Large Deviations Estimation of the Windfall and Shortfall Probabilities for Optimal Diversified Portfolios*

Ba Chu†
Carleton University
15th June 2010

Abstract

Many investors believe that they can effectively reduce risk by, among other ways, holding large combinations of investment assets. The purpose of this paper is to develop asymptotic approximations of the windfall and shortfall probabilities for an optimal portfolio of risky assets as the number of the assets becomes sufficiently large. We start by providing some heuristics to motivate our problem, then proceed to prove general large deviations theorems. We also present specific results with an application to the multivariate normal case. Our theoretical results justify the diversification tenet of the allocation strategies that many hedge funds and pension funds tend to adopt nowadays.

JEL classifications: C60; C13; G11

Keywords: Diversification; Large deviations; Shortfall probabilities; Windfall probabilities

*The author would like to sincerely thank Juan Carlos Escanciano, Kim P. Huỳnh, Joon Y. Park, and Phạm Đình Tạọ for many insightful suggestions. The author is indebted to Chi Mai Nguyễn for having carefully read the original manuscript and provided many constructive comments. The remaining errors are the responsibilities of the author. The usual disclaimer applies.

†Department of Economics, Carleton University, B-857 Loeb Building, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada. Phone: +1 (613) 520 2600 (ext. 1546). Fax: +1 (613) 520 3906. E-mail: ba_chu@carleton.ca. Web Page: [http://http-server.carleton.ca/~bchu/](http://http-server.carleton.ca/~bchu/)
1 Introduction

In financial economics the paradigm of diversification has appeared in various guises over the last twenty years, either as a portfolio problem (e.g. Samuelson [1967]), or more generally as an insurance problem where diversifiable risk can be eliminated from equilibrium allocation, and asset prices are determined by nondiversifiable risk factors (e.g. Malinvaud [1972]). The fundamental rationale for the consideration of infinite asset markets is that diversification pays – risk can be significantly reduced by spreading endowment across a large number of imperfectly correlated assets. The merits of diversification are clear in linear factor models for equity returns, which form the basis of the arbitrage pricing theory (APT) of Ross [1976, 1977] and its extensions to general large equilibrium asset markets (e.g. Chamberlain and Rothschild [1983]; Khan and Sun [2003]). The APT builds on the intuition that a large economy offers investors the opportunity to eliminate idiosyncratic risk through diversification of asset holdings, so that an investor who holds a well-diversified portfolio is effectively insured against idiosyncratic risk. In other words, if assets are allocated so that all investors hold well-diversified portfolios, then all investors are insured. As such, a well-diversified portfolio must have zero idiosyncratic risk.

Accordingly, diversification is an important investment vehicle. Indeed, this was the primary motivation for creation of the very first hedge fund with the aim of overperforming (or beating) a given benchmark return (usually proxied by the change in a stock market price index or a bond yield), especially in the periods when markets are volatile. Hence, maximization of the windfall probability (i.e., the probability that the hedge fund return will overperform a given benchmark return) or minimization of the shortfall probability (i.e., the probability that the hedge fund return will underperform this benchmark return) has, certainly, become a ultimate objective of many investment strategies, including hedge funds (see, e.g., Fishburn [1984], Stutzer [2003], or Basak et al. [2006], and many others). On the other side, most of investors generally tend to maximize their profit or minimize their loss. So, our first impression could be that the investor is naturally neutral to minimization of the shortfall probability and maximization of a profit function (or to
maximization of the windfall probability and minimization of a loss function). In the present paper, we endeavor to provide an analytically viable formalization of this intuition.

It is worth mentioning at this point that the shortfall probability is a special case of the lower-partial moments proposed by [Lee and Rao 1988]. And the shortfall probability stimulates diversification because it is a subadditive function (i.e., \( P(X + Y \leq r) \leq P(X \leq r) + P(Y \leq r) \)), notwithstanding a caveat that it does not take into account the severity of an incurred damage event as required by coherent risk measures \(^1\), which were proposed by [Artzner et al. 1999]. Hence, in the present paper we shall not be concerned with the coherence issue, but focus on the main theme – the risk of optimal diversified portfolios.

Furthermore, one can not determine the optimal weights of assets invested in a fund without the knowledge of the distribution of the returns for all possible portfolios. This issue becomes more complicated as the number of individual assets increases; and in most cases, the closed form of this return distribution does not exist, thus one needs to rely on a variety of asymptotic approximations. A line of work has focused on using the inverse Fourier transform and integral approximation methods. For instance, [Glasserman et al. 2002] propose an importance sampling algorithm to approximate an inverse Fourier integral used for computing the portfolio loss distribution when the underlying assets have a heavy-tailed distribution. However, this approach may be neither valid nor tractable when the size of portfolio is large, or the heavy-tailedness assumption is relaxed. The present paper pursues a different line by developing general large deviations approximations for the optimal windfall and shortfall probabilities. (Readers are referred to [Dembo and Zeitouni 1998] for an extensive account of the large deviations theory.)

As such the thrust of this paper is to develop a general framework for one-period optimal investments in diversified portfolios. This is a practically interesting problem for the following two reasons: first, as having explained from the outset, investors can, for most of the time, reduce idiosyncratic risks by holding diversified portfolios. Second, diversification is more relevant (and indeed less costly) for short-term investments than long-term ones. In the long term, diversification

\(^1\)The practical usefulness and pitfalls of coherent risk measures have been discussed a great deal in the risk management literature (see, e.g., [Heyde et al. 2007]).
is not necessary for the reason that risky assets appear safer over a longer time frame as idiosyncratic risks decline over time. And diminishing idiosyncratic risks are due to the fact that more information about firms become available at investors’ disposal – that is, the longer one times the horizon the more predictable returns become. Thus, reminiscent of the results of Barberis [2000], in a portfolio containing cash and a stock, the allocation to the stock increases with the horizon.

Therefore, the contribution of the present paper is twofold. First, we derive large deviations approximations of the windfall and shortfall probabilities for an optimal large portfolio based on dual problems on the profit and loss functions. The intuition is that there may exist a unbounded set of optimal portfolio choices that yields a maximum windfall probability (or a minimum shortfall probability) for the portfolio return, as the number of assets in these portfolios increases to infinity, thus every investor can always do best by maximizing the expectation of a particular endogenous profit function so as to achieve the same maximum windfall probability (or minimizing the expectation of a particular endogenous loss function so as to achieve the same minimum windfall probability). In fact, we show that an investor who has an exponential or power profit function can attain the maximum windfall probability; and an investor who has an exponential or power loss function can hit the minimum shortfall probability. In this sense, our results vindicate the basic premise of optimal investment that diversification in a conventional portfolio optimization problem (i.e., maximizing/minimizing the profit/loss of a final wealth) helps to reduce the shortfall risk.

Second, we apply the proposed approximations to derive the minimum shortfall probability in the well-known case whereby the joint distribution of risky assets’ returns is multivariate normal.

However, it is important to note that the results obtained in the present paper assume perfect knowledge of parameters – such as expected returns, variances and covariances. The issues of parameter uncertainty must also be addressed. But this falls outside the scope of the present paper, and needs to be addressed in future research.

The outline of the paper is as follows. Section 2 starts the exposition of the main theme with an

\footnote{For example, suppose that \( X \sim N(0, 1); Y \sim N(0, 1); \) and \( X \) and \( Y \) are independent. We can immediately see that \( 0.2X + 0.8Y \sim N(0, 0.2^2 + 0.8^2) \) and \( 0.8X + 0.2Y \sim N(0, 0.8^2 + 0.2^2) \). This implies that two portfolios, \((0.2, 0.8)\) and \((0.8, 0.2)\), yield the same windfall and shortfall probabilities.}
exact treatment of the multivariate normal case. Section 3 presents general results with discussions. Section 4 presents some specific applications. We shall note at this point that there are two main reasons for us to start with the general theory rather than with specific cases: on the one hand, the general theory implies the strongest cases (including the multivariate normal case) and, on the other hand, the proof argument of the specific cases cannot be immediately adapted to that of the general theory. Finally, Section 5 concludes this paper. For future references, some symbols and notations are tabulated as follows:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_i$</td>
<td>the return (or the profit) of asset $i$</td>
</tr>
<tr>
<td>$r$</td>
<td>a chosen benchmark return</td>
</tr>
<tr>
<td>$\alpha = {\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n}'$</td>
<td>a portfolio of $n$ risky assets</td>
</tr>
<tr>
<td>$1(x \in A)$</td>
<td>an indicator function, “equal to 1 if $x \in A$, and 0 otherwise”</td>
</tr>
<tr>
<td>$1$</td>
<td>a column vector of ones</td>
</tr>
<tr>
<td>$S_n^{(\alpha)} = \sum_{i=1}^n \alpha_i X_i$</td>
<td>the return of a portfolio of $n$ risky assets</td>
</tr>
<tr>
<td>$\tilde{\alpha}(\theta)$ (abbrev. $\tilde{\alpha}$)</td>
<td>a $n$-dimensional optimal [portfolio] control vector [if it exists] of the supremum $\sup_{\alpha \in A} \log E[\ell(S_n^{(\alpha)}, n\theta)]$ if $\theta &gt; 0$, and of the infimum $\inf_{\alpha \in A} \log E[\ell(S_n^{(\alpha)}, n\theta)]$ if $\theta &lt; 0$</td>
</tr>
<tr>
<td>$\tilde{\alpha}(\hat{\theta})$ (abbrev. $\tilde{\alpha}$)</td>
<td>a $n$-dimensional optimal [portfolio] control vector evaluated at an exposing point, $\hat{\theta}$</td>
</tr>
<tr>
<td>$\overset{a}{=} \quad \text{“asymptotically equal”}$</td>
<td></td>
</tr>
<tr>
<td>$B_r(\epsilon)$</td>
<td>a closed ball with the center $r$ and the diameter $\epsilon$</td>
</tr>
<tr>
<td>$(a, b)^+$</td>
<td>the maximum number out of ${a, b}$</td>
</tr>
<tr>
<td>$(a, b)^-$</td>
<td>the minimum number out of ${a, b}$</td>
</tr>
<tr>
<td>$\mathbb{P}$</td>
<td>the joint distribution of ${X_1, \ldots, X_n}$</td>
</tr>
<tr>
<td>$\mathbb{P}^*$</td>
<td>the conjugate distribution of $\mathbb{P}$</td>
</tr>
<tr>
<td>$\mathbb{R}$</td>
<td>the real line</td>
</tr>
<tr>
<td>$\mathbb{R}^+$</td>
<td>the positive real line</td>
</tr>
<tr>
<td>$\mathbb{R}^-$</td>
<td>the negative real line</td>
</tr>
</tbody>
</table>
Note that, if \( S_n^{(\alpha)} \) is the gross return, then \( S_n^{(\alpha)} \) is always nonnegative for all \( \alpha \in [0,1)^n \).

2 Fundamental Rationale

To place the results obtained and the viewpoints taken in this paper into proper perspective, we shall analyze a close-to-trivial, but extremely enlightening, case pertaining to the multivariate normal returns. For a given number of risky assets, \( n \), the z-standardization yields

\[
P \left( S_n^{(\alpha)} \leq r \right) = P \left( \frac{S_n^{(\alpha)} - \alpha' \mu}{\sqrt{\alpha' \Sigma \alpha}} \leq -\frac{\alpha' \mu - r}{\sqrt{\alpha' \Sigma \alpha}} \right) = \Phi \left( -\frac{\alpha' \mu - r}{\sqrt{\alpha' \Sigma \alpha}} \right),
\]

where \( \mu \) represents the vector of expected returns, \( \Sigma \) denotes the variance-covariance matrix of asset returns, and \( \Phi(\bullet) \) is the standard normal c.d.f. This relation shows that a minimal shortfall-probability (sp) portfolio maximizes the Sharpe Ratio, \( SR(\alpha) = \frac{\alpha' \mu - r}{\sqrt{\alpha' \Sigma \alpha}} \). As such, a portfolio with minimal sp is analogous to that with maximal Sharpe Ratio. To illuminate this intuition, it is to be noted that \( \Phi(z) \approx 1 - \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \). It follows that, since \( \frac{\partial \Phi(z)}{\partial z} = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left\{ 1 + \frac{1}{z^2} \right\} > 0 \), the relation \( \frac{\partial \Phi(-SR(\alpha))}{\partial \alpha} = \frac{\partial \Phi(z)}{\partial z} \bigg|_{z=-SR(\alpha)} \frac{\partial SR(\alpha)}{\partial \alpha} = 0 \) together with the constraint that the matrix

\[
\frac{\partial^2 \Phi(-SR(\alpha))}{\partial \alpha \partial \alpha'} = -\frac{\partial^2 SR(\alpha)}{\partial \alpha \partial \alpha'} \bigg|_{z=-SR(\alpha)} - \frac{\partial^2 \Phi(z)}{\partial z^2} \bigg|_{z=-SR(\alpha)} \left( \frac{\partial SR(\alpha)}{\partial \alpha} \right)^2
\]

is positive-definite imply that \( \frac{\partial SR(\alpha)}{\partial \alpha} = 0 \) and \( \frac{\partial^2 SR(\alpha)}{\partial \alpha \partial \alpha'} \) is negative-definite.

Because the Sharpe Ratio is a complex nonlinear function of portfolio weights, \( \alpha \), it would be impossible to derive a closed-form expression for the maximal Sharpe Ratio portfolio, or equivalently, the minimal sp portfolio. In this sense the Sharpe Ratio may contain multiple optimal points. Here we shall discuss two conventional mean-variance strategies and their optimality in terms of the minimal sp portfolio insofar as they are well-diversified in the sense of Chamberlain and Rothschild [1983], as defined below.
Definition 1 (well-diversification). The portfolio $\alpha$ is well-diversified if

$$\|\alpha\|_2 = \left( \sum_{i=1}^{n} \alpha_i^2 \right)^{1/2} \to 0 \text{ in probability.}$$

Remark 2.1. Well-diversification of a given portfolio, $\alpha$, implies that, by investing infinitesimal fractions of wealth in each of a large number of risky assets, the idiosyncratic risk – namely the contribution of $\epsilon_t$ in a factor model $X - 1r = E[X] + Bz + \epsilon$, where the vector $z$ contains factors and the matrix $B$ contains factor loadings, to the portfolio excess return $\alpha'(X - 1r)$ – vanishes in mean square. That is,

$$\text{var}(\alpha' \epsilon) = \alpha' G \alpha \to 0 \text{ in probability},$$

where $G$ is the variance-covariance matrix of $\epsilon$, if the maximum eigenvalue, $\rho(.)$, of $G$ does not grow too quickly since $\alpha'G\alpha \leq \|\alpha\|_2 \rho(G)$.

The first strategy is the global minimum-variance (gmv) portfolio defined by $\hat{\alpha}_{GMV} = \frac{1}{\Sigma^{-1}}$. (see Ingersoll [1987, p. 86]). One can immediately see that the Sharpe Ratio $SR(\hat{\alpha}_{GMV}) = \left(1'\Sigma^{-1}1\right)^{1/2} \left(\hat{\alpha}_{GMV}'\mu - r\right)$ diverges to infinity if $\mu \to \infty$ and $\Sigma \to 0$. Nevertheless, this scenario is very unlikely to occur because of the no-arbitrage constraint – an appealing and simple tenet behind the APT (see Ross [1982]). More specifically, if the market permits arbitrage opportunities then investors do not have to choose between mean and variance. They can, for a given price, acquire portfolios which have arbitrarily high expected returns and arbitrarily low variances. Accordingly the Sharpe Ratio must always be finite; thus it is impossible to eliminate the shortfall risk completely, or at least, to attain the lower bound of the sp with probability one, even when the number of risky assets becomes very large. In this sense the gmv portfolio is suboptimal. This portfolio is well-diversified if $\|\hat{\alpha}_{GMV}\|_2 = \left(1'\Sigma^{-1}1\right)^{-1/2} 1'\Sigma^{-2}1 \to 0$ as $n \to \infty$, which is valid as long as $1'\Sigma^{-1}1 = O(n)$ and $1'\Sigma^{-2}1 = O(n)$.

The second strategy is the ‘tangency’ portfolio defined by $\hat{\alpha}_T = \frac{\Sigma^{-1}(\mu - r)}{1'\Sigma^{-1}(\mu - r)1'}$. The ‘tangency’ portfolio plays an important role in the mean-variance portfolio analysis. Suppose that a portfolio contains a risky asset, then the two-fund separation theorem asserts that one can construct a mean-
variance portfolio by simply allocating the fund into the riskless asset with the riskless return $r$ and the portfolio with none of the riskless asset (or the ‘tangency’ portfolio). One can show, by using the no-arbitrage argument, that the Sharpe Ratio $SR(\hat{\alpha}_T) = (\mu - r1)'\Sigma^{-1}(\mu - r1)$ must always be finite, even when there is an infinite number of risky assets. In this sense the ‘tangency’ portfolio is suboptimal. This portfolio becomes well-diversified if $\|\hat{\alpha}_T\|_2 = (1'\Sigma^{-1}(\mu - r1))^{-1/2} (\mu - r1)'\Sigma^{-2}(\mu - r1) \to 0$, which is valid as long as $1'\Sigma^{-1}(\mu - r1) = O(n)$ and $(\mu - r1)'\Sigma^{-2}(\mu - r1) = O(n)$.

Hitherto, one may still wonder if a closed-form expression for the portfolio with the maximal Sharpe Ratio, or equivalently, with the minimal $sp$ exists as the size of this portfolio increases to infinity. We have shown that the portfolio of this type is indeed inferable and analogous to the mean-variance portfolio. The expression of this portfolio, to be stated in Proposition 1, is given by

$$\tilde{\alpha}(\hat{\theta}_r) = \frac{2}{\hat{\theta}_r^2 n^2} \left(\Sigma + \Sigma'\right)^{-1} (\eta 1 - n\hat{\theta}_r, \mu),$$

where $\eta = \frac{1}{n'\left(\Sigma + \Sigma'\right)^{-1}} \left(\hat{\theta}_r^2 n^2 + n\hat{\theta}_r 1' \left(\Sigma + \Sigma'\right)^{-1} \mu\right)$; and $\hat{\theta}_r$ is defined in Proposition 1. Clearly, this portfolio differs considerably from the gmv portfolio and the ‘tangency’ portfolio. The measure of diversification $\|\tilde{\alpha}(\hat{\theta}_r)\|_2$ converges to zero at some rate, which is determined by the asymptotic joint behavior of $\mu, \Sigma$ and $\hat{\theta}_r$.

It is to be stressed that the above heuristics concerning the multivariate normal returns is only motivating and expositional. Instead the main object of the present study is to analyze the non-normal case. In this general case, minimizing the $sp$ and maximizing the Sharpe Ratio do not often yield the same portfolio; the explanation is that, when returns do not follow the multivariate normal distribution, the c.d.f. $F \left( -\frac{\alpha' \mu - r}{\sqrt{\alpha' \Sigma \alpha}} \right)$ is not a pivotal quantity. In other words, this c.d.f. contains some unknown parameters. It then implies that minimizing the $sp$ may not be congruent with maximizing the Sharpe Ratio. Hereby the main results, to be derived in next sections, concern large-deviations approximations of the minimal $sp$ in the limiting case when the number of risky assets diverges.
3 General Results

Let $\ell(x, \theta)$ denote a nonnegative proper lower semicontinuous function of $\theta$ for every $x$ in a given marginal support of $\ell(x, \cdot)$, say $\mathcal{X}$. Supposing that the function $\ell(x, \theta)$ has the following monotonicity property:

\[
\{ \theta \in \mathbb{R} : \theta > 0 \} = \{ \theta \in \mathbb{R} : \frac{\partial \ell(x, \theta)}{\partial x} > 0 \},
\]

\[
\{ \theta \in \mathbb{R} : \theta < 0 \} = \{ \theta \in \mathbb{R} : \frac{\partial \ell(x, \theta)}{\partial x} < 0 \}.
\]

Hereafter, we shall refer to the function $\ell(x, \theta)$ as a profit function if $\theta$ is positive, and as a loss function if $\theta$ is negative. This profit function becomes a von Neumann - Morgenstern utility function if, given a positive $\theta$, it is concave for every $x \in \mathcal{X}$. It is worth noting that the term “profit function" has specific meaning in microeconomics; and here we have borrowed this term simply because of the natural congruence in the functional form, not because of the microeconomic interpretations.

Before stating the large deviations approximation of the supremum windfall probability, we now specify a set of minimal assumptions needed for our results.

It is important to note at this point that functions that satisfy these assumptions include the exponential functions, which are often used in the formulation of constant absolute risk aversion utility, and the polynomial functions, which are employed to specify constant relative risk aversion utility, and their combinations. In the sequel, we shall provide some specific examples of these functions following the proof of Theorem 1 stated below.

Assumption 2.1 (Convexity). The second-order derivative, $\frac{\partial^2 \ell(x, \theta)}{\partial \theta^2}$, of $\ell(x, \theta)$ is positive on $\{ \theta : \theta \in \mathbb{R}^+ \}$ for every $x \in \mathcal{X}$.

Assumption 2.2 (Pointwise Convergence). $\limsup_{\delta \to 0} \frac{\ell(x, \theta + \lambda \delta)}{\ell(x, \theta \delta)} = 1$ uniformly for every $x \in \mathcal{X}$. That is, $\ell(x, \theta + \lambda) \approx \ell(x, \theta \delta)\ell(x, \lambda \delta)$.

Assumption 2.3 (Existence of Feasible Portfolios). The set $\mathcal{A} \subseteq \mathbb{R}^n$ contains feasible portfolios such that a unique optimal portfolio $\bar{\alpha}$ exists in the interior of $\mathcal{A}$ for all $\theta > 0$.
Assumption 2.4 (Asymptotic Stability). In view of Assumption 2.3, let us define
\[
\ell^*(x, \theta) = \limsup_{n \to \infty} \frac{1}{n} \log \ell(x, n\theta), \quad \forall \ x \in \mathcal{X}.
\]
\[
\Lambda(\theta) = \limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in \mathcal{A}} \log E_P[\ell(S_n(\alpha), n\theta)] = \limsup_{n \to \infty} \frac{1}{n} \log E_P[\ell(S_n(\bar{\alpha}), n\theta)].
\]

Let \( \bar{\theta} = \inf\{\theta > 0 : \Lambda(\theta) = \infty\} \), then, for each \( x \in B_\epsilon(r) \), there exists a sufficiently large \( \theta_r \in (0, \bar{\theta}) \) such that \( \ell^*(x, \theta) = \ell^*(x, \theta_r) \theta + r(\epsilon, \theta_r) \) for all \( \theta \in [\theta_r, \bar{\theta}] \), where \( \ell^*(x, \theta_r) = \left. \frac{\partial \ell^*(x, \theta)}{\partial \theta} \right|_{\theta = \theta_r} \). And the remainder \( r(\epsilon, \theta_r) \) is infinitesimal (i.e., \( \lim_{\theta_r \to \infty, \epsilon \to 0} r(\epsilon, \theta_r) = 0 \)). [It is worth mentioning that, as we shall see below, this Taylor approximation allows us to apply the tools of Convex Analysis to the rate function.]

Remark 3.1. We now need to discuss on the validity of the conditions stated in Assumptions 2.1 - 2.4. This is unfortunately a delicate matter, because a few is known on this subject. Assumption 2.1 asserts that the marginal rate of profit increases proportionally with \( \theta \), namely the coefficient of profit desirability. Thus, the demand for profit increases without bound, as \( \theta \) diverges to infinity.

Assumption 2.2 states that a profit function, \( \ell(x, \theta) \), has an asymptotic multiplicatively-separable (ms) representation. Examples of this type of function include, but not necessarily limited to, the power, exponential, and power-exponential functions. It is worth emphasizing that these functions have exact ms representations. Assumption 2.3 is a trivial and mild condition permitting the existence of a set of feasible portfolios.

Assumption 2.4 postulates that both the logarithm of a profit function and that of the expected profit tend to measurable functions in the limit. And this limiting profit function must be asymptotically linear in the coefficient of profit desirability, \( \theta \), for \( \theta \) approaching the boundary of its relevant domain. This turns out to be true in the light of Assumption 2.1, which essentially allows the demand for profit to increase commensurately with \( \theta \).

By Assumption 2.1, the function \( \Lambda(\theta) \) must be a first-order differentiable convex function. We
define the following Fenchel-Legendre transform of $\Lambda(\theta)$:

$$\Lambda^*(s) = \sup_{\theta \in [\theta_r, \theta]} [\theta \ell''(s, \theta_r) - \Lambda(\theta)].$$

By the definition of the Fenchel-Legendre transform, the rate function $\Lambda^*(s)$ is a convex function. Following [den Hollander [2000]], it is not hard to prove that

$$\Lambda^*(s) = \begin{cases} 
\hat{\theta}_s \ell''(s, \theta_r) - \Lambda(\hat{\theta}_s) & \text{if } s \in \{s : \Lambda'(\theta_r) \leq \ell''(s, \theta_r) < \Lambda'(\bar{\theta})\}, \\
0 & \text{if } s \in \{s : \Lambda'(\theta_r) > \ell''(s, \theta_r)\},
\end{cases} \quad (3.1)$$

where $\hat{\theta}_s = \theta(s, \theta_r)$ is a unique solution [assuming that it always exists] to $\ell''(s, \theta_r) = \Lambda'(\theta)$, where $\Lambda'(\theta) = \frac{\partial \Lambda(\theta)}{\partial \theta}$, such that $\hat{\theta}_s \in [\theta_r, \bar{\theta}]$. (Note that the point $\hat{\theta}_s$ is also called as an exposing point to the hyperplane $\{\theta : \ell''(s, \theta_r) = \Lambda'(\theta)\}$ associated with the half-spaces $\{\theta : \ell''(s, \theta_r) > \Lambda'(\theta)\}$ and $\{\theta : \ell''(s, \theta_r) < \Lambda'(\theta)\}$.)

We need the following lemma for further developments.

**Lemma 1.** The following inequality:

$$\ell''(r, \theta_r)\hat{\theta}_r - \Lambda^*(r) > \ell''(s, \theta_r)\hat{\theta}_r - \Lambda^*(s),$$

holds for every $s \in B_\varepsilon(r)$.

**Proof.** Since $\ell''(s, \theta_r)\hat{\theta}_r - \Lambda(\hat{\theta}_r) < \Lambda^*(s)$, we have

$$\ell''(s, \theta_r)\hat{\theta}_r - \Lambda(\hat{\theta}_r) < \ell''(s, \theta_r)\hat{\theta}_s - \Lambda(\hat{\theta}_s)$$

$$\Leftrightarrow \Lambda(\hat{\theta}_r) > \Lambda(\hat{\theta}_s) + \ell''(s, \theta_r)\hat{\theta}_r - \ell''(s, \theta_r)\hat{\theta}_s$$

$$\Leftrightarrow \ell''(r, \theta_r)\hat{\theta}_r - \left(\ell''(r, \theta_r)\hat{\theta}_r - \Lambda(\hat{\theta}_r)\right) > \ell''(s, \theta_r)\hat{\theta}_r - \Lambda^*(s)$$

$$\Leftrightarrow \ell''(r, \theta_r)\hat{\theta}_r - \Lambda^*(r) > \ell''(s, \theta_r)\hat{\theta}_r - \Lambda^*(s).$$
The first theorem of this paper is stated below.

**Theorem 1.** If Assumptions 2.1-2.4 hold, then the supremum windfall probability can be approximated as follows:

\[
\limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) = -\Lambda^*(r) - \tau(0, \theta_r).
\]  

(3.2)

Theorem 1 postulates that an investor who holds a sufficiently large number of assets can always find an optimal combination of these assets by maximizing an expected endogenous profit function so as to ensure that the windfall probability is maximum. Moreover, although investors may have different optimal portfolios associated with their own profit functions, they can always achieve the same maximum windfall probability by increasing the size of their portfolio. This is a vindication of the diversification tenet that diversification will eliminate all the idiosyncratic risks, and everyone will benefit from diversification. As such, the only remaining risk is the market [nondiversifiable] risk, which every investor has to share. That is, whatever the choice of profit function, it is essential that the individual investors’ optimal portfolios guarantee the same windfall probability.

**Proof of Theorem 1.** We shall prove the following bounds:

- **Upper Bound:** Under Assumption 2.1, the Chebyshev inequality yields

\[
P(S_n^{(\alpha)} \geq r) \leq \frac{E_p[\ell(S_n^{(\alpha)}, n\theta)]}{\ell(r, n\theta)}
\]

\[\Leftrightarrow \log P(S_n^{(\alpha)} \geq r) \leq \log E_p[\ell(S_n^{(\alpha)}, n\theta)] - \log \ell(r, n\theta)
\]

\[\Leftrightarrow \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) \leq \sup_{\alpha \in A} \log E_p[\ell(S_n^{(\alpha)}, n\theta)] - \log \ell(r, n\theta)
\]

\[\Leftrightarrow \limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) \leq \limsup_{n \to \infty} \frac{1}{n} \log \sup_{\alpha \in A} E_p[\ell(S_n^{(\alpha)}, n\theta)] - \limsup_{n \to \infty} \frac{1}{n} \log \ell(r, n\theta),
\]

where the optimal portfolio, \( \tilde{\alpha} \), exists in view of Assumption 2.3. Moreover, by Assumption 2.4, we have

\[
\limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) \leq \Lambda(\theta) - \ell^*(r, \theta).
\]
Hence, it follows that
\[
\limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) \leq -\sup_{\theta \in \Theta} \left[ \theta \ell'(r, \theta_r) - \Lambda(\theta) \right] - v(0, \theta_r) = \Lambda^*(r) - v(0, \theta_r). \tag{3.3}
\]

**Lower Bound:** Given a benchmark return, \( r \), satisfying the constraint \( \Lambda'(\theta_r) \leq \ell'(r, \theta_r) < \Lambda'(\overline{\theta}) \), let us define a conjugate joint probability measure, \( \mathbb{P}^* \), on the product probability space \( (\Omega, \mathcal{F}_i, \mathbb{P}) \), where \( (\Omega, \mathcal{F}_i) \) denotes the sample space of \( X_i \), as follows:
\[
\frac{d\mathbb{P}^*}{d\mathbb{P}} = \frac{\ell(S_n^{(\hat{\alpha})}, n\hat{\theta}_r)}{E_\mathbb{P}[\ell(S_n^{(\hat{\alpha})}, n\hat{\theta}_r)]}. \tag{3.4}
\]

Then, choosing \( r_n \) such that \( s \in B_\epsilon(r_n) \) implies \( s \in \{ s : \Lambda'(\theta_r) \leq \ell'(s, \theta_r) < \Lambda'(\overline{\theta}) \} \), we have
\[
P(S_n^{(\hat{\alpha})} > r_n - \epsilon) = \int 1(S_n^{(\hat{\alpha})} > r_n - \epsilon) \frac{E_\mathbb{P}[\ell(S_n^{(\hat{\alpha})}, n\hat{\theta}_r)]}{\ell(S_n^{(\hat{\alpha})}, n\hat{\theta}_r)} d\mathbb{P}^*
\]
\[
\geq \frac{E_\mathbb{P}[\ell(S_n^{(\hat{\alpha})}, n\hat{\theta}_r)]}{\ell(r_n + \epsilon, n\hat{\theta}_r)} \int 1(S_n^{(\hat{\alpha})} \in (r_n - \epsilon, r_n + \epsilon)) d\mathbb{P}^*,
\]
thus
\[
\log P(S_n^{(\hat{\alpha})} > r_n - \epsilon) \geq -\left\{ \log \ell(r_n + \epsilon, n\hat{\theta}_r) - \log E_\mathbb{P}[\ell(S_n^{(\hat{\alpha})}, n\hat{\theta}_r)] \right\} + \log P^*(S_n^{(\hat{\alpha})} \in (r_n - \epsilon, r_n + \epsilon)).
\]

This gives
\[
\limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) \geq \liminf_{n \to \infty} \frac{1}{n} \log P(S_n^{(\hat{\alpha})} > r_n - \epsilon)
\]
\[
\geq -\left\{ \limsup_{n \to \infty} \frac{1}{n} \log \ell(r_n + \epsilon, n\hat{\theta}_r) - \limsup_{n \to \infty} \frac{1}{n} \log E_\mathbb{P}[\ell(S_n^{(\hat{\alpha})}, n\hat{\theta}_r)] \right\}
\]
\[
+ \liminf_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \in (r_n - \epsilon, r_n + \epsilon)).
\]
Hence, we have

\[
\limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) \geq -\ell^*(r + \epsilon, \hat{\theta}_r) - \Lambda(\hat{\theta}_r) + \liminf_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \in (r_n - \epsilon, r_n + \epsilon)).
\]

(3.5)

Now, in order to prove the lower bound we shall prove that

\[
\liminf_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \in (r_n - \epsilon, r_n + \epsilon)) = 0,
\]

which is equivalent to

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \notin (r_n - \epsilon, r_n + \epsilon)) < 0.
\]

(3.6)

Since \( P^*(S_n^{(\hat{\alpha})} \notin (r_n - \epsilon, r_n + \epsilon)) \leq (P^*(S_n^{(\hat{\alpha})} \geq r_n + \epsilon), P^*(S_n^{(\hat{\alpha})} \leq r_n - \epsilon))^+ \), in order to prove Eq. (3.6), we need to show that

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \geq r_n + \epsilon) < 0,
\]

(3.7)

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \leq r_n - \epsilon) < 0.
\]

(3.8)

We first prove Eq. (3.7). In view of Assumption 2.1, an application of the Tchebyshev in equality yields, for a \( \lambda \in (0, \overline{\theta} - \hat{\theta}_r) \),

\[
P^*(S_n^{(\hat{\alpha})} \geq r_n + \epsilon) \leq \frac{E_{P^*}[\ell(S_n^{(\hat{\alpha})}, n\lambda)]}{\ell(r_n + \epsilon, n\lambda)}
\]

\[
\Leftrightarrow \limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \geq r_n + \epsilon) \leq \frac{1}{n} \log \ell(r_n + \epsilon, n\lambda) - \limsup_{n \to \infty} \frac{1}{n} \log E_{P^*}[\ell(S_n^{(\hat{\alpha})}, n\lambda)]
\]

\[
= -[\ell^*(r + \epsilon, \lambda) - \tilde{\Lambda}(\lambda)],
\]

(3.9)

where, in view of Assumption 2.4, \( \ell^*(r + \epsilon, \lambda) = \limsup_{n \to \infty} \frac{1}{n} \log \ell(r_n + \epsilon, n\lambda) \) and \( \tilde{\Lambda}(\lambda) = \).
\[
\limsup_{n \to \infty} \frac{1}{n} \log E_x^*[\ell(S_n^{(\hat{\alpha})}, n\lambda)]. \text{ By Assumption 2.2, setting } \delta = 1/n, \text{ we obtain }
\]

\[
\bar{\Lambda}(\lambda) = \limsup_{n \to \infty} \frac{1}{n} \log E_x [\ell(S_n^{(\hat{\alpha})}, n\lambda)\ell(S_n^{(\hat{\alpha})}, n\hat{\theta}_r)] - \limsup_{n \to \infty} \frac{1}{n} \log E_x [\ell(S_n^{(\hat{\alpha})}, n\hat{\theta}_r)] = \Lambda(\lambda + \hat{\theta}_r) - \Lambda(\hat{\theta}_r).
\]

Moreover, one can immediately verify that Assumption 2.2 implies that \(\ell^*(x, \theta) = -\ell^*(x, -\theta)\) and \(\ell^*(x, \theta + \lambda) = \ell^*(x, \theta) + \ell^*(\lambda)\) for all \(x \in \mathcal{X}\). Thus, Eq. (3.9) can be rewritten as

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \geq r_n + \epsilon) \leq -[\ell^*(r + \epsilon, \lambda + \hat{\theta}_r) - \Lambda(\lambda + \hat{\theta}_r)] + [\ell^*(r + \epsilon, \hat{\theta}_r) - \Lambda(\hat{\theta}_r)].
\]

By Assumption 2.4, we have

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \geq r_n + \epsilon) \leq -[\ell'(r + \epsilon, \theta_r)(\lambda + \hat{\theta}_r) - \Lambda(\lambda + \hat{\theta}_r)] - \varphi(\epsilon, \theta_r) + [\ell'(r, \hat{\theta}_r) - \Lambda(\hat{\theta}_r)] - \ell'(r, \hat{\theta}_r)\hat{\theta}_r + \ell'(r + \epsilon, \hat{\theta}_r)\hat{\theta}_r + \varphi(\epsilon, \theta_r).
\]

Hence, it follows that

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \geq r_n + \epsilon) \leq -\sup_{\xi \in \theta(\bar{\theta}) \in \bar{\theta}} [\ell'(r + \epsilon, \theta_r)\xi - \Lambda(\xi)] + \Lambda^*(\lambda - \ell^*(r, \hat{\theta}_r)\hat{\theta}_r
\]

\[
+ \ell^*(r + \epsilon, \hat{\theta}_r)\hat{\theta}_r
\]

\[
= -\Lambda^*(r + \epsilon) + \Lambda^*(\lambda) - \ell^*(r, \hat{\theta}_r)\hat{\theta}_r + \ell^*(r + \epsilon, \hat{\theta}_r)\hat{\theta}_r < 0,
\]

where the last inequality follows from Lemma [3] by setting \(s = r + \epsilon\). More importantly, note that \(\Lambda^*(r + \epsilon) = \sup_{\xi \in \theta(\hat{\theta})} [\ell'(r + \epsilon, \theta_r)\xi - \Lambda(\xi)]\) because the exposing point \(\xi(r)\) is a nondecreasing function of \(r\). Therefore, Eq. (3.7) has been proved.

We now proceed to the proof of Eq. (3.8). In view of Assumption 2.1, an application of the
Tchebyshov inequality yields, for \( \lambda \in [\theta_r - \hat{\theta}_r, 0) \),

\[
P^*(S_n^{(\hat{\alpha})} \leq r_n - \epsilon) = P^*(\ell(S_n^{(\hat{\alpha})}, \lambda) \geq \ell(r_n - \epsilon, \lambda)) \\
\leq E_{\hat{\pi}^*}[\ell(S_n^{(\hat{\alpha})}, n\lambda)]/\ell(r_n - \epsilon, n\lambda).
\]

By Assumption 2.4, we have

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \leq r_n - \epsilon) \leq - \left[ \limsup_{n \to \infty} \frac{1}{n} \log \ell(r_n - \epsilon, n\lambda) - \lim_{n \to \infty} \frac{1}{n} \log E_{\hat{\pi}^*}[\ell(S_n^{(\hat{\alpha})}, n\lambda)] \right] \\
= [\ell^*(r - \epsilon, \lambda) - \tilde{\Lambda}(\lambda)],
\]

where \( \ell^*(r - \epsilon, \lambda) = \limsup_{n \to \infty} \frac{1}{n} \log \ell(r_n - \epsilon, n\lambda) \) and \( \tilde{\Lambda}(\lambda) = \lim_{n \to \infty} \frac{1}{n} \log E_{\hat{\pi}^*}[\ell(S_n^{(\hat{\alpha})}, n\lambda)] \).

Similarly, by Assumption 2.2, setting \( \delta = 1/n \), we obtain

\[
\tilde{\Lambda}(\lambda) = \Lambda(\lambda + \hat{\theta}_r) - \Lambda(\hat{\theta}_r),
\]

thus, Eq. (3.10) can be rewritten as

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \leq r_n - \epsilon) \leq - \left[ \ell^*(r - \epsilon, \lambda + \hat{\theta}_r) - \Lambda(\lambda + \hat{\theta}_r) \right] + [\ell^*(r - \epsilon, \hat{\theta}_r) - \Lambda(\hat{\theta}_r)] \\
= -\left[ \ell^*(r - \epsilon, \theta_r)(\lambda + \hat{\theta}_r) - \Lambda(\lambda + \hat{\theta}_r) \right] + [\ell^*(r, \theta_r)\hat{\theta}_r - \Lambda(\hat{\theta}_r)] \\
+ \ell^*(r - \epsilon, \hat{\theta}_r)\hat{\theta}_r - \ell^*(r, \hat{\theta}_r)\hat{\theta}_r,
\]

where the last equality follows from Assumption 2.4. Therefore, we obtain

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\hat{\alpha})} \leq r_n - \epsilon) \leq - \sup_{\xi \in [\theta_r, \hat{\theta}_r]} \left[ \ell^*(r - \epsilon, \theta_r)\xi - \Lambda(\xi) \right] + \Lambda^*(r) + \ell^*(r - \epsilon, \hat{\theta}_r)\hat{\theta}_r \\
- \ell^*(r, \hat{\theta}_r)\hat{\theta}_r \\
= -\Lambda^*(r - \epsilon) + \Lambda^*(r) + \ell^*(r - \epsilon, \hat{\theta}_r)\hat{\theta}_r - \ell^*(r, \hat{\theta}_r)\hat{\theta}_r < 0,
\]

where the last inequality follows from Lemma \( \square \) by setting \( s = r - \epsilon \). Eq. (3.8) has been
Returning to Eq. (3.3), let $\epsilon$ approach 0, in view of Assumption 2.4 we get the following lower bound:
\[
\limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) \geq -[\ell'(r, \theta_r)\hat{\theta}_r - \Lambda(\hat{\theta}_r)] - r(0, \theta_r).
\]
(3.11)

The theorem follows from Eqs (3.3) and (3.11). QED.

**Examples** of functions that satisfy Assumptions 2.1-2.4 are as follows:

- the exponential function $e^{\theta x}$: $X = \mathbb{R}$, $\theta_* = 0$, $\ell^*(x, \theta) = \theta x$, where $\theta \in (0, \bar{\theta})$, and $r(\epsilon, \theta_r) = 0$
- the power function $x^{\theta}$: $X = \mathbb{R}^+$, $\theta_* = 0$, $\ell^*(x, \theta) = \theta \log(x)$, and $r(\epsilon, \theta_r) = 0$
- the power exponential function $x^{\theta}e^{\theta x}$: $X = \mathbb{R}^+$, $\theta_* = 0$, $\ell^*(x, \theta) = \theta(\log(x) + x)$, and $r(\epsilon, \theta_r) = 0$
- $(x^2 - 2x + 2)^{\theta}$: $X = \mathbb{R}$, $\theta_* = 0$, $\ell^*(x, \theta) = \theta \log(x^2 - 2x + 2)$, and $r(\epsilon, \theta_r) = 0$
- $e^{\theta x} + \frac{e^x}{x}$: $X = \mathbb{R}$, $\theta_* = 0$, $\ell^*(x, \theta) = \limsup_{n \to \infty} \frac{1}{n} \log \left( e^{n\theta x} + \frac{e^{x}}{(n\theta)^2} \right) = \theta x$, and $r(\epsilon, \theta_r) = 0$

Assumption 2.2 can immediately be verified by noting that $\lim_{\delta \to 0} e^{\theta \delta} = 0$, $\forall x \in X$.

Now, in order to state the large deviations approximation of the infimum *shortfall* probability, we shall specify a set of minimal assumptions which are, to a certain extent, analogous to Assumptions 2.1 - 2.3 stated above, thus need no further interpretation.

**Assumption 2.1’ (Convexity).** The second-order derivative, $\frac{\partial^2 \ell(x, \theta)}{\partial \theta^2}$, of $\ell(x, \theta)$ is positive on $\{\theta : \theta \in \mathbb{R}^-\}$ for every $x \in X$.

**Assumption 2.3’ (Existence of Feasible Portfolios).** The set $A \subseteq \mathbb{R}^n$ contains feasible portfolios such that a *unique* optimal portfolio $\tilde{\alpha}$ exists in the interior of $A$ for all $\theta < 0$.

**Assumption 2.4’ (Asymptotic Stability).** In view of Assumption 2.3, let us define
\[
\ell^*(x, \theta) = \liminf_{n \to \infty} \frac{1}{n} \log \ell(x, n\theta), \ \forall x \in X.
\]
\[
\Lambda(\theta) = \liminf_{n \to \infty} \frac{1}{n} \inf_{\alpha \in A} \log E_{\mathbb{P}}[\ell(S_n^{\alpha}, n\theta)] = \liminf_{n \to \infty} \frac{1}{n} \log E_{\mathbb{P}}[\ell(S_n^{\tilde{\alpha}}, n\theta)].
\]
Let $\bar{\theta} = \inf\{\theta < 0 : \Lambda(\theta) = -\infty\}$, then, for each $x \in B_r(r)$, there exists a sufficiently small $\theta_r \in (\bar{\theta}, 0)$ such that $\ell^*(x, \theta) = \ell^*(x, \theta_r) + \epsilon(\epsilon, \theta_r)$ for all $\theta \in (\bar{\theta}, \theta_r]$, where $\ell^*(x, \theta_r) = \frac{\partial \ell^*(x, \theta)}{\partial \theta}|_{\theta = \theta_r}$. And the remainder $\epsilon(\epsilon, \theta_r)$ is infinitesimal (i.e., $\lim_{\theta_r \rightarrow \infty, \epsilon \downarrow 0} |\epsilon(\epsilon, \theta_r)| = 0$).

We define the following Fenchel-Legendre transform of $\Lambda(\theta)$:

$$\Lambda^*(s) = \sup_{\theta \in (\bar{\theta}, \theta_r]} [\theta \ell^*(s, \theta_r) - \Lambda(\theta)].$$

By the definition of the Fenchel-Legendre transform, the rate function $\Lambda^*(s)$ is a convex function.

Following Dembo and Zeitouni [1998, Lemma 2.3.9], it is not hard to prove that

$$\Lambda^*(s) = \begin{cases} \hat{\theta}_s \ell^*(s, \theta_r) - \Lambda(\hat{\theta}_s) & \text{if } s \in \{s : \Lambda'(\theta) < \ell^*(s, \theta_r) \leq \Lambda'(\theta_r)\}, \\ 0 & \text{if } s \in \{s : \Lambda'(\theta_r) < \ell^*(s, \theta_r)\}, \end{cases}$$

(3.12)

where $\hat{\theta}_s = \theta(s, \theta_r)$ is a unique solution [assuming that it always exists] to $\ell^*(s, \theta_r) = \Lambda'(\theta)$, where $\Lambda'(\theta) = \frac{\partial \Lambda(\theta)}{\partial \theta}$, such that $\hat{\theta}_s \in (\bar{\theta}, \theta_r]$. (Note that the point $\hat{\theta}_s$ is also called as an exposing point to the hyperplane $\{\theta : \ell^*(s, \theta_r) = \Lambda'(\theta)\}$ associated with the half-spaces $\{\theta : \ell^*(s, \theta_r) > \Lambda'(\theta)\}$ and $\{\theta : \ell^*(s, \theta_r) < \Lambda'(\theta)\}$.)

The second theorem of this paper is stated below.

**Theorem 2.** If Assumptions 2.1’, 2.2, 2.3’, and 2.4’ hold, the the infimum shortfall probability can be approximated as follows:

$$\lim\inf_{n \rightarrow \infty} \frac{1}{n} \inf_{\alpha \in \mathcal{A}} \log P(S_n^{(\alpha)} \leq r) = -\Lambda^*(r) - \epsilon(0, \theta_r).$$

(3.13)

Theorem 2 states that an investor who holds a sufficiently large number of risky assets can always find an optimal combination of these assets by minimizing an expected endogenous loss function so as to ensure that the shortfall probability is minimum. Furthermore, an average investor can do best by simply diversifying among all classes of risky assets and then minimizing this investor’s expected
endogenous loss function. Hence, diversification apparently benefits most investors endowed with certain endogenous convex loss functions.

**Proof of Theorem 2.** Since this theorem is dual to Theorem 1, the proof can be done with the same method used to prove the latter. QED.

4 Specific Results

If $\ell(x, \theta)$ is an exponential function (i.e., $\ell(x, \theta) = e^{\theta x}$), then Assumptions 2.1, 2.2, and 2.1' are obviously satisfied; Assumption 2.3 implies Assumption 2.4 with $\tau(\epsilon, \theta_r) = 0$ (i.e., the Taylor approximation is exact); and Assumption 2.3' implies Assumption 2.4' with $\tau(\epsilon, \theta_r) = 0$. The large deviations estimation of the optimal windfall and shortfall probabilities can be derived in the same spirit as the Gärtner-Ellis theorem (see, e.g., ten Hollander [2000, Chap. 5]). We now state the following corollary:

**Corollary 1.** If Assumption 2.3 holds, then the supremum windfall probability can be approximated as follows:

$$\limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} P(S_n^{(\alpha)} \geq r) = -\Lambda^*(r), \quad (4.1)$$

where $\Lambda^*(r) = \sup_{\theta \in (0, \bar{\theta})} [r \theta - \Lambda(\theta)]$ with $\Lambda(\theta) = \limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log E_{\mathbb{P}}[e^{n \theta S_n^{(\alpha)}}]$, for $r \in \mathbb{R} : \Lambda'(0) < r < \Lambda'(\bar{\theta})$.

Furthermore, if Assumption 2.3' holds, then the infimum shortfall probability can be approximated as follows:

$$\liminf_{n \to \infty} \frac{1}{n} \inf_{\alpha \in A} P(S_n^{(\alpha)} \leq r) = -\Lambda^*(r), \quad (4.2)$$

where $\Lambda^*(r) = \sup_{\theta \in (\theta, 0)} [r \theta - \Lambda(\theta)]$ with $\Lambda(\theta) = \liminf_{n \to \infty} \frac{1}{n} \inf_{\alpha \in A} \log E_{\mathbb{P}}[e^{n \theta S_n^{(\alpha)}}]$, for $r \in \mathbb{R} : \Lambda'(\theta) < r < \Lambda'(0)$.

Although the proof of the above corollary implicitly follows from that of Theorem 1, but for the sake of a lucid exposition of the large deviations theory, we shall provide the proof of Eq. (4.1), which is analogous to the proof of Theorem 1. The proof of Eq. (4.2) can be done in the same way.
Proof of Corollary [41]. We shall prove the upper and lower bounds.

- **Upper Bound**: The Tchebyshev inequality gives

  \[ P(S_n^{(\alpha)} \geq r) = E_{\mathbb{P}}[1(S_n^{(\alpha)} \geq r)] \leq \frac{E_{\mathbb{P}}[e^{\theta S_n^{(\alpha)}}]}{e^{\theta r}}, \quad \forall \theta \in (0, \theta). \]

  Linearizing the RHS of the above inequality, by Assumption 2.3 we obtain

  \[ \frac{1}{n} \sup_{\alpha \in \mathcal{A}} \log P(S_n^{(\alpha)} \geq r) \leq - \left\{ \theta r - \frac{1}{n} \sup_{\alpha \in \mathcal{A}} \log E_{\mathbb{P}}[e^{\theta S_n^{(\alpha)}}] \right\} \]

  \[ \Leftrightarrow \limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in \mathcal{A}} \log P(S_n^{(\alpha)} \geq r) \leq - \left\{ \theta r - \log E_{\mathbb{P}}[e^{\theta S_n^{(\alpha)}}] \right\}. \]

  It follows that

  \[ \limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in \mathcal{A}} \log P(S_n^{(\alpha)} \geq r) \leq - \sup_{\theta \in (0, \theta)} [\theta r - \Lambda(\theta)] \]

  \[ = -\Lambda^*(r), \quad (4.3) \]

  where \( \Lambda(\theta) = \limsup_{n \to \infty} \frac{1}{n} \log E_{\mathbb{P}}[e^{\theta S_n^{(\alpha)}}] \) and \( r \in \{ r \in \mathbb{R} : \Lambda'(0) < r < \Lambda'(\theta) \} \).

- **Lower Bound**: Let us define a conjugate joint probability measure, \( \mathbb{P}^* \), on the product probability space \((\Omega, \otimes_{i=1}^n \mathcal{F}_i, \mathbb{P})\) as follows:

  \[ \frac{d\mathbb{P}^*}{d\mathbb{P}} = e^{\hat{\theta}_r S_n^{(\alpha)} - \log E_{\mathbb{P}}[e^{\hat{\theta}_r S_n^{(\alpha)}}]}, \]

  where \( \hat{\theta}_r \) is a unique exposing point [assuming that it exists] to the hyperplane \( \{ \theta > 0 : \Lambda'(\theta) = r \} \).

  Pick up a sequence, \( r_n \), such that \( \lim_{n \to \infty} r_n = r \) and \( [r_n - \epsilon, r_n + \epsilon] \in \{ r \in \mathbb{R} : \Lambda'(0) < r < \} \)
Λ′(\tilde{\theta})\}, we obtain

\[
P(S_n^{\tilde{\alpha}} > r_n - \epsilon) = \int 1(S_n^{\tilde{\alpha}} > r_n - \epsilon) e^{\log E_{P}|e^{\hat{\theta} r S_n^{\tilde{\alpha}}}| - n\hat{\theta}_r S_n^{\tilde{\alpha}}} \, dP^* \\
\geq \int 1(S_n^{\tilde{\alpha}} \in (r_n - \epsilon, r_n + \epsilon)) e^{\log E_{P}|e^{\hat{\theta} r S_n^{\tilde{\alpha}}}| - n\hat{\theta}_r S_n^{\tilde{\alpha}}} \, dP^* \\
\geq e^{\log E_{P}|e^{\hat{\theta} r S_n^{\tilde{\alpha}}}| - n\hat{\theta}_r (r_n + \epsilon) P^*(S_n^{\tilde{\alpha}} \in (r_n - \epsilon, r_n + \epsilon)).
\]

Hence, it is straightforward to show that

\[
\limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{\alpha} \geq r) \geq \liminf_{n \to \infty} \frac{1}{n} \log P(S_n^{\tilde{\alpha}} > r_n - \epsilon) \\
\geq - (\hat{\theta}_r (r + \epsilon) - \Lambda(\hat{\theta}_r)) \\
+ \liminf_{n \to \infty} \frac{1}{n} \log P^*(S_n^{\tilde{\alpha}} \in (r_n - \epsilon, r_n + \epsilon)) \quad (4.4)
\]

In view of Eq. (4.4), in order to show the lower bound, we need to prove that

\[
\liminf_{n \to \infty} \frac{1}{n} \log P^*(S_n^{\tilde{\alpha}} \in (r_n - \epsilon, r_n + \epsilon)) = 0.
\]

This is equivalent to show

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{\tilde{\alpha}} \notin (r_n - \epsilon, r_n + \epsilon)) < 0.
\]

Hence, it is necessary to show that

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{\tilde{\alpha}} \geq r_n + \epsilon) < 0 \quad (4.5)
\]

and

\[
\limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{\tilde{\alpha}} \leq r_n - \epsilon) < 0. \quad (4.6)
\]
To prove Eq. (4.5), an application of the Tchebyshev inequality gives

\[ P^*(S_n^{(\alpha)} \geq r_n + \epsilon) \leq e^{-n[\lambda(r_n + \epsilon) - \log E_{\bar{P}^*}[e^{n\lambda S_n^{(\alpha)}}]]} \]

\[ \iff \limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\alpha)} \geq r_n + \epsilon) \leq -\{(r_n + \epsilon)\lambda - \limsup_{n \to \infty} \frac{1}{n} \log E_{\bar{P}^*}[e^{n\lambda S_n^{(\alpha)}}]\} \]

where \( \Lambda(\lambda) = \limsup_{n \to \infty} \frac{1}{n} \log E_{\bar{P}^*}[e^{n\lambda S_n^{(\alpha)}}] \) for every \( \lambda \in (0, \theta - \hat{\theta}) \).

Moreover, since \( \tilde{\Lambda}(\lambda) = \Lambda(\lambda + \hat{\theta}_r) - \Lambda(\hat{\theta}_r) \) we obtain

\[ \limsup_{n \to \infty} \frac{1}{n} \log P^*(S_n^{(\alpha)} \geq r_n + \epsilon) \leq -\{(r_n + \epsilon)\lambda - \tilde{\Lambda}(\lambda)\} - \Lambda(\lambda + \hat{\theta}_r) + \hat{\theta}_r \epsilon \]

where the last inequality follows from Lemma 1. Hence, we have proved Eq. (4.5). And Eq. (4.6) can be proved in the same way.

In view of Eq. (4.4), we have

\[ \limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) \geq -(\hat{\theta}_r(r + \epsilon) - \Lambda(\hat{\theta}_r)). \]

By letting \( \epsilon \) go to 0, we obtain the following lower bound:

\[ \limsup_{n \to \infty} \frac{1}{n} \sup_{\alpha \in A} \log P(S_n^{(\alpha)} \geq r) \geq -\Lambda^*(r). \] (4.7)

Eqs (4.3) and (4.7) implies Eq. (4.1). QED.

As an application of Corollary 1, we state the following proposition:
**Proposition 1.** Suppose that \{X_1, \ldots, X_n\}' is a multivariate normal random vector, \(N(\mu', \Sigma)\), we have

\[
\liminf \frac{1}{n} \inf_{\alpha \in \mathbb{R}^n \text{ s.t. } 1'\alpha = 1} \log P(S_n^{(\alpha)} \leq r) = -\Lambda^*(r),
\]

where

\[
\Lambda^*(r) = r\hat{\theta}_r - \Lambda(\hat{\theta}_r) \text{ for some } r \in (-\infty, \Lambda'(0)),
\]

\[
\Lambda(\theta) = 2 \left\{ \frac{1}{4} \theta^2 \frac{C}{B^2} + \frac{1}{2} \theta \left( \frac{A}{B} + 2 \frac{AC}{B^2} - 2 \frac{D}{B} \right) + \frac{A^2}{B} + \frac{A^2C}{B^2} - 2 \frac{AD}{B} + F - E \right\},
\]

\[
\hat{\theta}_r = \frac{B^2}{C} \left[ r - \left( \frac{A}{B} + 2 \frac{AC}{B^2} - 2 \frac{D}{B} \right) \right],
\]

and

\[
\frac{\mu' (\Sigma + \Sigma')^{-1} \mathbf{1}}{n} \quad = \quad A,
\]

\[
\frac{1'(\Sigma + \Sigma')^{-1} \mathbf{1}}{n} \quad = \quad B,
\]

\[
\frac{1'(\Sigma + \Sigma')^{-1} \Sigma(\Sigma + \Sigma')^{-1} \mathbf{1}}{n} \quad = \quad C,
\]

\[
\frac{1'(\Sigma + \Sigma')^{-1} \Sigma(\Sigma + \Sigma')^{-1} \mu}{n} \quad = \quad D,
\]

\[
\frac{\mu' (\Sigma + \Sigma')^{-1} \mathbf{1}}{n} \quad = \quad E,
\]

\[
\frac{\mu' (\Sigma + \Sigma')^{-1} \Sigma(\Sigma + \Sigma')^{-1} \mu}{n} \quad = \quad F.
\]

**Proof of Proposition 1.** Note that the set of feasible portfolios is \(\mathcal{A} = \{\alpha \in \mathbb{R}^n : 1'\alpha = 1\}\) and

\[
\log E[e^{n\theta S_n^{(\alpha)}}] = n\theta \left( \alpha' \mu + \frac{1}{2} n\theta \alpha' \Sigma \alpha \right).
\]

Since \(\frac{\partial^2 \log E[e^{n\theta S_n^{(\alpha)}}]}{\partial \alpha \partial \alpha} = \frac{n^2 \theta^2}{2} (\Sigma + \Sigma') > 0\), Assumption 2.3' is true. First, an application of Eq. (4.2)
in Corollary \[\Box\] and some tedious algebra manipulations yield the following optimal portfolio vector:

\[
\tilde{\alpha}(\theta) = \frac{2}{\theta^2 n^2} (\Sigma + \Sigma')^{-1}(\eta 1 - n\theta \mu),
\]

where

\[
\eta = \frac{1}{\theta^2 n^2} \left( \frac{\theta^2 n^2}{2} + n\theta 1'(\Sigma + \Sigma')^{-1}\mu \right).
\]

Next, we can derive the rate function \( \Lambda^*(r) \) by initially substituting \( \tilde{\alpha}(\theta) \) into

\[
\Lambda(\theta) = \limsup_{n \to \infty} \frac{1}{n} \log E[e^{n\theta S_n(\tilde{\alpha})}].
\]

Unfortunately, the only known result is the following cumbersome formula:

\[
\Lambda(\theta) = \frac{2}{\theta^2} \left( \frac{1}{n^2} \sum_{\omega} \left[ \frac{1}{2} \eta \left( \frac{1'}{1' + \theta \mu} \right)^2 + \frac{1}{2} n\theta \left( \frac{1'}{1' + \theta \mu} \right)^2 \right] \right)
\]

\[
+ \frac{1'}{1'} \left( \frac{1}{1'} + \theta \mu \right) \left[ \frac{1}{1'} + \theta \mu \right] - \frac{1'}{1'} \left( \frac{1}{1'} + \theta \mu \right) \left[ \frac{1}{1'} + \theta \mu \right] - \frac{1'}{1'} \left( \frac{1}{1'} + \theta \mu \right) \left[ \frac{1}{1'} + \theta \mu \right]
\]

\[
- \frac{1'}{1'} \left( \frac{1}{1'} + \theta \mu \right) \left[ \frac{1}{1'} + \theta \mu \right] - \frac{1'}{1'} \left( \frac{1}{1'} + \theta \mu \right) \left[ \frac{1}{1'} + \theta \mu \right]
\]

\[
- \frac{1'}{1'} \left( \frac{1}{1'} + \theta \mu \right) \left[ \frac{1}{1'} + \theta \mu \right] - \frac{1'}{1'} \left( \frac{1}{1'} + \theta \mu \right) \left[ \frac{1}{1'} + \theta \mu \right]
\]

\[
= 2 \left[ \frac{1}{4} \frac{\theta^2 C}{B^2} + \frac{1}{2} \theta \left( A + 2 \frac{AC}{B^2} - 2 \frac{D}{B} \right) + \frac{A^2}{B} + \frac{A^2 C}{B^2} - 2 \frac{AD}{B} + F - E \right],
\]

where the last equality follows from Eq. (4.11). Thus Eq. (4.9) has been proved. Finally, solving the equation \( \Lambda'(\theta) = r \), we can obtain the exposing point \( \hat{\theta}_r \) given in Eq. (4.11). QED.
5 Conclusion

In this paper, we present the large deviations approximations for the windfall and shortfall probabilities of the [one-period-ahead] return of a diversified optimal portfolio. Moreover, we show that, in a sufficiently large portfolio, an optimal invested portfolio, which yields the maximum windfall probability or the minimum shortfall probability, can also be obtained by maximizing an endogenous profit function or minimizing an endogenous loss function, respectively. The results in this paper suggest that, whatever the choice of investment vehicle, it is essential to have a sufficient diversified portfolio. This is because investing in only one or two assets is extremely risky and entirely inappropriate for the majority of investors. For instance, as far as hedge funds are concerned, holding funds of hedge funds has now become popular.

Our analytical framework, although not taking into consideration dynamic strategies, proves tractable in solving the problems posed above and illustrates the potential insights offered by this type of approach. Future research shall focus on studying dynamic investment strategies for diversified portfolios. We are aware that this is a quite challenging problem because of two reasons: First, we need to include some potentially important features of the data such as serial dependence and heteroscedasticity. Second, handling those features of the data by using well-known mathematical apparatuses such as stochastic calculus is nontrivial in the high dimension.
References


