU model is more flexible than (6.1), and deals not only with probability distributions but with the corresponding events structures as well. The description of this model which coincides with that of (6.1) in the case of so called coalescing, may be found in Luce (2000); one of the most recent axiomatic systems - in Marley and Luce (1992). In the present paper we restrict ourselves to (6.1), and hence to orders on spaces of probability distributions.

There are several axiomatic justifications of the criterion (6.1); all of them can be found in books Wakker (1989), Quiggin (1993), and Luce (2000). A key axiom is either the trade-off consistency requirement [Wakker (1994)] or the different, though in a certain sense similar, ordinal dependence axiom [Green and Jullien (1988), Quiggin (1989), Segal (1989)]. The latter axiom requires that if two probability distribution functions (d.f.'s) coincide on an interval, their values on that interval should not affect their ranking. In other words, if $F(x)$ and $G(x)$ are two d.f.'s, and $F(x) = G(x)$ for all x's from a set A, then when changing $F(x)$ and $G(x)$ only on A, keeping them to be equal to each other, we do not change the order relation between F and G. This is very much similar to the Savage sure-thing principle. Certainly such an axiom is considerably weaker than the independence axiom.

Consider now how Ψ can look.

6.2 Possible properties of Ψ.

In Sections 1.1.2 and 1.2 we have considered already the simple case when Ψ is a power function, more specifically when $\Psi(p)$ equals either p^β or $1-(1-p)^\beta$. Such weighting functions arise in the case of the so called first order reduction of compound gambles; see for a definition and comments, e.g., Luce (2000, p.84). In the modern theory this is viewed as "too simple".

In the last ten years there has been a great deal of investigation - theoretical and experimental - of more sophisticated types of the transformation function Ψ. Usually answers are being given in terms of the weighting function w. Surveys of many results may be found in Quiggin (1993), and especially in Luce (2000), including some empirical evidence; see also Wu and Gonzales (1996, 1999), and references therein.

We list below typical situations. (The terminology is from Lopes (1987, 1990); the interpretation is based on the fact that the "transformation of the expected utility" by Ψ in (6.1) is due to the values of its derivatives rather than to the values of Ψ themselves. When saying that the subject underestimates the probability of an event, we mean that the subject takes to less extent into account that this event may occur; in the extreme case the subject just neglects the possibility of such an event.)

- $\Psi(p) > p$ and concave (hence $w(p) < p$ and convex): the larger a given level for values of the random variable under consideration, the less the subject takes into account the possibility that the random variable will reach this level; the subject is security minded.

- $\Psi(p) < p$ and convex ($w(p) > p$ and concave): the opposite case; the subject is potential-minded.

- $\Psi(p)$ is inverse S-shaped ($w(p)$ is of the same shape): cautiously hopeful; roughly speaking the subject overestimates the probabilities of "very large" and "very small" levels.

- $\Psi(p)$ is S-shaped ($w(p)$ is also S-shaped): the subject is oriented to moderate values of the random variable, and underestimates the probabilities of very large and very small values.

In theorems of this paper the function $\Psi(p)$ is close to that of an S-shape for small and large p's, though it should be noted that many experiments presented in the literature testify rather to the inverse S-shaped pattern [see, e.g., Wu and Gonzales (1996, 1999) and the survey in Luce (2000)]. In the author's opinion, it can be connected with the fact that all experiments mentioned (usually dealing with students) concern - to my knowledge - one-time gains or investments, and the amounts of money involved are not large. In such a situation it is not surprising that people can count to some extent on large deviations even overestimating real probabilities of their occurrence. However, in long run investment, dealing with significant amounts of money, and in situations when these amounts really matter for the investor (say, we talk about pension plans), the investor may exhibit an opposite behavior based on moderate values of the wealth rather than possibilities of large deviations.

6.3 Prelec's representation.

For an explicit representation for Ψ we appeal to a though relatively recent but already well known paper by Prelec (1998), who provided an axiomatic system leading to the weighting function

$$w(p) = w_{\beta\eta}(p) = \exp\left\{-[-\beta\ln(p)]^{\eta}\right\}, \tag{6.4}$$

where parameters $\beta, \eta > 0$. Actually, Prelec writes $w_{\beta\eta}(p) = \exp\left\{-\beta[-\ln(p)]^{\eta}\right\}$. We prefer equivalent (6.4) which seems more convenient for a number of reasons; for example it allows to consider, in a natural way, the case $\eta = \infty$ which corresponds to truncation (see below), or for $\eta = 1$ directly leads to p^{β}.

The corresponding transformation function $\Psi_{\beta\eta}(p) = 1 - w_{\beta\eta}(1-p)$.

The most essential axiom among those which lead to (6.4) is that of N-compound invariance for $N = 2,3$; see Prelec (1998), or Luce (2000) for details. Luce (2000, 2001) suggested another axiom, namely N-reduction invariance for $N = 2,3$, which leads to the same representation. The latter axiom is formulated in terms of binary gambles where one of the consequences is no change from the status quo; see Luce (2000, 2001) for details. In my opinion, the latter type of axioms is simpler. See also a neat discussion and generalizations in Luce (2000).

We discuss briefly properties of $w_{\beta\eta}$.

- First, (6.4) includes the case of power functions: $w_{\beta 1}(p) = p^{\beta}$.

- If $\eta \neq 1$, function $w_{\beta\eta}$ is asymmetric, and have a unique fixed point $p_{\mathrm{f}} = \exp\{-\beta^{\eta/(1-\eta)}\}$, and a unique inflection point $p_{\mathrm{i}} = e^{-c}$ where c is a solution to the equation $\eta\beta c^{\beta} = \eta - 1 + c$. If $\beta = 1$ both points equal $1/e$.

- If $\eta > 1$, $w_{\beta\eta}$ is S-shaped: convex on $[0,\ p_{\mathrm{i}}]$ and concave beyond it. In this case $w_{\beta\eta}(p) \to 0$, as $p \to 0$, faster than *any* power function.

- If $\eta < 1$, $w_{\beta\eta}$ is inverse-S-shaped: concave on $[0,\ p_{\mathrm{i}}]$ and convex beyond it. In this case $w_{\beta\eta}(p) \to 0$, as $p \to 0$, slower than *any* power function.

In accordance of what was said about the S-shape above, we will consider (6.4) for $\eta > 1$. It is worth emphasizing however that for our purposes we do not need Prelec's representation to be true exactly: $w(p)$ will serve merely as a *bound* for $w(p)$ and only for p's close to zero and/or to one.

Respectively, axioms implying (6.4) are not necessary for our purposes, and we can view them as those describing a limiting allowed case of the behavior in the area of large deviations.

In conclusion note that, to include truncation into Prelec's framework, one may set

$$w(p) = w_{\beta\eta}(p) = \begin{cases} w_{\beta\eta}(p) & \text{for } p \le p_{\mathrm{f}}, \\ p & \text{for } p > p_{\mathrm{f}}, \end{cases}$$

(where p_{f} is the fixed point of $w_{\beta\eta}$). Then, as $\eta \to \infty$,

$$w_{\beta\eta}(p) \to \begin{cases} 0 & \text{for } p \le e^{-\beta} \\ p & \text{for } p > e^{-\beta} \end{cases},$$

which leads to (6.2) if $\beta = -\ln q$.

Acknowledgments: I am thankful to E.Omberg for calling my attention

to papers [24] and [15], and to R.D.Luce and H.Markowitz for very useful discussions.

References

[1] Algoet, P. and Cover, T (1988), Asymptotic Optimality and Asymptotic Equipartition Properties of Log-Optimum Investment, Annals of Probability, 1988, **16**, 876–898.

[2] Arnold, V.M., Evstigneev, I.V. and Gundalach, V.M. (2000) Convex-valued random dynamic systems: A variational principle for equilibrium states, Random Operators and Stochastic Equations 7, 23-38.

[3] Breeden, D. (1979), An intemporal asset pricing model with stochastic consumption and investment opportunities, Journal of Financial Economics, **7**, 265-296.

[4] Breiman, L. (1961), Optimal gambling systems for favorable games. Proc. Fourth Berkeley Symp. Math. Statist. Probab. **1**, 65-78, California Press.

[5] Dow, James & Sérgio R.C. Werlang (1992), Excess

Volatility of Stock Prices and Knightian Uncertainty, European Economic Review 36, 631-638.

[6] Duffie, D. (1992) Asset pricing theory. Princeton: Princeton Univ.Press.

[7] Epstein, Larry G. & Tan Wang (1994), Intertemporal Asset Pricing under Knightian Uncertainty, Econometrica 62, 283-322.

[8] Evstigneev, I.V. and Taksar, M.I. (2001) Rapid growth paths in convex-valued random dynamical systems, Stochastics and Dynamics, v. 1, 493 - 509.

[9] Feller, W. (1971) An Introduction into Probability theory and its Applications, V.2, New York : Wiley.

[10] Green, J. and Jullien, B. (1988) Ordinal independence in non-linear utility theory, Quarterly Journal of Economics 102 (4), 785-796.

[11] Goldman, Barry (1974), A negative report on the ‘near optimality’ of the max-expected-log policy as applied to bounded utilities for long-lived programs, Journal of Financial Economics 1, 97-103.

[12] Hakansson, N. Multi-period mean-variance analysis: Toward a general theory of portfolio choice, Journal of Finance, **26,** 868-870

[13] Ibragimov, I.A. and Linnik, Yu.V. (1971) Independent and stationary sequences of random variables. Groningen, Wolters-Noordhoff.

[14] Kahneman, D. and Tversky, A.(1979), Prospect Theory: Analysis of decision under risk, Econometrica, **47,** 263-291.

[15] Kim, T., Omberg, E., and Russell, T. (1993), The Log-Optimum and Expected-Utility Maximizing Portfolio Strategies for Long-Horizon Investors. Working paper, San Diego State University.

[16] Latané, H. (1959), Criteria for choice among risky ventures, Journal of Political Economy 67, 144-155.

[17] Latané, H. (1979), The geometric mean criterion continued, Journal of Banking and Finance, **3**, 309-311.

[18] Loeve, M. (1977) Probability Theory, 4th ed. New York : Springer-Verlag.

[19] Lopes, L.L. (1987), Between hope and fear: the psychology of risk. In. L. Berkowitz (ed.). Advances in Experimental Social Psychology, vol. 20, N.-Y.: Academic Press, pp. 225-295.

[20] Lopes, L.L. (1990), Re-modeling risk aversion. In G.M.Furstenberg (ed.) Acting under Uncertainty,: Multidisciplinary Conceptions. Boston: Kluwer, pp. 267-299.

[21] Luce, R.D. (2001) Reduction invariance and Prelec’s weighting function. J. of Mathematical psychology, **45**, 167-179

[22] Luce, R.D. (2000) Utility of Gains and Losses. Erlbaum.

[23] Markowitz, H. (1959), Portfolio Selection, Willey, New York.

[24] Markowitz, H. (1976) Investment for the long run: New evidence for an old rule. Journal of Finance, **31**, 1273-1286.

[25] Marley, A.A and Luce, R.D. (2002), A simple axiomatization of binary rank-dependent expected utility of gains (losses), Journal of Mathematical Psychology, **46**, 40-55.

[26] Merton, R. (1973), An intemporal capital asset pricing model, Econometrica, **41**, 867-887.

[27] Merton, R. and Samuelson P. (1974), Fallacy of the log-normal approximation to optimal portfolio decision making over many periods, Journal of Financial Economics, **1**, 67-94.

[28] Merton, R. (1992) Continuous time finance, Blackwell Publishes.

[29] Mukerji, Sujoy & Jean-Marc Tallon (1998), Ambiguity Aversion and Incompleteness of Financial Markets.

[30] Ophir, Tsvi (1978), The geometric-mean principle revisited, Journal of Banking and Finance 2, 103-107.

[31] Ophir, Tsvi (1979), The geometric-mean principle revisited: a reply to a 'reply', Journal of Banking and Finance 3, 301-303.

[32] Prelec, D. (1998) The Probability Weighting Function, Econometrica, **66**, No.3, 497-527.

[33] Quiggin, J. (1993) Generalized Expected Utility Theory. Kluwer.

[34] Quiggin, J. (1982) A theory of anticipated utility, Journal of Economic Behavior and Organization, **3**, 324-43.

[35] Quiggin, J. (1989) Stochastic dominance in regret theory, Review of Economic Studies 57 (2), 503-511.

[36] Roell, A. (1987) Risk aversion in Quiggin and Yaari's rank-order model of choice under uncertainty, Economic Journal 97, 143-159.

[37] Rotar, V.I. (1998) Probability Theory, World Scientific.

[38] Samuelson, P. (1969) Life-time portfolio selection by dynamic stochastic programming, Review of Economics and Statistics, **51**, 239-246.

[39] Samuelson P. (1971), The "fallacy" of maximizing the geometric mean in long sequences of investing or gambling, Proc. Nation. Acad. Sci. USA, **68**, 2493-2496.

[40] Samuelson P. (1979), Why we should not make mean log of wealth big though years to act are long, Journal of Banking and Finance 3, 305-307.

[41] Samuelson P. (1988), Long-run risk tolerance when equity returns are mean regressing: Pseudoparadoxes and vindication of 'business man's risk', The May 7-8 1988 Yale Symposium in honor of James Tobin.

[42] Segal, U. A. (1989) Anticipated utility: a measure representation approach, Annals of Operation Research 19,359-374.

[43] Seneta, E. (1976) Regularly varying functions. Lecture Notes in Mathematics, Vol. 508.

[44] Simonsen, M. & Sérgio R.C. Werlang (1991), Subadditive Probabilities and Portfolio Inertia, Revista de Econometria 11, 1- 19.

[45] Tallon, Jean-Marc (1998), Do Sunspots Matter when Agents Are Choquet-Expected-Utility Maximizers, Journal of Economic Dynamics and Control 22, 357-368.

[46] Wakker, P.P. (1989) Additive Representations of Preferences. Kluwer.

[47] Wakker, P.P. (1993), Unbounded utility for Savage's "Foundations of Statistics", and other models, Mathematics of Operations Research, **18**, 446-485.

[48] Wakker, P. (1994), Separating Marginal Utility and Probabilistic Risk Aversion, Theory and Decision **36**, 1-44.

[49] Wu, G. and Gonzales, R. (1996) Curvature of probability weighting function, Management Sciences **42**, 1676-1690.

[50] Wu, G. and Gonzales, R. (1999) Non-linear decision weights in choice under uncertainty, Management Sciences **45**, 74-85.

[51] Yaari, M.E. (1987), The dual theory of choice under risk, Econometrica, **55**, 95-115.