

# Taylor Series Approximations to Expected Utility and Optimal Portfolio Choice\*

Lorenzo Garlappi<sup>†</sup>      Georgios Skoulakis<sup>‡</sup>

September 2, 2011

## Abstract

This paper revisits the subject of Taylor series approximations to expected utility and investigates the applicability of the technique to optimal portfolio choice problems. We first provide conditions under which the approximate expected utility of a given portfolio converges to its exact counterpart. We then extend the analysis to the optimal portfolio choice setting and provide conditions on the distribution of asset returns under which the solution to the approximate portfolio choice problem converges to its exact counterpart. Finally, we show that, when asset returns are skewed, one can improve the precision and efficiency of the Taylor expansion by applying a simple nonlinear transformation to asset returns designed to symmetrize the transformed return distribution and shrink its support.

---

\*We wish to thank participants at the 2008 Latin American Meeting of the Econometric Society (LAMES), Rio de Janeiro, Brazil, Carlos Berenstein, Su Li, Oleg Rytchkov, Konstantinos Spiliopoulos, and especially Harry Tamvakis for valuable input and suggestions. Any errors are our sole responsibility. Comments and suggestions are welcome at [lorenzo.garlappi@sauder.ubc.ca](mailto:lorenzo.garlappi@sauder.ubc.ca) or [gskoulak@rhsmith.umd.edu](mailto:gskoulak@rhsmith.umd.edu).

<sup>†</sup>Finance Department, Sauder School of Business, University of British Columbia, Vancouver, BC, Canada; Email: [lorenzo.garlappi@sauder.ubc.ca](mailto:lorenzo.garlappi@sauder.ubc.ca).

<sup>‡</sup>Finance Department, Robert H. Smith School of Business, University of Maryland, College Park, MD, USA; Email: [gskoulak@rhsmith.umd.edu](mailto:gskoulak@rhsmith.umd.edu).

# 1 Introduction

Since the seminal work of Arrow (1964) and Pratt (1964), there has been a voluminous literature on the use of polynomial approximations for the computation of expected utility.<sup>1</sup> Broadly speaking, polynomial approximations to expected utility can be thought of as a computational tool for approximating integrals. Although alternative numerical techniques, such as quadrature<sup>2</sup> and Monte Carlo simulation,<sup>3</sup> for approximating integrals have become more sophisticated and powerful since the initial work of Arrow (1964) and Pratt (1964), more recently the subject of polynomial approximations has re-emerged in the context of dynamic portfolio choice (see, e.g., Wu (2003), Brandt, Goyal, Santa-Clara, and Stroud (2005), and Garlappi and Skoulakis (2010)). It turns out that, due to the numerical complexity of these problems, it is often computationally efficient to approximate the utility function by a polynomial obtained via a Taylor expansion.

There are, however, several issues that need to be dealt with before safely relying on Taylor series for the approximation of expected utility. First, it is important to ensure that a series converges to the exact expected utility as more terms are added. Obviously, convergence depends on the type of utility function: for some utility functions such as the exponential, the series converges for all possible level of returns (i.e., the series has an infinite *radius of convergence*) while for other utility functions, such as those in the power class, convergence is guaranteed only on a specific range of portfolio returns. Second, the use of a Taylor series inevitably raises the question of how many terms one needs to incorporate in the Taylor expansion. Perhaps the harshest criticism on the use of Taylor series approximation is offered by Hlawitschka (1994) who, building on the earlier work of Loistl (1976), argues that the usefulness of Taylor series approximations is strictly an empirical issue unrelated to the convergence properties of the infinite series, and, most importantly, that *even for a convergent series* adding more terms does not necessarily improve the quality of the approximation.

Despite the long history of polynomial approximations to utility functions, the literature to date provides remarkably little guidance for a correct implementation of these techniques. The purpose of this paper is to fill this gap by establishing a set of conditions under which Taylor series can be used as a sound computational tool for both the evaluation of expected utility of a given portfolio as well as the solution to a portfolio choice problem. Our results are derived for the class of HARA utility functions and under the assumption that the distribution of asset returns has bounded support.

In the first part of the paper (section 2), we analyze the Taylor series approximations to the expected utility of a *given* portfolio of assets. This is the problem studied by the early literature cited above. If the distribution of future wealth is bounded, we show that, by choosing an expansion point in the upper half of the return distribution, the Taylor series is guaranteed

---

<sup>1</sup>A partial list of papers include Feldstein (1969), Samuelson (1970), Tsiang (1972), Borch (1974), Loistl (1976), Levy and Markowitz (1979), Pulley (1981), Pulley (1983), Kroll, Levy, and Markowitz (1984), Markowitz (1991), Hlawitschka (1994), and many others.

<sup>2</sup>See, for instance, Judd (1998) for a comprehensive textbook treatment and Tauchen and Hussey (1991) for applications to asset pricing models.

<sup>3</sup>See the textbooks of Judd (1998) and Glasserman (2004).

to converge to the exact expected utility. Moreover, by choosing a sufficiently large center of expansion, it is possible to obtain a Taylor series for which the sequence of partial sums with *even* number of terms converges *monotonically* to the exact expected utility. This addresses the concern raised by Hlawitschka (1994) that adding more terms in a series does not necessarily improve the approximation.

In the second part of the paper (section 3), we address the more challenging issue of providing conditions for the sound use of Taylor series approximations for solving portfolio choice problems. The complexity of this problem arises from the fact that in order to be able to correctly use Taylor approximations, it is important to make sure that the proposed approximation converges *uniformly* across all admissible portfolio allocations. We determine conditions on the shocks to asset returns that guarantee convergence of the solution of the approximate optimal portfolio choice to the exact optimal portfolio. As for the case of expected utility calculation, the choice of the center of expansion is crucial for ensuring convergence of the approximate optimal allocations. In light of recent work that has relied on Taylor approximations for the solution of dynamic portfolio choice problems, this result is important because it provides a way to assess when such approximations are a sound computational tool and when they should be avoided.

The key to a successful implementation of the Taylor approximation is to minimize the size of the deviations of realized returns from the chosen center of expansion. The ideal setup would be the one in which the distribution of portfolio return is *symmetric* and has *narrow support*. Building on a principle widely used in statistics and econometrics (e.g., Box and Cox (1964)), we demonstrate that the quality of the Taylor approximation can be improved by performing a suitable *nonlinear transformation* of the portfolio return that makes the distribution of the transformed return (close to) symmetric and shrinks its support. This change-of-variable approach relates to, and can be viewed as an extension of, the loglinearization technique that has been used in the literature to obtain approximate analytical solutions to portfolio choice models (see, e.g., Campbell and Viceira (1999), Campbell and Viceira (2002), and Campbell, Chan, and Viceira (2003)). While loglinearization amounts to a first-order Taylor expansion with respect to the logarithm of the variable of interest, the approach pursued in this paper is based on a generic Taylor expansion with respect to a nonlinear transformation of asset returns. Our numerical implementations indeed illustrate that, when asset returns are skewed and/or exhibit excess kurtosis, Taylor approximations applied to the transformed asset returns converge faster and are more accurate.

The convergence results in this paper require that the support of the risky asset return distribution be bounded. This assumption deserves some discussion. From a modeling perspective, the return distribution for a specific investment horizon can reasonably be assumed to have bounded support. Indeed, given the finite amount of data that we observe in practice, it seems impossible to statistically reject the assumption that a distribution of asset returns has bounded support. From a purely technical point of view, compactness appears to be a standard requirement for rigorously establishing convergence results in the context of numerical solutions of asset pricing models (see, among others, Tauchen and Hussey (1991), Rust (1997), Marcet and Sin-

gleton (1999), and Koo (1999)). For instance, the convergence of the quadrature approximation, a commonly used numerical procedure for computing integrals, has so far been formally proven only under the assumption of shock distributions with compact support (see, e.g. Tauchen and Hussey (1991)). However, in practice, researchers, including Tauchen and Hussey (1991) in their applications, use the quadrature method under the assumption of normally distributed shocks which is clearly at odds with the theoretical requirement of compact support. In justifying this practice, Tauchen and Hussey (1991) make the following statement: “. . . we follow the tradition in this area of research of applying results deduced for the case of bounded support to models with unbounded support.” (p. 380) We are not aware of theoretical results that directly establish convergence of numerical methods in this context without the compactness assumption. Nevertheless, in the final section of the paper, we show how one can approximate the solution to a portfolio choice problem with returns having unbounded support by solving a sequence of problems in which the support of the return distribution is suitably truncated, and to which the Taylor approximation approach is applicable.

The theoretical results in this paper are proved using complex-theoretic techniques. The proofs utilize a number of auxiliary results that are presented in Appendix A. All proofs are presented in Appendix B.

## 2 Expected utility

In this section, we consider Taylor series expansions for approximating the expected utility of a given portfolio. We first provide general conditions that guarantee convergence of the approximate expected utility to the exact expected utility. We then show that, by choosing an appropriate nonlinear transformation of portfolio returns, it is possible to improve accuracy of the Taylor series approximation. We conclude the section with a set of numerical examples.

### 2.1 Taylor approximation based on a linear decomposition

Consider an investor with initial wealth  $W_0$  invested in a portfolio that earns a stochastic gross return  $R_p$  over a given planning horizon. Let  $U(W_1)$  denote the utility of terminal wealth  $W_1 = W_0 R_p$ . We restrict attention to utility functions in the HARA class

$$U(W_1) = \frac{1}{1-\gamma} \left( W_1 + \frac{\gamma}{\alpha} \right)^{1-\gamma}, \quad (1)$$

where  $\alpha \neq 0$ ,  $\gamma \neq 1$ ,  $\gamma > 0$  and  $W_1 + \frac{\gamma}{\alpha} > 0$ .<sup>4</sup> Without loss of generality, we assume that  $W_0 = 1$  and focus on the utility of portfolio return  $R_p$ <sup>5</sup>

$$U(R_p) = \frac{1}{1-\gamma} \left( R_p + \frac{\gamma}{\alpha} \right)^{1-\gamma}. \quad (2)$$

The Taylor series approximation to the utility  $U(R_p)$  is obtained as follows. The starting point is the linear decomposition  $R_p = c_p + d_p$ , where  $c_p$  is the deterministic constant to be used as the center of the Taylor expansion and  $d_p = R_p - c_p$  is the associated stochastic deviation. The utility (2) as a function of the deviation  $d_p$  then reads

$$\tilde{U}(d_p) = \frac{1}{1-\gamma} \left( c_p + d_p + \frac{\gamma}{\alpha} \right)^{1-\gamma}. \quad (3)$$

Because  $\tilde{U}(d_p) = U(R_p)$  and  $\tilde{U}(0) = U(c_p)$ , the  $M$ -th order Taylor approximation  $U_M(R_p)$  to  $U(R_p)$  can be expressed as  $U_M(R_p) = \sum_{m=0}^M \frac{1}{m!} U^{(m)}(c_p) d_p^m$ , or equivalently

$$\tilde{U}_M(d_p) = \sum_{m=0}^M \frac{1}{m!} \tilde{U}^{(m)}(0) d_p^m, \quad (4)$$

where

$$U^{(m)}(c_p) = \tilde{U}^{(m)}(0) = (-\gamma)_{m-1} \left( c_p + \frac{\gamma}{\alpha} \right)^{1-\gamma-m}, \quad m = 1, 2, \dots \quad (5)$$

and  $(x)_k$  denotes the falling factorial  $(x)_k \equiv x(x-1)\cdots(x-k+1)$ . The resulting approximate expected utility is

$$\mathbb{E}[\tilde{U}_M(d_p)] = \sum_{m=0}^M \frac{1}{m!} \tilde{U}^{(m)}(0) \mathbb{E}[d_p^m]. \quad (6)$$

Throughout this section, we will consider portfolio returns with support satisfying the following assumption.

**Assumption 2.1** *The support of the distribution of portfolio return  $R_p$  is the closed interval  $[\underline{R}_p, \overline{R}_p]$ , where  $\underline{R}_p > -\frac{\gamma}{\alpha}$ .*

The following proposition<sup>6</sup> provides conditions under which the approximate expected utility derived from equation (4) converges to the exact expected utility  $\mathbb{E}[U(R_p)]$ , as the number of Taylor terms  $M$  tends to infinity.

---

<sup>4</sup>The HARA class contains a number of frequently used utility functions as special cases. As  $\gamma \rightarrow 1$  and  $\alpha \rightarrow \infty$  in (1), we obtain the logarithmic utility function. When  $\gamma \neq 1$  and  $\alpha \rightarrow \infty$ , we obtain the power utility function with CRRA coefficient  $\gamma$ . Finally, when  $\alpha > 0$  and  $\gamma \rightarrow \infty$ , we obtain the exponential utility function with CARA coefficient  $\alpha$ .

<sup>5</sup>To obtain this normalization the parameters of the HARA utility function need to be adjusted in an obvious fashion. From  $W_1 = W_0 R_p$  we have  $U(W_1) = W_0^{1-\gamma} [1/(1-\gamma)(R_p + \gamma/\alpha W_0)^{1-\gamma}]$ , where the quantity in square bracket is of the HARA form (1).

<sup>6</sup>See also Theorem 1 and Table 1.1 in Jurczenko and Maillet (1996) for similar results.

**Proposition 2.2** *Let  $R_p$  satisfy Assumption 2.1,  $c_p$  be the center of expansion of the Taylor approximation to utility (2), and  $d_p = R_p - c_p$  be the associated portfolio return deviation. If  $|d_p| \leq D_p$ , with  $D_p < D_p^* = c_p + \frac{\gamma}{\alpha}$ , then  $\tilde{U}_M(d_p)$  converges to  $\tilde{U}(d_p)$  uniformly in  $d_p$  and the approximate expected utility  $\mathbb{E}[\tilde{U}_M(d_p)]$  converges to the exact expected utility  $\mathbb{E}[\tilde{U}(d_p)]$ , as  $M \rightarrow \infty$ .*

The following corollary shows how to select the center of expansion in order to guarantee convergence of the Taylor series.

**Corollary 2.3** *Let  $R_p$  satisfy Assumption 2.1. If  $c_p > \frac{1}{2}(\bar{R}_p - \frac{\gamma}{\alpha})$ , then there exists a bound  $D_p < c_p + \frac{\gamma}{\alpha}$  so that Proposition 2.2 applies. Moreover,  $c_p \geq \frac{1}{2}(\underline{R}_p + \bar{R}_p)$  implies  $c_p > \frac{1}{2}(\bar{R}_p - \frac{\gamma}{\alpha})$ .*

The second part of the corollary implies that choosing the midpoint of the support of the return distribution always guarantees convergence of the Taylor series. Note that if the distribution of  $R_p$  is *symmetric*, then the midpoint  $\frac{1}{2}(\underline{R}_p + \bar{R}_p)$  equals the expected portfolio return  $\mu_p$  which is the common choice of expansion point for Taylor approximations to expected utility (see, e.g., Loistl (1976)). If the distribution of  $R_p$  is *negatively skewed*, then the mean  $\mu_p$  is larger than the midpoint  $\frac{1}{2}(\underline{R}_p + \bar{R}_p)$ . In these two cases, selecting the mean  $\mu_p$  as the center of expansion ensures that the sufficient condition in Corollary 2.3 is satisfied and therefore convergence is guaranteed. However, if the portfolio return distribution is *positively skewed*, choosing the mean  $\mu_p$  as the center of expansion *might* lead to violation of the sufficient condition in Proposition 2.2. If the condition in Proposition 2.2 is indeed violated and the portfolio return takes values outside the radius of convergence with positive probability then, as the next proposition demonstrates, the Taylor approximation to expected utility does not converge to the exact expected utility. This result shows that, within the context of HARA utility functions, the assumption of bounded support is actually necessary for establishing convergence of the Taylor approximation.

**Proposition 2.4** *Let  $c_p$  be the center of expansion of the Taylor approximation to utility (2), and  $d_p = R_p - c_p$  be the associated portfolio return deviation. If  $\mathbb{P}[|d_p| \geq L] = q > 0$ , where  $L > D_p^* = c_p + \frac{\gamma}{\alpha}$ , then the series representing the approximate expected utility in (6) does not converge.*

Knowing that a Taylor series converges is not sufficient to guarantee that including additional terms to the expansion improves the approximation. As Hlawitschka (1994) points out, a Taylor expansion with more terms may provide a worse approximation of the exact expected utility than one with fewer terms. Therefore, it is important, for practical purposes, to obtain *a priori* conditions under which including additional terms improves the quality of the approximation. The approximate expected utility in (6) can be expressed as follows

$$\mathbb{E}[U_M(R_p)] = \sum_{m=0}^M \frac{1}{m!} U^{(m)}(c_p) \mathbb{E}[d_p^m]. \quad (7)$$

Note that, according to (5),  $U^{(m)}(\cdot) > 0$  if  $m$  is odd and  $U^{(m)}(\cdot) < 0$  if  $m$  is even. Therefore, the contribution of the  $m$ -th term in (7) is determined by the sign of the  $m$ -th moment  $\mathbb{E}[d_p^m]$ . This moment is positive for even  $m$  and therefore all even-order terms in the approximation (7) are negative. In contrast, for odd  $m$  the moment  $\mathbb{E}[d_p^m]$  can be either positive or negative. If the moment is negative (positive) then the corresponding term in the expansion is also negative (positive). Hence, if the odd-order moments are positive then the sequence of partial Taylor sums is oscillating and the convergence of the approximate expected utility to its exact counterpart is not monotonic. These observations provide a theoretical explanation for the issue of non-monotonicity highlighted in Hlawitschka (1994). It also suggests that, by focusing attention on partial Taylor sums of even order, it is possible to achieve monotone convergence. The following proposition makes this argument precise and provides a sufficient condition for monotone convergence of the Taylor approximation.<sup>7</sup>

**Proposition 2.5** *Suppose that the portfolio return deviation  $d_p$  satisfies the condition  $|d_p| \leq D_p$  with  $D_p < D^* = c_p + \frac{\gamma}{\alpha}$ . If the moments of  $d_p$  are such that, for  $K = 1, 2, \dots$ ,*

$$U^{(2K-1)}(c_p)\mathbb{E}[d_p^{2K-1}] + \frac{1}{2K}U^{(2K)}(c_p)\mathbb{E}[d_p^{2K}] < 0, \quad (8)$$

*then the sequence  $\mathbb{E}[U_{2K}(R_p)]$  converges monotonically from above to the exact expected utility  $\mathbb{E}[U(R_p)]$ . For the case of HARA utility, the above condition is equivalent to*

$$2K \left( c_p + \frac{\gamma}{\alpha} \right) \mathbb{E}[(R_p - c_p)^{2K-1}] < (\gamma + 2K - 2)\mathbb{E}[(R_p - c_p)^{2K}]. \quad (9)$$

The intuition for the condition in the above proposition is simple: because even-order terms are negative, the partial sum of order  $2K - 1$  has a larger value than the partial sum of order  $2K$ . To guarantee that the sequence of partial Taylor sums of even order converges monotonically, it suffices that the sum of the  $2K - 1$  and  $2K$  order terms be negative, for  $K = 1, 2, \dots$ . This is exactly what conditions (8) and (9) state. Importantly, the choice of the center of expansion  $c_p$  is crucial for ensuring that (9) holds. As  $c_p$  increases,  $\mathbb{E}[(R_p - c_p)^{2K-1}]$  decreases and eventually becomes negative. In contrast, the right-hand side in (9) is always positive. Therefore, when the number of Taylor terms is even, say  $M = 2L$ , one can choose a large enough  $c_p$  so that condition (9) holds for all  $K = 1, \dots, L$ . Selecting  $c_p$  in such a fashion yields monotonic convergence of the Taylor approximation, if one considers partial sums of even order. However, one should avoid choosing an excessively large value for  $c_p$  since this might result in slow convergence.

## 2.2 Taylor approximation based on a nonlinear decomposition

A successful implementation of any approximation scheme relies on controlling and minimizing the approximation error. In the context of Taylor expansions used to approximate expected utility of portfolio returns, this means minimizing the deviations of realized returns from the

---

<sup>7</sup>Similar conditions are also derived in Jurczenko and Maillet (1996) and Jondeau and Rockinger (2006).

chosen center of expansion. To this purpose, the ideal setup for Taylor expansion is one in which the distribution of portfolio return is *symmetric* and with *narrow support*. Symmetry helps because, when choosing the midpoint of the support of the return distribution as the center of expansion, it is possible to make the highest possible deviation from the center as small as possible. A narrow distribution helps because it ensures that deviations from the center of expansion, and hence approximation errors, are not too large.

In this subsection, we argue that, by using a proper *nonlinear transformation* of portfolio returns, it is possible to both reduce the degree of asymmetry of the return distribution and shrink its support. In the next subsection, we show, through a set of numerical experiments, that using the Taylor series expansion on the transformed variable improves the accuracy of the approximation. The motivation for using such a transformation is related to the practice, widely used in statistics and econometrics, of transforming skewed data via a nonlinear mapping with the goal of approaching symmetry/normality. The Box-Cox transformation is a prominent example of such transformations (see, e.g., Box and Cox (1964), Draper and Cox (1969), Hinkley (1975), and Sakia (1992) among others). The transformation we propose can also be thought as a generalization of the loglinearization technique advanced by Campbell and his authors for providing approximate analytical solutions for asset pricing models.

To illustrate the main intuition, consider the case of modeling gross returns as lognormal random variables, as frequently used in the life-cycle portfolio choice literature (see, e.g., Campbell and Viceira (1999), Campbell and Viceira (2002), Campbell, Chan, and Viceira (2003)). Under this assumption, gross returns are positively skewed. Because  $R_p = e^{\log(R_p)}$ , the log return  $\log(R_p)$  is, by definition, normal and therefore symmetric. Moreover, because the exponential is a convex function, the support of  $\log(R_p)$  is *narrower* than the support of  $R_p$ , i.e., the log transformation shrinks the support of  $R_p$ .

In the context of the HARA utility (2), the above intuition suggests working with the following transformation

$$r_p = \log \left( R_p + \frac{\gamma}{\alpha} \right). \quad (10)$$

To proceed with the Taylor expansion, we decompose the transformed variable  $r_p$  as  $r_p = \kappa_r + \delta_r$ , where  $\kappa_r$  is the deterministic quantity to be used as center of the expansion and  $\delta_r = r_p - \kappa_r$  is the associated stochastic deviation. Note that  $\delta_r$  does not necessarily have mean zero. From equation (2), the utility and its derivatives, defined in terms of the deviation  $\delta_r$ , are<sup>8</sup>

$$\widehat{\mathcal{U}}(\delta_r) = \frac{1}{1-\gamma} e^{(1-\gamma)(\kappa_r + \delta_r)}, \quad (11)$$

$$\widehat{\mathcal{U}}^{(m)}(\delta_r) = \frac{1}{1-\gamma} (1-\gamma)^m e^{(1-\gamma)(\kappa_r + \delta_r)}, \quad m = 1, 2, \dots, \quad (12)$$

and the  $M$ -th order Taylor approximation is

$$\widehat{\mathcal{U}}_M(\delta_r) = \sum_{m=0}^M \frac{1}{m!} \widehat{\mathcal{U}}^{(m)}(0) \delta_r^m = \frac{e^{(1-\gamma)\kappa_r}}{1-\gamma} \sum_{m=0}^M \frac{1}{m!} (1-\gamma)^m \delta_r^m. \quad (13)$$

---

<sup>8</sup>We introduce the notation  $\widehat{\cdot}$ , as opposed to  $\widetilde{\cdot}$ , to differentiate between the linear decomposition in 2.1 and the present nonlinear decomposition.

It is well known that the Taylor series approximation to the exponential function converges over the entire real line. Hence, the above *pointwise* Taylor approximation (13) converges for *any* value of the deviation  $\delta_r$ . It follows that, if  $|\delta_r|$  is bounded by *any* positive real number  $\Delta$ , then the approximate expected utility based on a Taylor series with respect to  $r_p$  converges to the exact expected utility.

The logarithmic transformation is of practical importance when portfolio returns are *positively* skewed since it reduces the skewness and, hence, makes the distribution of the log variable more symmetric. This feature has implications for the speed of convergence and accuracy of the Taylor approximation. However, if returns are *negatively* skewed, this transformation will make the log variable even more negatively skewed and, hence, worsen the accuracy of the approximation. Even if asset returns are positively skewed, a negatively skewed portfolio return can easily emerge if we allow short positions. Therefore, when implementing a Taylor expansion, it is important to reduce the asymmetry of return distribution also in the case of negatively skewed returns. To this purpose, we need a sufficiently flexible nonlinear transformation that can symmetrize the distribution and shrink the support of both positive and negatively skewed returns.

Among the many possible choices, we propose a parsimonious transformation that implicitly defines the transformed return  $y$  from the original portfolio return  $R_p$  through the following nonlinear relation

$$R_p = A + Be^{Cy}. \quad (14)$$

From the functional form of (14), we can assume, with no loss of generality, that the transformed random variable  $y$  has zero mean and unit variance. Note that the structure of the transformation (14) can accommodate different values of skewness. In particular, assuming  $C > 0$ , a positively (negatively) skewed distribution for  $R_p = A + Be^{Cy}$  can be accommodated by a symmetric or positively skewed distribution for  $y$  and  $B > 0$  ( $< 0$ ). In this case, the transformed variable used in the Taylor expansion will be  $y = \frac{1}{C} \log\left(\frac{R_p - A}{B}\right)$ . As before, we use a deterministic quantity  $\kappa_y$  as the center of expansion and  $\delta_y = y - \kappa_y$  as the associated stochastic deviation. Note that the logarithmic transformation in equation (10) is a special case of the transformation in equation (14) for  $A = -\gamma/\alpha$ ,  $B = 1$ , and  $C = 1$ .

From equation (2), the utility defined in terms of the deviation  $\delta_y$  is

$$\widehat{U}(\delta_y) = \frac{1}{1-\gamma} \left( A + Be^{Cy} + \frac{\gamma}{\alpha} \right)^{1-\gamma}, \quad (15)$$

and the  $M$ -th order Taylor approximation with respect to  $\delta_y$  is

$$\widehat{U}_M(\delta_y) = \sum_{m=0}^M \frac{1}{m!} \widehat{U}^{(m)}(0) \delta_y^m. \quad (16)$$

Note that  $\widehat{U}^{(m)}(0)$  is the derivative of the composite function  $U(g(\delta_y))$  where  $g(\delta_y) = A + Be^{C\kappa_y + C\delta_y} + \frac{\gamma}{\alpha}$ , evaluated at  $\delta_y = 0$ . Derivatives of composite functions can be efficiently computed using the recursive scheme known as the Faà di Bruno formula; see Savits (2006) for

the underlying theory and Garlappi and Skoulakis (2010) for applications to portfolio choice problems.

The next proposition formally characterizes the domain of convergence of the Taylor series expansion with respect to the transformed variable  $y$ .

**Proposition 2.6** *Let  $R_p$  satisfy Assumption 2.1, define  $y = \frac{1}{C} \log\left(\frac{R_p - A}{B}\right)$ , and let  $\Theta = \frac{\gamma/\alpha + A}{Be^{C\kappa_y}}$ . Then, the Taylor approximation (16) with respect to the deviation  $\delta_y = y - \kappa_y$  converges to the exact utility (15) for all  $\delta_y$  with  $|\delta_y| < \Delta_p^*$  where*

$$\Delta_p^* = \begin{cases} \sqrt{[\log(\Theta)]^2 + \pi^2}, & \text{if } \Theta > 0 \\ |\log(-\Theta)|, & \text{if } \Theta < 0. \end{cases} \quad (17)$$

The convergence is uniform on any interval  $[-\Delta_p, \Delta_p]$ , where  $0 < \Delta_p < \Delta_p^*$ , and, hence, the approximate expected utility  $\mathbb{E}[\widehat{U}_M(\delta_y)]$  converges to the exact expected utility  $\mathbb{E}[\widehat{U}(\delta_y)]$ , as  $M \rightarrow \infty$ .

Note that, in the case of the logarithmic transformation in equation (10), we have  $A = -\gamma/\alpha$  and hence, by (17), the domain of convergence of the Taylor series with respect to the log returns is the entire real line.

To make the preceding proposition operational one needs to determine the parameters  $A$ ,  $B$ , and  $C$  used in the transformation (14). An intuitive as well as computationally tractable approach to this task is to treat the transformed variable  $y$  as normally distributed (with mean 0 and variance 1). Using well-known properties of the lognormal distribution, it is possible to specify  $A$ ,  $B$ , and  $C$  such that the first three moments of  $A + Be^{Cy}$  match the first three moments of  $R_p$ . Specifically, given target values  $(\mu_p, \sigma_p, \xi_p)$  of the mean, standard deviation, and skewness of  $R_p$ , we can find  $A$ ,  $B$ , and  $C$  such that

$$\begin{aligned} \mu_p &= A + Be^{\frac{1}{2}C^2}, \\ \sigma_p^2 &= B^2 e^{C^2} (e^{C^2} - 1), \\ \xi_p &= \text{sign}(B) (e^{C^2} - 1)^{1/2} (e^{C^2} + 2). \end{aligned}$$

Solving the system for  $A$ ,  $B$ , and  $C$  yields

$$A = \mu_p - Be^{C^2/2}, \quad (18)$$

$$B = \text{sign}(\xi_p) \frac{\sigma_p}{\sqrt{e^{C^2}(e^{C^2} - 1)}}, \quad (19)$$

$$C = \sqrt{\log\left(\Gamma - 1 + \frac{1}{\Gamma}\right)}, \quad \Gamma = \left(\frac{\xi_p^2 + \xi_p \sqrt{4 + \xi_p^2} + 2}{2}\right)^{1/3}. \quad (20)$$

Note that, while the assumption of normality for  $y$  is clearly at odds with the bounded support condition required by Proposition 2.6, it simply provides a practical way of obtaining the parameters  $A$ ,  $B$ , and  $C$  and implementing the nonlinear transformation  $y = \frac{1}{C} \log\left(\frac{R_p - A}{B}\right)$ . In next subsection, through numerical examples, we illustrate the effectiveness of the nonlinear decomposition discussed above.

### 2.3 Numerical examples

We now turn to some numerical examples to illustrate the practical implications of the theoretical issues discussed above. Although the full advantage of Taylor approximation becomes more evident in multi-period portfolio choice problems (see Garlappi and Skoulakis (2010)), in this section we limit ourselves to the case of single-period expected utility calculations, to fully reflect the theoretical arguments of the previous two subsections.

Broadly speaking, Taylor approximation to expected utility simply facilitates an approximate computation of an expectation. Several alternative methods are available in the literature such as Monte Carlo simulation (e.g., Glasserman (2004)) or quadrature methods (e.g., Judd (1998)). It is beyond the scope of our paper to thoroughly discuss the advantages and disadvantages of each of these approaches. It is worth mentioning, however, that, in general, Monte Carlo methods can be computationally inefficient, especially if one wishes to achieve high accuracy, and quadrature methods require specific distributional assumptions, such as normality, and also suffer from the curse of dimensionality. In contrast, the Taylor approximation results derived above are valid under *any* portfolio return distribution, provided it has a bounded support. Furthermore, the implementation of the method is computationally efficient, an aspect that becomes evident in dynamic portfolio choice applications.

We consider the problem of assessing the expected utility  $\mathbb{E}[U(R_p)]$ , where  $U(\cdot)$  belongs to the CRRA class with risk aversion  $\gamma$  (special case of the HARA class in (2), when  $\alpha \rightarrow \infty$ ). We consider two types of distribution for log portfolio returns  $r_p$ : (i) normal distribution and (ii) mixture of normal distributions. The latter is used because it can easily incorporate skewness and excess kurtosis. We consider two Taylor approximations and present the results in Table 1. The first decomposition, labeled “Gross return decomposition”, is based on the decomposition of the gross portfolio return  $R_p = c_p + d_p$  while the second, labeled “Log return decomposition”, is based on the decomposition of the log portfolio return  $r_p = \kappa_r + \delta_r$ . Since our theoretical results are derived under the assumption of a bounded support for the distribution of shocks, we numerically implement the Taylor approximation to expected utility by using the Gauss-Hermite quadrature approximation when the distribution of  $r_p$  is assumed to be normal. When the distribution of  $r_p$  is assumed to be a mixture of normal distributions, we use the same approximation for each component of the mixture. In all cases, we use five nodes in the quadrature approximation. When  $r_p$  is assumed to follow a normal distribution, we set the center of expansion  $c_p$  equal to the midpoint of the support of the quadrature-based approximation to the distribution of  $R_p$ , and the center of expansion  $\kappa_r$  equal to the mean of  $r_p$ . When  $r_p$  is assumed to follow a mixture of normal distributions, we apply the same procedure to each component of the mixture. In the latter case, the expected utility is equal to the weighted average of expected utilities that correspond to the components of the mixture.

Utility is an ordinal concept and unique only up to positive affine transformations, and, hence, in order to assess the quality of an approximation to expected utility, we cannot compare the *levels* of exact and approximate utility. Instead, following Levy and Markowitz (1979) and Hlawitschka (1994), we rank the portfolios based on their exact and approximate expected utility,

and then compute the Spearman correlation between the two sets of expected utilities. High values of the correlation, i.e., close to 1, imply better quality of the approximation.

Table 1 reports the Spearman correlations between the exact and approximate expected utilities based on the rankings of 11 portfolios. The parameters that determine the distribution of the log return  $r_p$  are selected at random to roughly correspond to quarterly equity returns. When the distribution of  $r_p$  is assumed to be normal, the mean  $\mu_r$  and the standard deviation  $\sigma_r$  of  $r_p$  are drawn from the intervals  $[\underline{\mu}, \bar{\mu}] = [0.03, 0.06]$  and  $[\underline{\sigma}, \bar{\sigma}] = [0.10, 0.15]$ , respectively. When the distribution of  $r_p$  is assumed to be a mixture of normal distributions, we use mixtures with four components. The vector of mixing weights is proportional to  $\exp(\mathbf{u})$  where  $\mathbf{u}$  is a four-dimensional vector of uniform (0,1) numbers drawn at random independently. For  $k = 1, 2, 3, 4$ , the mean  $\mu_{r,k}$  and standard deviation  $\sigma_{r,k}$  of the  $k$ -th normal component are drawn from the intervals  $[\underline{\mu}_k, \bar{\mu}_k]$  and  $[\underline{\sigma}_k, \bar{\sigma}_k]$ , respectively. The intervals used in the simulation are:  $[\underline{\mu}_1, \bar{\mu}_1] = [0.01, 0.03]$  and  $[\underline{\sigma}_1, \bar{\sigma}_1] = [0.20, 0.25]$ ,  $[\underline{\mu}_2, \bar{\mu}_2] = [0.03, 0.05]$  and  $[\underline{\sigma}_2, \bar{\sigma}_2] = [0.15, 0.20]$ ,  $[\underline{\mu}_3, \bar{\mu}_3] = [0.05, 0.07]$  and  $[\underline{\sigma}_3, \bar{\sigma}_3] = [0.10, 0.15]$ ,  $[\underline{\mu}_4, \bar{\mu}_4] = [0.07, 0.09]$  and  $[\underline{\sigma}_4, \bar{\sigma}_4] = [0.05, 0.10]$ . The resulting mixtures of normal distributions exhibit skewness and excess kurtosis.<sup>9</sup> For different values of the coefficient of risk aversion  $\gamma$  and order  $M$  of the Taylor expansion, the table reports the median Spearman correlation across 10,000 simulation repetitions.

There are several observations that emerge from Table 1. For both distributional assumptions (normal and mixture of normals) and both decompositions (with respect to gross and log return), we observe that (i) the correlations are monotonically increasing in the order of the Taylor expansion  $M$ , and (ii) the approximation is deteriorating as the curvature of the utility function increases (i.e., it is worse for higher  $\gamma$ ).<sup>10</sup> Most importantly, for both distributional assumptions (normal and mixture of normals), the correlation coefficients for the log return decomposition are larger than those obtained by using the gross return decomposition. The reason is that, in the case of the log return decomposition, the magnitude of the potential shocks is reduced, thus resulting in a better behaved Taylor approximation. Finally, the low level of correlation for low values of  $M$  indicates that the quality of low order expansions quickly deteriorates as the level of risk aversion increases. This suggests that the results shown by Hlawitschka (1994), who only uses CRRA preferences with risk aversion between 0.1 and 0.9, are not robust. This feature is more pronounced for the case of mixture of normals for which the correlations, for  $M = 2$  and  $\gamma$  between 6 and 10, are actually negative illustrating the poor performance of the quadratic utility approximation.

---

<sup>9</sup>Specifically, the three quartiles of the distribution of the mean, standard deviation, skewness and kurtosis are (0.047, 0.050, 0.053), (0.155, 0.162, 0.169), (-0.300, -0.255, -0.215), and (4.015, 4.260, 4.550), respectively.

<sup>10</sup>Note that the correlations are monotonically increasing in the order of expansion  $M$  for both the gross and log return decomposition. However, this need not be the case in general. In fact, in unreported results we find that, using the mean of the distribution of  $R_p$  as a center of expansion  $c_p$  may lead to *divergence* of the approximation, i.e., Spearman correlations decrease as  $M$  increases. This is a consequence of the fact that, for positively skewed distributions, the choice of the mean as a center of expansion might violate the convergence condition established in Proposition 2.2 and Corollary 2.3.

### 3 Optimal portfolio choice

Having addressed the convergence properties of Taylor approximation to expected utility, we can now focus on the properties of the approximation in the context of optimal portfolio choice. Recent papers use Taylor approximations as a convenient method to simplify the solution to dynamic portfolio choice problems (e.g., Brandt, Goyal, Santa-Clara, and Stroud (2005) and Garlappi and Skoulakis (2010)). However, the literature so far has provided little guidance to assess whether the use of such approximations is theoretically sound. In particular, it is not clear whether the portfolio obtained from solving the approximate portfolio choice problem converges to the exact optimal portfolio. Even in a single-period problem, this is a challenging task. The main complication resides in the fact that, unlike the problem of evaluating expected utility of a *given* portfolio, the portfolio allocation is a priori *not known* but endogenously determined as the solution to an optimization problem. Hence, to justify the use of Taylor approximations within this optimization problem, it is necessary to derive conditions on the shocks to asset returns that guarantee convergence of the Taylor series uniformly across *all* admissible portfolio allocation vectors.

As in the previous section, we will require that asset return distributions have bounded support. Our goal is to determine the largest possible deviations in asset return that guarantee (i) the convergence of the approximate expected utility to the exact expected utility and (ii) the convergence of the approximate optimal portfolio allocation to its exact counterpart. We will analyze separately the cases of portfolio choice with and without short-selling restrictions, and, as in the previous section we will demonstrate that, when asset returns are skewed, it is possible to improve the empirical properties of the approximation by using a Taylor series with respect to a suitable nonlinear transformation of the primitive asset returns.

Consider an investor with current wealth  $W_0$  and preferences over terminal wealth  $W_1$  described by the HARA utility function (1). As before, without loss of generality, we set  $W_0 = 1$  and concentrate on the portfolio return  $R_p$ . The investment opportunity set available to the investor consists of a risk-free asset with gross rate of return  $R_f$  and  $N$  risky assets with *gross* returns  $R_{g,n}$ ,  $n = 1, \dots, N$ . The vector of *excess* returns is denoted by  $\mathbf{R} = (R_1, \dots, R_N)'$  where  $R_n = R_{g,n} - R_f$ . The allocation to the risky assets is denoted by  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_N)'$ , the portfolio return is  $R_p(\boldsymbol{\omega}) = R_f + \boldsymbol{\omega}'\mathbf{R}$  and the resulting HARA utility is given by

$$U(R_f + \boldsymbol{\omega}'\mathbf{R}) = \frac{1}{1-\gamma} \left( R_f + \boldsymbol{\omega}'\mathbf{R} + \frac{\gamma}{\alpha} \right)^{1-\gamma}. \quad (21)$$

The set  $\boldsymbol{\Omega} \subset \mathbb{R}^N$  of admissible portfolio allocation vectors is assumed to be either bounded, closed, and convex (subsection 3.1) or finite (subsection 3.2).

Our objective is to solve the portfolio choice problem

$$\boldsymbol{\omega}^* = \arg \max_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} V(\boldsymbol{\omega}) \quad (22)$$

where

$$V(\boldsymbol{\omega}) = \mathbb{E} [U(R_f + \boldsymbol{\omega}'\mathbf{R})] \quad (23)$$

is the exact expected utility as a function of the portfolio allocation vector  $\boldsymbol{\omega} \in \boldsymbol{\Omega}$ .

Throughout this section, we make the following assumption that guarantees the existence of the utility function in (21) for all possible portfolio return realizations.

**Assumption 3.1** For all  $n = 1, \dots, N$ , the support of the distribution of the excess return  $R_n$  is the closed interval  $[\underline{R}_n, \overline{R}_n]$ . Moreover, for all return realizations  $\mathbf{R}$  and admissible portfolio weights  $\boldsymbol{\omega} \in \boldsymbol{\Omega}$ , we have  $R_f + \boldsymbol{\omega}'\mathbf{R} + \frac{\gamma}{\alpha} \geq \epsilon$ , where  $\epsilon$  is an arbitrarily small positive constant.

### 3.1 Taylor approximation based on a linear decomposition

The Taylor approximation method for solving the portfolio choice problem proceeds as follows. We decompose the vector of excess returns  $\mathbf{R}$  as  $\mathbf{R} = \mathbf{c} + \mathbf{d}$  where  $\mathbf{c} = (c_1, \dots, c_N)'$  is a vector of deterministic constants to be used as the center of the Taylor expansion and  $\mathbf{d} = (d_1, \dots, d_N)'$  is the corresponding vector of stochastic deviations. Naturally, we assume that  $c_n \in [\underline{R}_n, \overline{R}_n]$ . We then define the utility  $\tilde{U}(\mathbf{d}; \boldsymbol{\omega})$  in terms of the deviation vector  $\mathbf{d}$  as follows

$$\tilde{U}(\mathbf{d}; \boldsymbol{\omega}) = U(R_f + \boldsymbol{\omega}'\mathbf{R}) = \frac{1}{1-\gamma} \left( \left( R_f + \boldsymbol{\omega}'\mathbf{c} + \frac{\gamma}{\alpha} \right) + \boldsymbol{\omega}'\mathbf{d} \right)^{1-\gamma}. \quad (24)$$

The  $M$ -th order Taylor approximation to  $\tilde{U}(\mathbf{d}; \boldsymbol{\omega})$  with respect to  $\mathbf{d}$  centered around  $\mathbf{0}_N$ , the  $N$ -dimensional vector of zeros, is

$$\tilde{U}_M(\mathbf{d}; \boldsymbol{\omega}) = \sum_{|\mathbf{m}| \leq M} \frac{1}{\mathbf{m}!} \tilde{U}^{(\mathbf{m})}(\mathbf{0}_N; \boldsymbol{\omega}) \mathbf{d}^{\mathbf{m}} \quad (25)$$

where  $\mathbf{m} = (m_1, \dots, m_N)' \in \mathbb{N}^N$ ,  $\mathbf{m}! = m_1! \cdots m_N!$ ,  $|\mathbf{m}| = m_1 + \cdots + m_N$ , and  $\mathbf{x}^{\mathbf{m}} = x_1^{m_1} \cdots x_N^{m_N}$  for  $\mathbf{x} = (x_1, \dots, x_N)' \in \mathbb{R}^N$ , and  $\tilde{U}^{(\mathbf{m})}(\mathbf{0}_N; \boldsymbol{\omega}) = \frac{\partial^{m_1 + \cdots + m_N}}{\partial d_1^{m_1} \cdots \partial d_N^{m_N}} \tilde{U}(\mathbf{0}_N; \boldsymbol{\omega})$ . From the last equation, we then obtain the approximate expected utility based on the  $M$ -th order Taylor expansion as

$$\tilde{V}_M(\boldsymbol{\omega}) = \mathbb{E} \left[ \tilde{U}_M(\mathbf{d}; \boldsymbol{\omega}) \right] = \sum_{0 \leq |\mathbf{m}| \leq M} \frac{1}{\mathbf{m}!} \tilde{U}^{(\mathbf{m})}(\mathbf{0}_N; \boldsymbol{\omega}) \mathbb{E}[\mathbf{d}^{\mathbf{m}}]. \quad (26)$$

Finally, the approximate portfolio choice allocation is obtained by solving the following maximization problem

$$\tilde{\boldsymbol{\omega}}_M \in \arg \max_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} \tilde{V}_M(\boldsymbol{\omega}). \quad (27)$$

For implementation purposes, a convenient way to obtain the Taylor expansion above is the following. First, define the deterministic constant  $c_p(\boldsymbol{\omega}) = R_f + \boldsymbol{\omega}'\mathbf{c}$  and the portfolio return deviation  $d_p(\boldsymbol{\omega}) = \boldsymbol{\omega}'\mathbf{d}$  so that the decomposition  $R_p(\boldsymbol{\omega}) = c_p(\boldsymbol{\omega}) + d_p(\boldsymbol{\omega})$  holds. Next, define the utility  $\tilde{u}(d_p(\boldsymbol{\omega}))$  in terms of  $d_p(\boldsymbol{\omega})$  as

$$\tilde{u}(d_p(\boldsymbol{\omega})) = U(R_f + \boldsymbol{\omega}'\mathbf{R}) = \frac{1}{1-\gamma} \left( \left( c_p(\boldsymbol{\omega}) + \frac{\gamma}{\alpha} \right) + d_p(\boldsymbol{\omega}) \right)^{1-\gamma} \quad (28)$$

and use it to obtain the following *one-dimensional* Taylor approximation with respect to  $d_p(\boldsymbol{\omega})$  centered around 0:

$$\tilde{u}_M(d_p(\boldsymbol{\omega})) = \sum_{m=0}^M \frac{1}{m!} \tilde{u}^{(m)}(0) [d_p(\boldsymbol{\omega})]^m. \quad (29)$$

To compute the power  $[d_p(\boldsymbol{\omega})]^m$ , we make use of the multinomial formula

$$[d_p(\boldsymbol{\omega})]^m = (\boldsymbol{\omega}'\mathbf{d})^m = \sum_{|\mathbf{m}|=m} \frac{m!}{\mathbf{m}!} \boldsymbol{\omega}^{\mathbf{m}} \mathbf{d}^{\mathbf{m}}.$$

Hence, we can express the approximate expected utility based on the  $M$ -th order Taylor expansion as

$$\tilde{V}_M(\boldsymbol{\omega}) = \mathbb{E}[\tilde{u}_M(d_p(\boldsymbol{\omega}))] = \sum_{m=0}^M \tilde{u}^{(m)}(0) \sum_{|\mathbf{m}|=m} \frac{1}{\mathbf{m}!} \boldsymbol{\omega}^{\mathbf{m}} \mathbb{E}[\mathbf{d}^{\mathbf{m}}]. \quad (30)$$

Note that the two expressions for  $\tilde{V}_M(\boldsymbol{\omega})$  in (26) and (30) are identical since  $\tilde{U}(\mathbf{d}; \boldsymbol{\omega}) = \tilde{u}(\boldsymbol{\omega}'\mathbf{d})$  and therefore, by the chain rule, we have

$$\tilde{U}^{(\mathbf{m})}(\mathbf{d}; \boldsymbol{\omega}) = \boldsymbol{\omega}^{\mathbf{m}} \tilde{u}^{(|\mathbf{m}|)}(\boldsymbol{\omega}'\mathbf{d}).$$

Before addressing the issue of convergence of the Taylor series approximation, we derive, in the next proposition, some useful properties of the exact expected utility  $V(\boldsymbol{\omega})$ .

**Proposition 3.2** *Let the vector of excess returns  $\mathbf{R}$  satisfy Assumption 3.1. Furthermore, assume there is no redundant asset, i.e., there exists no  $\boldsymbol{\alpha} \in \mathbb{R}^N \setminus \{\mathbf{0}_N\}$  such that  $\mathbb{P}[\boldsymbol{\alpha}'\mathbf{R} = 0] = 1$ . Then,  $V$  is continuous and strictly concave on  $\boldsymbol{\Omega}$ , and hence admits a unique maximum at  $\boldsymbol{\omega}^*$ .*

The following proposition provides sufficient conditions on the distribution of asset returns that guarantee convergence of the approximate optimal portfolio weight vector  $\tilde{\boldsymbol{\omega}}_M$  to the exact optimal portfolio weight vector  $\boldsymbol{\omega}^*$  for an arbitrary (bounded, closed, and convex) set  $\boldsymbol{\Omega}$  of admissible portfolio allocations.

**Proposition 3.3** *If the excess return deviations  $d_n = R_n - c_n$ ,  $n = 1, \dots, N$ , satisfy the condition  $|d_n| \leq D$  with  $D < D^*$  and*

$$D^* = \min_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} \frac{R_f + \frac{\gamma}{\alpha} + \boldsymbol{\omega}'\mathbf{c}}{|\boldsymbol{\omega}'\mathbf{1}_N|}, \quad (31)$$

where  $\mathbf{1}_N$  is the  $N$ -dimensional vector of ones, then the approximate expected utility  $\tilde{V}_M(\boldsymbol{\omega})$  converges to the exact expected utility  $V(\boldsymbol{\omega})$  uniformly with respect to the allocation  $\boldsymbol{\omega}$ . Moreover, the sequence of the approximate optimal weights  $\tilde{\boldsymbol{\omega}}_M$  converges to the exact optimal weight  $\boldsymbol{\omega}^*$ , as  $M \rightarrow \infty$ .

Since, by assumption, the center of expansion  $\mathbf{c}$  belongs to the support of  $\mathbf{R}$  and  $\boldsymbol{\Omega}$  is compact, Assumption 3.1 guarantees that  $D^* > 0$ .

The following corollary provides sufficient conditions under which Proposition 3.3 is applicable. The conditions are not optimal but they are relatively simple and easy to check in practical applications.

**Corollary 3.4** Define  $\underline{\mathbf{R}} = (\underline{R}_1, \dots, \underline{R}_N)'$ ,  $\overline{\mathbf{R}} = (\overline{R}_1, \dots, \overline{R}_N)'$ ,  $\underline{R} = \min \{\underline{R}_n : n = 1, \dots, N\}$ , and  $\overline{R} = \max \{\overline{R}_n : n = 1, \dots, N\}$ . If the condition

$$\frac{\overline{R} - \underline{R}}{2} < \min_{\boldsymbol{\omega} \in \Omega} \frac{R_f + \frac{\gamma}{\alpha} + \boldsymbol{\omega}' \left( \frac{\underline{\mathbf{R}} + \overline{\mathbf{R}}}{2} \right)}{|\boldsymbol{\omega}' \mathbf{1}_N|} \quad (32)$$

is satisfied, then, under the choice  $c_n = \frac{\underline{R}_n + \overline{R}_n}{2}$ ,  $n = 1, \dots, N$ , the condition  $|d_n| \leq D$ ,  $n = 1, \dots, N$  in Proposition 3.3 is satisfied with  $D = \frac{\overline{R} - \underline{R}}{2}$ .

While not available explicitly in closed form, the right hand side in (32) can be computed numerically for a given set  $\Omega$  of admissible allocations. Then one can easily verify whether the condition in (32) is satisfied, in which case using the midpoint of the support of the return distribution as center of the expansion ensures convergence.

### 3.2 Taylor approximation based on a nonlinear decomposition

As discussed in section 2.2, when the distribution of portfolio returns is skewed, one can significantly improve the accuracy of the Taylor approximation to expected utility by using a nonlinear transformation that symmetrizes and shrinks the distribution of portfolio returns. In this section, we apply the same principle in the context of optimal portfolio choice when individual asset returns have skewed distributions.<sup>11</sup> We specify conditions on the shocks to the transformed asset returns that guarantee convergence of the approximate optimal portfolio allocation to its exact counterpart, as the number of Taylor terms tends to infinity. In the next subsection, we illustrate numerically how a nonlinear transformation can improve the accuracy of Taylor series approximation in the presence of skewed asset returns.

Let us implicitly define the *transformed* random variable  $\mathbf{y} = (y_1, \dots, y_N)'$  through the following nonlinear transformation

$$\frac{R_n}{R_f} = A_n + B_n e^{C_n y_n}, \quad n = 1, \dots, N. \quad (33)$$

The excess return  $R_n$  is then given by  $R_n = R_f (A_n + B_n e^{C_n y_n})$ . Note that Assumption 3.1 implies that  $y_n$  also has compact support, which we denote by  $[y_n, \overline{y}_n]$ , for  $n = 1, \dots, N$ . One can specify the values of  $A_n$ ,  $B_n$ , and  $C_n$  following the analysis in subsection 2.2. Specifically, treating  $y_n$  as normally distributed, we can select  $A_n$ ,  $B_n$ , and  $C_n$  to match the first three moments of  $R_n$  and obtain similar expressions as in (18)–(20). Specifically, if  $\mu_n$ ,  $\sigma_n$ ,  $\xi_n$  represent the mean, standard deviation, and skewness of the excess return  $R_n$ ,  $n = 1, \dots, N$ , the appropriate values

---

<sup>11</sup>If the distribution of asset returns is symmetric, then the Taylor expansion based on the linear decomposition considered in the previous subsection clearly provides the optimal approximation scheme.

of  $A_n$ ,  $B_n$ , and  $C_n$  are then determined by

$$A_n = \frac{\mu_n}{R_f} - B_n e^{C_n^2/2}, \quad (34)$$

$$B_n = \text{sign}(\xi_n) \frac{\sigma_n}{R_f \sqrt{e^{C_n^2} (e^{C_n^2} - 1)}}, \quad (35)$$

$$C_n = \sqrt{\log \left( \Gamma_n - 1 + \frac{1}{\Gamma_n} \right)}, \quad \Gamma_n = \left( \frac{\xi_n^2 + \xi_n \sqrt{4 + \xi_n^2 + 2}}{2} \right)^{1/3}. \quad (36)$$

Let  $\kappa_n$  denote the deterministic constant to be used as the center of expansion for the return on the  $n$ -th asset and  $\delta_n = y_n - \kappa_n$  be the corresponding stochastic deviation so that  $y_n = \kappa_n + \delta_n$ . Naturally, we assume  $\kappa_n \in [\underline{y}_n, \bar{y}_n]$ ,  $n = 1, \dots, N$ . Since the goal of the transformation is to symmetrize the distribution, in empirical applications it makes sense to set  $\kappa_n$  equal to the midpoint of  $[\underline{y}_n, \bar{y}_n]$  in order to minimize the magnitude of the deviation  $\delta_n$ . Note that the deviation  $\delta_n$  does not necessarily have mean zero. The portfolio return  $R_p(\boldsymbol{\omega})$  is then expressed as

$$R_p(\boldsymbol{\omega}) = R_f + \boldsymbol{\omega}' \mathbf{R} = R_f \left[ 1 + \sum_{n=1}^N \omega_n \left( A_n + B_n e^{C_n(\kappa_n + \delta_n)} \right) \right] \quad (37)$$

and the corresponding utility as a function of  $\boldsymbol{\delta} \equiv (\delta_1, \dots, \delta_N)'$  and  $\boldsymbol{\omega}$  is given by

$$\widehat{U}(\boldsymbol{\delta}; \boldsymbol{\omega}) = \frac{1}{1-\gamma} R_f^{1-\gamma} \left[ \left( \frac{\gamma}{\alpha R_f} + \left( 1 + \sum_{n=1}^N \omega_n A_n \right) \right) + \sum_{n=1}^N (\omega_n B_n e^{C_n \kappa_n}) e^{C_n \delta_n} \right]^{1-\gamma}. \quad (38)$$

The  $M$ -th order Taylor approximation with respect to  $\boldsymbol{\delta}$  centered around  $\mathbf{0}_N$  is

$$\widehat{U}_M(\boldsymbol{\delta}; \boldsymbol{\omega}) = \sum_{|\mathbf{m}| \leq M} \frac{1}{\mathbf{m}!} \widehat{U}^{(\mathbf{m})}(\mathbf{0}_N; \boldsymbol{\omega}) \boldsymbol{\delta}^{\mathbf{m}} \quad (39)$$

where  $\widehat{U}^{(\mathbf{m})}(\mathbf{0}_N; \boldsymbol{\omega}) = \frac{\partial^{m_1 + \dots + m_N}}{\partial \delta_1^{m_1} \dots \partial \delta_N^{m_N}} \widehat{U}(\mathbf{0}_N; \boldsymbol{\omega})$ . Hence, the approximate expected utility is given by

$$\widehat{V}_M(\boldsymbol{\omega}) = \mathbb{E} \left[ \widehat{U}_M(\boldsymbol{\delta}; \boldsymbol{\omega}) \right] = \sum_{|\mathbf{m}| \leq M} \frac{1}{\mathbf{m}!} \widehat{U}^{(\mathbf{m})}(\mathbf{0}_N; \boldsymbol{\omega}) \mathbb{E} [\boldsymbol{\delta}^{\mathbf{m}}]. \quad (40)$$

The set of feasible portfolio allocation vectors  $\boldsymbol{\Omega}$  is taken to be a finite subset of  $\mathbb{R}^N$ .<sup>12</sup> The exact and approximate optimal portfolio allocations, denoted by  $\boldsymbol{\omega}^*$  and  $\widehat{\boldsymbol{\omega}}_M$ , are obtained as the solutions to the following maximization problems<sup>13</sup>

$$\boldsymbol{\omega}^* \in \arg \max_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} V(\boldsymbol{\omega}), \quad (41)$$

and

$$\widehat{\boldsymbol{\omega}}_M \in \arg \max_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} \widehat{V}_M(\boldsymbol{\omega}). \quad (42)$$

<sup>12</sup>This is a technical assumption required for establishing the convergence result in Proposition 3.5.

<sup>13</sup>Note that, since  $\boldsymbol{\Omega}$  is a finite set, neither the exact nor the approximate maximization problem necessarily have a unique solution.

Proposition 3.5 below, which is the analogue to Proposition 3.3, provides sufficient conditions for the convergence of the Taylor series expansion under the nonlinear transformation (33). The proposition states the general version of the result and is formulated in terms of arbitrary distributions of excess returns and sets of admissible portfolio allocations. Two subsequent corollaries, motivated by empirical relevance, focus on the portfolio choice problem with positively skewed asset returns and no short selling constraints. Before stating the results, we introduce some notation.

Let  $\mathbf{A} = (A_1, \dots, A_N)'$ ,  $L_n = B_n e^{C_n \kappa_n}$ ,  $n = 1, \dots, N$ , and  $\mathbf{L} = (L_1, \dots, L_N)'$ . Define the functions

$$\mathcal{R}_1(w) = \frac{\sqrt{1+w^2}}{w} \arctan(w), \quad w \in (0, \infty), \quad (43)$$

$$\mathcal{R}_2(w) = \frac{\sqrt{1+w^2}}{w} (\pi - \arctan(w)), \quad w \in (0, \infty). \quad (44)$$

and let  $\mathcal{S}_1$  and  $\mathcal{S}_2$  denote the inverse functions of  $\mathcal{R}_1$  and  $\mathcal{R}_2$  defined on  $(1, \frac{\pi}{2})$  and  $(\frac{\pi}{2}, \infty)$ , respectively. Lemma A.5 shows that the functions  $\mathcal{S}_1$  and  $\mathcal{S}_2$  are well-defined. Furthermore, define

$$\mathcal{T}(a, b; x) = e^{ax} \cos(bx), \quad x \in \mathbb{R}, \quad a^2 + b^2 \leq 1, \quad (45)$$

and let

$$\mathcal{E}(x) = \min\{\mathcal{T}(a, b; x) : a^2 + b^2 \leq 1\}, \quad x \in \mathbb{R}, \quad (46)$$

$$\mathcal{M}(x) = \max\{\mathcal{T}(a, b; x) : a^2 + b^2 \leq 1\}, \quad x \in \mathbb{R}. \quad (47)$$

Lemma A.6 shows that

$$\mathcal{E}(x) = \begin{cases} \exp(-|x|), & \text{if } |x| \leq 1, \\ \exp\left(-\frac{1}{\sqrt{1+[\mathcal{S}_1(|x|)]^2}}|x|\right) \cos\left(\frac{\mathcal{S}_1(|x|)}{\sqrt{1+[\mathcal{S}_1(|x|)]^2}}|x|\right), & \text{if } 1 < |x| < \frac{\pi}{2}, \\ 0, & \text{if } |x| = \frac{\pi}{2}, \\ \exp\left(\frac{1}{\sqrt{1+[\mathcal{S}_2(|x|)]^2}}|x|\right) \cos\left(\frac{\mathcal{S}_2(|x|)}{\sqrt{1+[\mathcal{S}_2(|x|)]^2}}|x|\right), & \text{if } |x| > \frac{\pi}{2}, \end{cases} \quad (48)$$

and

$$\mathcal{M}(x) = e^{|x|}, \quad x \in \mathbb{R}. \quad (49)$$

Finally, let  $\gamma > 0$ ,  $\lambda_0 \in \mathbb{R}$ ,  $\lambda^- \leq 0$ , and  $\lambda^+ \geq 0$  where at least one of  $\lambda^-$  or  $\lambda^+$  is different from 0, and assume that  $\lambda = \lambda_0 + \lambda^- + \lambda^+ > 0$ . Denote  $\mathbf{\Lambda} = (\lambda_0, \lambda^-, \lambda^+)'$  and define the function

$$\mathcal{I}(\Delta; \mathbf{\Lambda}) = \lambda_0 + \lambda^+ \mathcal{E}(\Delta) + \lambda^- \mathcal{M}(\Delta), \quad \Delta \in [0, \infty), \quad (50)$$

where  $\mathcal{E}(\cdot)$  and  $\mathcal{M}(\cdot)$  are defined by (46) and (47), and given explicitly by (48) and (49), respectively.

**Proposition 3.5** Let  $\Omega$  be a finite subset of  $\mathbb{R}^N$ . For all  $\omega \in \Omega$ , let

$$\lambda_0(\omega) = 1 + \frac{\gamma}{\alpha R_f} + \omega' \mathbf{A}, \quad \lambda_n(\omega) = \omega_n L_n, \quad n = 1, \dots, N, \quad (51)$$

$$\lambda^-(\omega) = \sum_{n=1}^N \min(\lambda_n(\omega), 0), \quad \lambda^+(\omega) = \sum_{n=1}^N \max(\lambda_n(\omega), 0), \quad (52)$$

and

$$\mathbf{\Lambda}(\omega) = (\lambda_0(\omega), \lambda^-(\omega), \lambda^+(\omega))'. \quad (53)$$

If the deviations  $\delta_n$ ,  $n = 1, \dots, N$ , satisfy the condition  $|C_n \delta_n| \leq \Delta$  where  $0 < \Delta < \Delta^*$  and

$$\Delta^* = \min_{\omega \in \Omega} \mathcal{I}^{-1}(0; \mathbf{\Lambda}(\omega)), \quad (54)$$

where  $\mathcal{I}(\cdot; \mathbf{\Lambda})$  is defined by (50), then the approximate expected utility  $\widehat{V}_M(\omega)$  converges to the exact utility  $V(\omega)$ , for all  $\omega \in \Omega$ . Moreover,  $V(\widehat{\omega}_M) \rightarrow V(\omega^*)$ , i.e., the certainty equivalent loss associated with  $\widehat{\omega}_M$  converges to 0, as  $M \rightarrow \infty$ .

Next, we address the portfolio choice problem with positively skewed asset returns and no short selling constraints. If  $R_n$  is positively skewed, then its skewness  $\xi_n$  is positive and, by equation (35) we have  $B_n > 0$ . Note that this case covers the widely used transformation of log excess returns  $r_n$ , implicitly defined by  $\frac{R_{g,n}}{R_f} = e^{r_n}$ , for  $A_n = -1$  and  $B_n > 0$ .

**Corollary 3.6** Assume that no short selling is allowed, i.e.,

$$\Omega = \{\omega = (\omega_1, \dots, \omega_N)' : 0 \leq \omega_n \leq 1, \quad n = 1, \dots, N, \quad 0 \leq 1 - \omega' \mathbf{1}_N \leq 1\}. \quad (55)$$

Consider the transformed variable  $\mathbf{y} = (y_1, \dots, y_N)'$  introduced in (33) with  $B_n > 0$ , for  $n = 1, \dots, N$ . Then, the bound  $\Delta^*$  in Proposition 3.5 becomes

$$\Delta^* = \min_{\omega \in \Omega} \mathcal{E}^{-1} \left( -\frac{1 + \frac{\gamma}{\alpha R_f} + \omega' \mathbf{A}}{\omega' \mathbf{L}} \right). \quad (56)$$

where  $\mathcal{E}(\cdot)$  is given by (48).

We conclude this subsection by presenting the special case of the preceding corollary for CRRA utility and  $\mathbf{y}$  equal to the log excess return vector  $\mathbf{r} = (r_1, \dots, r_N)'$ , where  $\frac{R_{g,n}}{R_f} = e^{r_n}$ ,  $n = 1, \dots, N$ . In this case,  $\alpha \rightarrow \infty$  and  $A_n = -1$ ,  $n = 1, \dots, N$  and hence since  $\mathcal{E}$  is strictly decreasing we have

$$\min_{\omega \in \Omega} \mathcal{E}^{-1} \left( -\frac{1 + \frac{\gamma}{\alpha R_f} + \omega' \mathbf{A}}{\omega' \mathbf{L}} \right) = \mathcal{E}^{-1} \left( -\min_{\omega \in \Omega} \frac{1 - \omega' \mathbf{1}_N}{\omega' \mathbf{L}} \right) = \mathcal{E}^{-1}(0) = \frac{\pi}{2}$$

where the last equality follows from equation (48). This establishes the following

**Corollary 3.7** *Assume that no short selling is allowed. Then, for CRRA utility, i.e., when  $\alpha \rightarrow \infty$ , and  $\mathbf{y}$  equal to the log excess return vector  $\mathbf{r}$ , the bound  $\Delta^*$  in Proposition 3.5 becomes  $\Delta^* = \frac{\pi}{2}$ .*

We close this subsection by briefly illustrating the practical implications of the condition in the last corollary. Suppose that the support of the distribution of the log excess return on the  $n$ -th asset,  $r_n$ , is  $[\underline{r}_n, \bar{r}_n]$ . Then, the condition in Corollary 3.7 is satisfied if (i)  $\bar{r}_n - \underline{r}_n < \pi$  and (ii) we select the midpoint of the support as the center of expansion:  $\kappa_r = \frac{\underline{r}_n + \bar{r}_n}{2}$ . Note that the condition  $\bar{r}_n - \underline{r}_n < \pi$  is satisfied for a large set of values of  $\underline{r}_n$  and  $\bar{r}_n$  that imply a rather wide support for the primitive asset return distribution. For instance, for  $\underline{r}_n = -1.65$  and  $\bar{r}_n = 1.45$  we have  $\bar{r}_n - \underline{r}_n = 3.1 < \pi$ . Assuming  $R_f = 1.05$  (calibrated to a horizon of one year), we obtain that the implied simple return,  $R_{g,n} - 1$ , lies between  $-80\%$  and  $347\%$ . This examples illustrates that the Taylor approximation converges for a wide range of asset return distributions.

### 3.3 Numerical examples

We conclude this section by providing a few numerical examples to illustrate the practical performance of the Taylor approximation in the context of portfolio choice. We consider the problem of maximizing the expected utility  $\mathbb{E}[U(R_p)]$ , where  $U(\cdot)$  belongs to the CRRA class with risk aversion  $\gamma$ . We assume that the investment opportunity set consists of a risk-free asset and  $N = 3$  risky assets. We consider two types of distribution for the vector of log excess returns  $\mathbf{r} \equiv \log(\mathbf{R}) - \log(R_f)\mathbf{1}_N$ : (i) multivariate normal distribution and (ii) mixture of multivariate normal distributions. The latter is used because it can easily incorporate skewness and excess kurtosis. The parameters that determine the joint distribution of  $\mathbf{r}$  are selected at random to roughly correspond to quarterly equity returns.

When the distribution of  $\mathbf{r}$  is assumed to be multivariate normal, the means and standard deviations of the risky asset log excess returns,  $r_n$ ,  $n = 1, 2, 3$ , are drawn at random independently from the intervals  $[\underline{\mu}, \bar{\mu}] = [0.03, 0.06]$  and  $[\underline{\sigma}, \bar{\sigma}] = [0.10, 0.15]$ , respectively. The correlation matrix of  $\mathbf{r}$  is of the form  $\mathbf{C} = \mathbf{U}\mathbf{U}'$  where  $\mathbf{U}$  is an  $N \times 2N$  matrix, the  $n$ -th row of which,  $\mathbf{u}_n$ , is a unit vector of the form  $\mathbf{u}_n = \mathbf{q}_n / \|\mathbf{q}_n\|$  with  $\mathbf{q}_n$  drawn from the standard  $2N$ -dimensional normal distribution.

When the distribution of  $\mathbf{r}$  is assumed to be a mixture of multivariate normal distributions, we use mixtures with four components. The vector of mixing weights is proportional to  $\exp(\mathbf{u})$  where  $\mathbf{u}$  is a four-dimensional vector of uniform (0,1) numbers drawn at random independently. For  $k = 1, 2, 3, 4$ , the means and standard deviations of the  $k$ -th normal component are drawn from the intervals  $[\underline{\mu}_k, \bar{\mu}_k]$  and  $[\underline{\sigma}_k, \bar{\sigma}_k]$ , respectively. The intervals used in the simulation are:  $[\underline{\mu}_1, \bar{\mu}_1] = [0.01, 0.03]$  and  $[\underline{\sigma}_1, \bar{\sigma}_1] = [0.20, 0.25]$ ,  $[\underline{\mu}_2, \bar{\mu}_2] = [0.03, 0.05]$  and  $[\underline{\sigma}_2, \bar{\sigma}_2] = [0.15, 0.20]$ ,  $[\underline{\mu}_3, \bar{\mu}_3] = [0.05, 0.07]$  and  $[\underline{\sigma}_3, \bar{\sigma}_3] = [0.10, 0.15]$ ,  $[\underline{\mu}_4, \bar{\mu}_4] = [0.07, 0.09]$  and  $[\underline{\sigma}_4, \bar{\sigma}_4] = [0.05, 0.10]$ . The resulting mixtures of normal distributions exhibit skewness and excess kurtosis. The correlation matrix of each component is selected in the same fashion as for the case of the normal distribution described above. Moreover, we set the annualized gross risk-free rate equal to 1.05.

We consider two different types of Taylor approximation. The first, labeled “Excess return decomposition”, is based on the decomposition of the excess return  $R_n = c_n + d_n$ ,  $n = 1, 2, 3$ , while the second, labeled “Log excess return decomposition”, is based on the decomposition of the log excess return  $r_n = \kappa_n + \delta_n$ ,  $n = 1, 2, 3$ . As before, since our theory relies on the assumption of a bounded support of the distribution of shocks, we implement the Taylor approximation to expected utility calculations by using a Gauss-Hermite quadrature approximation, when the distribution of  $\mathbf{r}$  is assumed to be normal. When the distribution of  $\mathbf{r}$  is assumed to be a mixture of normal distributions, we use the same approximation for each component of the mixture. The center of the Taylor expansion is determined as follows. When  $\mathbf{r}$  is assumed to follow a normal distribution, we set  $c_n$  equal to the midpoint of the support of the quadrature-based approximation to the distribution of  $R_n$ , and  $\kappa_n$  equal to the mean of  $r_n$ . When  $\mathbf{r}$  is assumed to follow a mixture of normal distributions, we apply the same procedure to each component of the mixture.

To assess the performance of the Taylor approximation, we compute the associated certainty equivalent losses using the quadrature-based solution as benchmark. In our implementation, we use Gauss-Hermite quadrature with five nodes in each dimension. The certainty equivalent losses, reported in Table 2, are stated in annualized basis points. For different values of the coefficient of risk aversion  $\gamma$  and order  $M$  of the Taylor expansion, the table reports the median certainty equivalent loss across 1,000 simulation repetitions.

The table shows the clear superior performance of the Taylor series expansion based on the log excess return decomposition. The certainty equivalent losses for the log excess return decomposition are uniformly lower than those based on excess return decomposition. Overall, the results are consistent with the results on expected utility presented in subsection 2.3. In particular, we observe that, for both distributional assumptions (normal and mixture of normals) and both decompositions (with respect to excess and log excess return), the certainty equivalent losses are monotonically increasing in the order of the Taylor expansion  $M$ . Finally, the high values of certainty equivalent losses for low values of  $M$  indicates that the quality of low order approximation, e.g., quadratic utility, can be quite poor.

## 4 Portfolio Choice with Unbounded Risky Return Distribution

The Taylor approximation method developed in subsection 3.1 cannot be directly applied to solve the portfolio choice problem if the support of return distribution is unbounded. In this section, we show how one can approximate the solution to such a problem solving a sequence of portfolio choice problems in which the support of the return distribution is suitably truncated. Before doing so, we provide conditions for the existence of a unique solution to the portfolio choice problem when the return distribution has unbounded support.

Consider the static portfolio choice problem for an investor with a strictly concave utility function  $U(\cdot)$  over wealth with domain  $\mathcal{W}$ , taken to be a convex subset of  $\mathbb{R}$ . We assume that  $U(\cdot)$  has a uniform sign; that is, we have either  $U(W) \geq 0$  for all  $W \in \mathcal{W}$  or  $U(W) \leq 0$  for all  $W \in \mathcal{W}$ . This condition is satisfied by the HARA utility function considered in Sections 2

and 3. The investment opportunity set consists of a risk-free asset with gross return  $R_f$  and  $N$  risky assets with excess return vector  $\mathbf{R}$ . The unbounded support of the distribution of the excess return vector  $\mathbf{R}$  is denoted by  $\mathbf{S}$ . Denoting by  $\boldsymbol{\omega}$  the vector of portfolio weights in the risky assets, we can express terminal wealth as  $W_1 = W_0 (R_f + \boldsymbol{\omega}'\mathbf{R})$ , where  $W_0$  is the initial wealth. The set of admissible portfolio weights for which expected utility of terminal wealth is well defined, denoted by  $\boldsymbol{\Omega}^A$ , consists of all  $\boldsymbol{\omega} \in \mathbb{R}^N$  such that (i)  $W_1 = W_0 (R_f + \boldsymbol{\omega}'\mathbf{R}) \in \mathcal{W}$  for all  $\mathbf{R} \in \mathbf{S}$  and (ii)  $\mathbb{E}[|U(W_0 (R_f + \boldsymbol{\omega}'\mathbf{R}))|] < \infty$ . It follows that  $\boldsymbol{\Omega}^A$  is a convex subset of  $\mathbb{R}^N$ . However,  $\boldsymbol{\Omega}^A$  could be open, closed or neither depending on the excess return support  $\mathbf{S}$ . To ensure that the portfolio choice problem admits a solution, we restrict the choice set to be a compact set  $\boldsymbol{\Omega}$  so that  $\boldsymbol{\Omega} \subset \boldsymbol{\Omega}^o$ , where  $\boldsymbol{\Omega}^o$  is the interior of  $\boldsymbol{\Omega}^A$ . Hence, the problem of the investor reads

$$\boldsymbol{\omega}^* \in \arg \max_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} V(\boldsymbol{\omega}) \quad (57)$$

where

$$V(\boldsymbol{\omega}) = \mathbb{E} [U(W_0(R_f + \boldsymbol{\omega}'\mathbf{R}))]. \quad (58)$$

As in Proposition 3.2, we assume there is no redundant risky asset, i.e., there exists no  $\boldsymbol{\alpha} \in \mathbb{R}^N \setminus \{\mathbf{0}_N\}$  such that  $\mathbb{P}[\boldsymbol{\alpha}'\mathbf{R} = 0] = 1$ . Under this assumption, it follows that  $V(\cdot)$  is strictly concave on  $\boldsymbol{\Omega}$  and so, by Theorem 10.4 in Rockafellar (1970), we have that  $V(\cdot)$  is continuous on  $\boldsymbol{\Omega}$ . Hence, there is a unique solution  $\boldsymbol{\omega}^*$  to the portfolio choice problem described in (57)-(58).

We next show how the above solution can be obtained as the limit of a sequence of solutions to portfolio choice problems with return distributions having bounded support. For any positive integer  $K$ , we consider the truncation of the excess return vector  $\mathbf{R}$  on the interval  $[-K, K]$  in each dimension to obtain the truncated random variable  $\mathbf{R}_K$ . That is, we have  $\mathbb{P}[\mathbf{R}_K \in \mathbf{B}] = \mathbb{P}[\mathbf{R} \in \mathbf{B}] / C_K$  for all Borel  $\mathbf{B} \subseteq \mathbf{D}_K$ , where  $\mathbf{D}_K = [-K, K]^N$  and  $C_K = \mathbb{P}[\mathbf{R} \in \mathbf{D}_K]$ . Define the function associated with the truncated problem

$$V_K(\boldsymbol{\omega}) = \mathbb{E} [U(W_0(R_f + \boldsymbol{\omega}'\mathbf{R}_K))] = \frac{1}{C_K} V_K^0(\boldsymbol{\omega}),$$

where

$$V_K^0(\boldsymbol{\omega}) = \mathbb{E} [U(W_0(R_f + \boldsymbol{\omega}'\mathbf{R})) \mathbf{1}_{[\mathbf{R} \in \mathbf{D}_K]}].$$

Let  $\boldsymbol{\omega}_K^*$  denote an optimal solution to the truncated problem, that is

$$\boldsymbol{\omega}_K^* \in \arg \max_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} V_K^0(\boldsymbol{\omega}).$$

Note that an approximation to  $\boldsymbol{\omega}_K^*$  can be obtained using the Taylor approximation approach developed in subsection 3.1. Finally, we illustrate that, as  $K \rightarrow \infty$ , the sequence  $\boldsymbol{\omega}_K^*$  converges to the optimal solution  $\boldsymbol{\omega}^*$  of the portfolio choice problem described in (57)-(58). Using the continuity of the utility function and the dominated convergence theorem, we obtain that  $V_K^0(\boldsymbol{\omega})$  converges to  $V(\boldsymbol{\omega})$  pointwise in  $\boldsymbol{\omega} \in \boldsymbol{\Omega}$ . Since  $U(\cdot) \geq 0$  or  $U(\cdot) \leq 0$ , we have that the convergence is monotonic and therefore by Dini's theorem (see Proposition 11, Section 2, Chapter 9 in Royden (1988)), we have that  $V_K^0(\boldsymbol{\omega})$  converges to  $V(\boldsymbol{\omega})$  uniformly in  $\boldsymbol{\omega} \in \boldsymbol{\Omega}$ . It follows from Lemma A.12 that  $\boldsymbol{\omega}_K^*$  converges to  $\boldsymbol{\omega}^*$ , as  $K \rightarrow \infty$ .

## 5 Conclusion

In this paper, we revisit a number of issues related to Taylor approximations to expected utility. We focus on the HARA utility function that is widely used in economics and finance. In the context of portfolio evaluation, we present conditions under which the approximate expected utility converges to the exact utility and it does so monotonically. Moreover, we illustrate that using Taylor expansion with respect to a variable introduced through an appropriate nonlinear transformation of portfolio return substantially improves the quality of the approximation. In the context of optimal portfolio choice, we provide conditions under which the approximate optimal portfolio allocation converges to its exact counterpart, as the order of the Taylor expansion tends to infinity. Finally, we illustrate that performing the Taylor approximation with respect to a nonlinear transformation of asset returns makes the approximation more accurate. Numerical examples are used to demonstrate the relevance of the theoretical results. The natural extension of the present work, that we plan to pursue in the future, is to study the convergence properties of Taylor series approximations in the context of dynamic portfolio choice problems.

## A Auxiliary Results

In this appendix, we state a number of theorems from complex analysis that are used repeatedly in the subsequent proofs. Furthermore, we state and prove a number of lemmata that are crucial in establishing the theoretical results in the main body of the paper.

In the context of the following four theorems,  $z, z_0$  and  $a_n, b_n, n = 0, 1, \dots$  are complex numbers and  $\mathbb{D}(z_0; r)$  stands for the open disk in the complex plane with center  $z_0$  and radius  $r$ , i.e.,  $\mathbb{D}(z_0; r) = \{z \in \mathbb{C} : |z - z_0| < r\}$ . The open disk with radius 1 centered at 0, referred to as the unit disk, is denoted by  $\mathbb{D}$ , i.e.,  $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ . The closed disk with radius 1 centered at 0 is denoted by  $\overline{\mathbb{D}}$ , i.e.,  $\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leq 1\}$ .

**Theorem A.1** [Theorem 9.20 in Apostol (1974)] *Given a power series  $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ , let  $\lambda = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$  and  $r = \frac{1}{\lambda}$ , (where  $r = 0$  if  $\lambda = +\infty$  and  $r = +\infty$  if  $\lambda = 0$ ). Then the series converges absolutely if  $|z - z_0| < r$  and diverges if  $|z - z_0| > r$ . Furthermore, the series converges uniformly on every compact subset interior to the disk of convergence.*

**Theorem A.2** [Theorem 9.23 in Apostol (1974)] *Assume that the series  $\sum_{n=0}^{\infty} a_n(z - z_0)^n$  converges for each  $z$  in the open disk  $\mathbb{D}(z_0; r)$ . Then the function  $f$  defined by the equation  $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n, z \in \mathbb{D}(z_0; r)$ , has a derivative  $f'(z)$  for each  $z$  on  $\mathbb{D}(z_0; r)$ , given by  $f'(z) = \sum_{n=1}^{\infty} n a_n(z - z_0)^{n-1}$ .*

**Theorem A.3** [Theorem 9.25 in Apostol (1974)] *Consider two power series expansions about the origin, say  $f(z) = \sum_{n=0}^{\infty} a_n z^n$ , where  $z \in \mathbb{D}(0; r)$  and  $g(z) = \sum_{n=0}^{\infty} b_n z^n$ , where  $z \in \mathbb{D}(0; R)$ . If, for a fixed  $z$  in  $\mathbb{D}(0; R)$ , we have  $\sum_{n=0}^{\infty} |b_n z^n| < r$ , then for this  $z$  we can write  $f[g(z)] = \sum_{k=0}^{\infty} c_k z^k$  where the coefficients  $c_k$  are given by  $c_k = \sum_{n=0}^{\infty} a_n b_k(n)$ ,  $k = 0, 1, \dots$  and the numbers  $b_k(n)$  are defined by the equation  $g(z)^n = (\sum_{k=0}^{\infty} b_k z^k)^n = \sum_{k=0}^{\infty} b_k(n) z^k$ .*

**Theorem A.4** [Theorem 16.20 in Apostol (1974)] *Assume that the function  $f$  is analytic on an open set  $S$  in  $\mathbb{C}$ , and let  $z_0$  be any point of  $S$ . Then all derivatives  $f^n(z_0)$  exist, and  $f$  can be represented by the convergent power series  $f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n$  on every open disk  $\mathbb{D}(z_0; R)$  whose closure lies in  $S$ .*

The following two lemmata, proved through standard calculus techniques, will be used in the sequel.

**Lemma A.5** *The functions  $\mathcal{R}_1(\cdot)$  and  $\mathcal{R}_2(\cdot)$  defined in (43) and (44), respectively, satisfy the following properties: (i)  $\mathcal{R}_1$  is strictly increasing,  $\lim_{w \rightarrow 0} \mathcal{R}_1(w) = 1$ , and  $\lim_{w \rightarrow \infty} \mathcal{R}_1(w) = \frac{\pi}{2}$  and (ii)  $\mathcal{R}_2$  is strictly decreasing,  $\lim_{w \rightarrow 0} \mathcal{R}_2(w) = \infty$ , and  $\lim_{w \rightarrow \infty} \mathcal{R}_2(w) = \frac{\pi}{2}$ . Therefore, the inverse functions of  $\mathcal{R}_1$  and  $\mathcal{R}_2$ , denoted by  $\mathcal{S}_1$  and  $\mathcal{S}_2$ , are well-defined on  $(1, \frac{\pi}{2})$  and  $(\frac{\pi}{2}, \infty)$ , respectively. Moreover,  $x \mathcal{S}_2(x) \rightarrow \pi$ , as  $x \rightarrow \infty$ .*

**Proof.** Using L'Hospital's rule, we obtain  $\lim_{w \rightarrow 0} \mathcal{R}_1(w) = \lim_{w \rightarrow 0} \frac{\arctan(w)}{w} = \lim_{w \rightarrow 0} \frac{1}{1+w^2} = 1$ . Moreover,  $\lim_{w \rightarrow \infty} \mathcal{R}_1(w) = \lim_{w \rightarrow \infty} \arctan(w) = \frac{\pi}{2}$ . Similarly, the corresponding limits for

$\mathcal{R}_2$  are  $\lim_{w \rightarrow 0} \mathcal{R}_2(w) = \infty$  and  $\lim_{w \rightarrow \infty} \mathcal{R}_2(w) = \frac{\pi}{2}$ . The derivative of  $\mathcal{R}_1$  and  $\mathcal{R}_2$  are  $\mathcal{R}'_1(w) = \frac{w - \arctan(w)}{w^2 \sqrt{1+w^2}}$  and  $\mathcal{R}'_2(w) = -\frac{\pi + w - \arctan(w)}{w^2 \sqrt{1+w^2}}$ , respectively. Letting  $\mathcal{P}(w) = w - \arctan(w)$ , we have  $\mathcal{P}'(w) = \frac{w^2}{1+w^2} > 0$  and, therefore,  $\mathcal{P}(w) > \mathcal{P}(0) = 0$  for all  $w > 0$ . Hence,  $\mathcal{R}'_1(w) > 0$  and  $\mathcal{R}'_2(w) < 0$  for all  $w > 0$ . It follows that  $\mathcal{R}_1(\cdot)$  is strictly increasing on  $(0, \infty)$  with  $1 < \mathcal{R}_1(w) < \frac{\pi}{2}$  and that  $\mathcal{R}_2(\cdot)$  is strictly decreasing on  $(0, \infty)$  with  $\frac{\pi}{2} < \mathcal{R}_2(w) < \infty$ . Finally, let  $x_n$  be a sequence with  $x_n \rightarrow \infty$  as  $n \rightarrow \infty$ , and define  $w_n = \mathcal{S}_2(x_n)$ . Since  $\mathcal{R}_2(w)$  is strictly decreasing and converges to  $\infty$  as  $w \rightarrow 0$ , we have  $w_n \rightarrow 0$ . Then  $x_n \mathcal{S}_2(x_n) = \mathcal{R}_2(w_n) w_n = \frac{\sqrt{1+w_n^2}}{w_n} (\pi - \arctan(w_n)) w_n \rightarrow \pi$  since  $\lim_{w \rightarrow 0} \arctan(w) = 0$ . This completes the proof. ■

**Lemma A.6** Consider the functions  $\mathcal{E}(\cdot)$  and  $\mathcal{M}(\cdot)$  defined by (46) and (47), respectively. Then, the functions  $\mathcal{E}(\cdot)$  and  $\mathcal{M}(\cdot)$  are explicitly given by expressions (48) and (49), respectively. Furthermore, the function  $\mathcal{E}(x)$  is strictly decreasing in  $|x|$ , and  $\mathcal{E}(x) \rightarrow -\infty$ , as  $|x| \rightarrow \infty$ , while the function  $\mathcal{M}(x)$  is strictly increasing in  $|x|$  and  $\mathcal{M}(x) \rightarrow \infty$ , as  $|x| \rightarrow \infty$ .

**Proof.** The result holds trivially for  $x = 0$ . Let  $x \neq 0$ . The system of equations  $\frac{\partial \mathcal{T}(a,b;x)}{\partial a} = \frac{\partial \mathcal{T}(a,b;x)}{\partial b} = 0$  is equivalent to  $\cos(bx) = \sin(bx) = 0$  which is not feasible. Therefore, the minimum and the maximum of  $\mathcal{T}$  are not attained in the interior of the set  $\{(a,b) : a^2 + b^2 \leq 1\}$  and, hence, they are attained on the boundary. Moreover,  $\mathcal{T}(a,b;0) = 1$  for all  $a$  and  $b$ . Hence,

$$\begin{aligned} \mathcal{E}(x) &= \min\{\tau(a;x) : -1 \leq a \leq 1\}, \\ \mathcal{M}(x) &= \max\{\tau(a;x) : -1 \leq a \leq 1\}, \end{aligned}$$

where

$$\tau(a;x) = e^{ax} \cos\left(x\sqrt{1-a^2}\right).$$

Note that  $\tau(a;x) = \tau(-a;-x)$  and therefore  $\mathcal{E}(x) = \mathcal{E}(-x)$  and  $\mathcal{M}(x) = \mathcal{M}(-x)$ , i.e.,  $\mathcal{E}(\cdot)$  and  $\mathcal{M}(\cdot)$  are even functions. Furthermore, note that  $\mathcal{E}(0) = \mathcal{M}(0) = 1$  and  $\mathcal{E}(-\frac{\pi}{2}) = \mathcal{E}(\frac{\pi}{2}) = 0$ .

The function  $\mathcal{M}(x)$  is easily determined as follows. Note that  $e^{ax} \cos\left(x\sqrt{1-a^2}\right) \leq e^{ax} \leq e^{|x|}$  and that the bound  $e^{|x|}$  is attainable when  $a = \text{sign}(x) \in \{-1, 1\}$ . Therefore, the global maximizer  $\bar{a}(x)$  is given by  $\bar{a}(x) = \text{sign}(x)$  and so  $\mathcal{M}(x) = e^{|x|}$ . Clearly,  $\mathcal{M}(x)$  is strictly increasing in  $|x|$  and  $\mathcal{M}(x) \rightarrow \infty$ , as  $|x| \rightarrow \infty$ .

To determine the function  $\mathcal{E}(\cdot)$ , we need to consider a number of separate cases: (i)  $0 < x \leq 1$ , (ii)  $1 < x < \frac{\pi}{2}$ , and (iii)  $x > \frac{\pi}{2}$ . Note that the derivative of  $\tau$  with respect to  $a$  is given by

$$\tau'(a;x) = xe^{ax} \left[ \cos\left(x\sqrt{1-a^2}\right) + \frac{a}{\sqrt{1-a^2}} \sin\left(x\sqrt{1-a^2}\right) \right], \quad -1 < a < 1.$$

First, consider  $0 < x < \frac{\pi}{2}$  (cases (i) and (ii)). Then, we have  $\cos\left(x\sqrt{1-a^2}\right) \geq 0$  for all  $a \in [-1, 1]$ . Recalling that  $\tau(a;x) = e^{ax} \cos\left(x\sqrt{1-a^2}\right)$ , we observe that, in this case, the global minimizer  $\underline{a}(x)$  has to be negative. To see this, note that  $\tau'(0;x) = x \cos(x) > 0$  and so 0 cannot be a minimizer and further note that if  $a > 0$  then  $\tau(-a;x) < \tau(a;x)$ . If the global minimizer

$\underline{a}(x)$  is in the open interval  $(-1, 0)$ , it should satisfy the FOC  $\tau'(a; x) = 0$ . Otherwise, we should have  $\underline{a}(x) = -1$ .

To identify the global minimizer  $\underline{a}(x)$ , we consider the FOC  $\tau'(a; x) = 0$  where  $a \in (-1, 0)$ . Letting  $w = -\frac{\sqrt{1-a^2}}{a} \in (0, \infty)$ , we have  $a = -\frac{1}{\sqrt{1+w^2}}$  and  $\sqrt{1-a^2} = \frac{w}{\sqrt{1+w^2}}$  and so we can rewrite the FOC in terms of  $w$  as

$$\cos\left(x\frac{w}{\sqrt{1+w^2}}\right) - \sin\left(x\frac{w}{\sqrt{1+w^2}}\right)w = 0 \Leftrightarrow \tan\left(x\frac{w}{\sqrt{1+w^2}}\right) = w.$$

Therefore,  $x\frac{w}{\sqrt{1+w^2}} = \arctan(w) + k\pi$  for some integer  $k$ . But  $x \in (0, \frac{\pi}{2})$  and so  $x\frac{w}{\sqrt{1+w^2}} \in (0, \frac{\pi}{2})$  which implies  $k = 0$  since  $\arctan(w) \in (-\frac{\pi}{2}, \frac{\pi}{2})$ . Hence, we have  $x = \frac{\sqrt{1+w^2}}{w} \arctan(w)$ . Therefore, the FOC  $\tau'(a; x) = 0$  with  $a \in (-1, 0)$  is equivalent to the equation  $x = \mathcal{R}_1(w)$  where  $\mathcal{R}_1$  is defined by (43). But  $\mathcal{R}_1(w) > 1$  for all  $w \in (0, \infty)$  and therefore the FOC does not have a solution in  $(-1, 0)$  if  $x \leq 1$ . Therefore,  $\underline{a}(x) = -1$  for  $x \in (0, 1]$ . If  $1 < x < \frac{\pi}{2}$ , then  $w = \mathcal{S}_1(x)$  where  $\mathcal{S}_1$  is the inverse function of  $\mathcal{R}_1$  defined by (43). Therefore,  $\underline{a}(x) = -\frac{1}{\sqrt{1+[\mathcal{S}_1(x)]^2}}$  for  $x \in (1, \frac{\pi}{2})$ .

Next consider  $x > \frac{\pi}{2}$  (case (iii)). Note that  $\cos(x\sqrt{1-a^2})$  takes negative values as  $a$  ranges from  $-1$  to  $1$ . Recalling that  $\tau(a; x) = e^{ax} \cos(x\sqrt{1-a^2})$  and  $\tau(-a; x) = e^{-ax} \cos(x\sqrt{1-a^2})$ , we observe that, in this case, the global minimizer  $\underline{a}(x)$  has to be positive. Moreover,  $\tau(1; x) = e^x > 0$  and so the global minimizer  $\underline{a}(x)$  should be in the open interval  $(0, 1)$  and it should satisfy the FOC  $\tau'(a; x) = 0$ . Letting  $w = \frac{\sqrt{1-a^2}}{a} \in (0, \infty)$ , we have  $a = \frac{1}{\sqrt{1+w^2}}$  and  $\sqrt{1-a^2} = \frac{w}{\sqrt{1+w^2}}$  and so we can rewrite the FOC in terms of  $w$  as

$$\cos\left(x\frac{w}{\sqrt{1+w^2}}\right) + \sin\left(x\frac{w}{\sqrt{1+w^2}}\right)w = 0 \Leftrightarrow \tan\left(x\frac{w}{\sqrt{1+w^2}}\right) = -w.$$

Therefore, the FOC is equivalent to  $x = \frac{\sqrt{1+w^2}}{w} (-\arctan(w) + k\pi)$  where  $k$  is an integer.

If  $x \in (\frac{\pi}{2}, \frac{3\pi}{2})$ , then we should have  $k = 1$  since  $\arctan(w) \in (-\frac{\pi}{2}, \frac{\pi}{2})$ . In this case, the FOC becomes  $x = \frac{\sqrt{1+w^2}}{w} (\pi - \arctan(w))$ .

Next, let  $x \geq \frac{3\pi}{2}$ . In this case,  $k$  should be a positive integer. Define the function  $\mathcal{J}(a; x) = \cos(x\sqrt{1-a^2})$  for  $a \in (0, 1)$ . The function  $\mathcal{J}(a; x)$  takes the value  $-1$  when  $a = \sqrt{1 - \frac{n^2\pi^2}{x^2}} \in (0, 1)$  where  $n$  is a positive integer with  $n < \frac{x}{\pi}$ . The largest such  $a$  is obtained for  $n = 1$  and is given by  $a_1 = \sqrt{1 - \frac{\pi^2}{x^2}} < 1$ . The derivative of  $\mathcal{J}$  is  $\mathcal{J}'(a; x) = \frac{ax}{\sqrt{1-a^2}} \sin(x\sqrt{1-a^2})$ . For  $a \in (a_1, 1)$ , we have  $x\sqrt{1-a^2} < \pi$  which implies  $\mathcal{J}'(a; x) > 0$ , i.e.,  $\mathcal{J}(\cdot; x)$  is strictly increasing in  $(a_1, 1)$ . Moreover,  $\mathcal{J}(a; x) \rightarrow 1$  as  $a \rightarrow 1$ . Hence, there exists a unique  $a_2$  in  $(a_1, 1)$  with  $\mathcal{J}(a; x) = 0$  so that  $\mathcal{J}(a; x)$  takes values over the entire interval  $(-1, 0)$  as  $a$  ranges over the open interval  $(a_1, a_2)$ . Since  $e^{ax}$  is strictly increasing in  $a$  over  $(0, 1)$ , it follows that the global minimizer  $\underline{a}(x)$  of  $\tau(a; x) = e^{ax} \mathcal{J}(a; x)$  should belong to the interval  $(a_1, a_2)$ . Note that the FOC in terms of  $w$  can be written as  $\frac{w}{\sqrt{1+w^2}}x = (k\pi - \arctan(w))$ . If  $a \in (a_1, a_2)$ , then  $x\sqrt{1-a^2} < \pi$  which implies  $\frac{w}{\sqrt{1+w^2}}x < \pi$ . Hence, for  $a \in (a_1, a_2)$  the FOC is feasible only when  $k = 1$  since  $\arctan(w) \in (-\frac{\pi}{2}, \frac{\pi}{2})$ . It follows that  $\underline{a}(x)$  corresponds to the solution  $w$  of the equation  $x = \frac{\sqrt{1+w^2}}{w} (\pi - \arctan(w))$ .

Hence, in both cases  $x \in (\frac{\pi}{2}, \frac{3\pi}{2})$  and  $x \in (\frac{3\pi}{2}, \infty)$ , the global minimizer  $\underline{a}(x)$  corresponds to the solution  $w$  of the equation  $x = \frac{\sqrt{1+w^2}}{w}(\pi - \arctan(w)) = \mathcal{R}_2(w)$  where  $\mathcal{R}_2$  is defined by (44). Since  $x > \frac{\pi}{2}$ , we can uniquely solve for  $w$  and obtain  $w = \mathcal{S}_2(x)$  where  $\mathcal{S}_2$  is the inverse of  $\mathcal{R}_2$  defined by (44). Therefore,  $\underline{a}(x) = \frac{1}{\sqrt{1+[\mathcal{S}_2(x)]^2}}$  for  $x > \frac{\pi}{2}$ .

Since  $\mathcal{E}(x)$  is defined as the minimum of a continuous function over a compact support, it is continuous. Hence, to show that  $\mathcal{E}(x)$  is strictly decreasing over  $(0, \infty)$ , it suffices to show it is strictly decreasing over the intervals  $(0, 1)$ ,  $(1, \frac{\pi}{2})$ , and  $(\frac{\pi}{2}, \infty)$ . Clearly  $\mathcal{E}(x)$  is strictly decreasing over  $(0, 1)$ . Next, consider the interval  $(1, \frac{\pi}{2})$ . Define the function

$$\mathcal{Q}_1(w) = \log(\mathcal{E}(\mathcal{R}_1(w))), \quad w \in (0, \infty).$$

Since  $\cos(\arctan(w)) = \frac{1}{\sqrt{1+w^2}}$ , we have  $\mathcal{Q}_1(w) = -\frac{\arctan(w)}{w} - \frac{1}{2} \log(1+w^2)$ . It follows that  $\mathcal{Q}'_1(w) = -\frac{w-\arctan(w)}{w^2} < 0$  for all  $w \in (0, \infty)$ . Hence,  $\mathcal{Q}_1(\cdot)$  is strictly decreasing over  $(0, \infty)$ . Let  $1 < x < x' < \frac{\pi}{2}$ . Then, since  $\mathcal{R}_1$  is strictly increasing, there exist  $0 < w < w'$  with  $x = \mathcal{R}_1(w)$  and  $x' = \mathcal{R}_1(w')$ . Then,  $\mathcal{Q}_1(w) > \mathcal{Q}_1(w')$  which implies  $\mathcal{E}(\mathcal{R}_1(w)) > \mathcal{E}(\mathcal{R}_1(w')) \Leftrightarrow \mathcal{E}(x) > \mathcal{E}(x')$ . Hence,  $\mathcal{E}(\cdot)$  is strictly decreasing over  $(1, \frac{\pi}{2})$ . To deal with the interval  $(\frac{\pi}{2}, \infty)$ , define the function

$$\mathcal{Q}_2(w) = \log(-\mathcal{E}(\mathcal{R}_2(w))), \quad w \in (0, \infty).$$

Since  $\cos(\pi - \arctan(w)) = -\frac{1}{\sqrt{1+w^2}}$ , we have  $\mathcal{Q}_2(w) = \frac{\pi - \arctan(w)}{w} - \frac{1}{2} \log(1+w^2)$ . It follows that  $\mathcal{Q}'_2(w) = -\frac{w + \pi - \arctan(w)}{w^2} = -\frac{(w + \pi) - \arctan(w + \pi)}{w^2} < 0$  for all  $w \in (0, \infty)$ . Hence,  $\mathcal{Q}_2(\cdot)$  is strictly decreasing. Let  $\frac{\pi}{2} < x < x'$ . Then, since  $\mathcal{R}_2$  is strictly decreasing, there exist  $0 < w' < w$  with  $x = \mathcal{R}_2(w)$  and  $x' = \mathcal{R}_2(w')$ . Then,  $\mathcal{Q}_2(w') > \mathcal{Q}_2(w)$  which implies  $-\mathcal{E}(\mathcal{R}_2(w')) > -\mathcal{E}(\mathcal{R}_2(w)) \Leftrightarrow \mathcal{E}(x') < \mathcal{E}(x)$ . Hence,  $\mathcal{E}(\cdot)$  is strictly decreasing over  $(\frac{\pi}{2}, \infty)$ . Finally, note that as  $|x| \rightarrow \infty$ , Lemma A.5 shows that  $|x|\mathcal{S}_2(|x|) \rightarrow \pi$  and hence  $\mathcal{S}_2(|x|) \rightarrow 0$ . This limits imply that  $\cos\left(\frac{\mathcal{S}_2(|x|)}{\sqrt{1+[\mathcal{S}_2(|x|)]^2}}|x|\right) \rightarrow -1$  and therefore  $\mathcal{E}(x) \rightarrow -\infty$  as  $|x| \rightarrow \infty$ . This completes the proof. ■

The following lemmata, obtained through complex-analytic techniques, address the convergence of Taylor series approximations to certain functions. The following notation will be used. For  $\mathbf{x} = (x_1, \dots, x_N)' \in \mathbb{R}^N$  and  $\mathbf{m} = (m_1, \dots, m_N)' \in \mathbb{N}^N$ , where  $\mathbb{N}$  is the set of nonnegative integers, define  $\mathbf{x}^{\mathbf{m}} = x_1^{m_1} \dots x_N^{m_N}$ ,  $\mathbf{m}! = m_1! \dots m_N!$ ,  $|\mathbf{m}| = m_1 + \dots + m_N$ . Furthermore, for a function  $G$  defined on  $\mathbb{R}^N$  denote its partial derivatives by  $G^{(\mathbf{m})}(\mathbf{x}) = \frac{\partial^{m_1 + \dots + m_N}}{\partial x_1^{m_1} \dots \partial x_N^{m_N}} G(x)$ .

**Lemma A.7** *Let  $\gamma > 0$  and  $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_N)' \in \mathbb{R}^{N+1}$  where  $\lambda_0 > 0$  and  $\lambda_n \neq 0$  for at least one  $n \in \{1, \dots, N\}$ . Define the quantities*

$$\lambda^- = \sum_{n=1}^N \min(\lambda_n, 0), \quad \lambda^+ = \sum_{n=1}^N \max(\lambda_n, 0),$$

and

$$D^*(\boldsymbol{\lambda}) = \frac{\lambda_0}{\lambda^+ - \lambda^-}.$$

Furthermore, define the function  $G(\cdot)$  on  $\mathbb{R}^N$  by

$$G(\mathbf{x}) = \begin{cases} \frac{1}{1-\gamma}(\lambda_0 + \lambda_1 x_1 + \cdots + \lambda_N x_N)^{1-\gamma} & \text{if } \gamma \neq 1, \\ \log(\lambda_0 + \lambda_1 x_1 + \cdots + \lambda_N x_N) & \text{if } \gamma = 1, \end{cases}$$

where  $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$ . Then, the function  $G(\cdot)$  is well-defined and the Taylor series approximation

$$G_M(\mathbf{x}) = \sum_{|\mathbf{m}| \leq M} \frac{1}{\mathbf{m}!} G^{(\mathbf{m})}(\mathbf{0}_N) \mathbf{x}^{\mathbf{m}}$$

converges uniformly to  $G(\mathbf{x})$ , as  $M \rightarrow \infty$ , for all  $\mathbf{x} \in [-D, D]^N$ , where  $0 < D < D^*(\boldsymbol{\lambda})$ .

**Proof.** First note that for  $\mathbf{x} \in [-D, D]^N$ , we have

$$\lambda_0 + \sum_{n=1}^N \lambda_n x_n \geq \lambda_0 + \lambda^- D + \lambda^+ (-D) = \lambda_0 + D(\lambda^- - \lambda^+) > \lambda_0 + D^*(\boldsymbol{\lambda})(\lambda^- - \lambda^+) = 0$$

and therefore  $G(\cdot)$  is well-defined for  $\mathbf{x} \in [-D, D]^N$ . Define the function  $F_\gamma(t; \mathbf{x})$  for  $t \in [-1, 1]$  by

$$\begin{aligned} F_\gamma(t; \mathbf{x}) &= G(tx_1, \dots, tx_N) \\ &= \begin{cases} \frac{1}{1-\gamma}(\lambda_0 + \lambda_1(tx_1) + \cdots + \lambda_N(tx_N))^{1-\gamma}, & \text{if } \gamma \neq 1, \\ \log(\lambda_0 + \lambda_1(tx_1) + \cdots + \lambda_N(tx_N)), & \text{if } \gamma = 1. \end{cases} \end{aligned}$$

It follows that the  $m$ -th order derivative of  $F_\gamma(t; \mathbf{x})$  with respect to  $t$  is given

$$F_\gamma^{(m)}(t; \mathbf{x}) = \sum_{|\mathbf{m}|=m} \frac{m!}{\mathbf{m}!} G^{(\mathbf{m})}(tx_1, \dots, tx_N) x_1^{m_1} \cdots x_N^{m_N}$$

and the  $M$ -th order Taylor series of  $F_\gamma(\cdot; \mathbf{x})$  centered at 0 is

$$F_{\gamma, M}(t; \mathbf{x}) = \sum_{m=0}^M \frac{1}{m!} F_\gamma^{(m)}(0; \mathbf{x}) t^m = \sum_{m=0}^M \left[ \sum_{|\mathbf{m}|=m} \frac{1}{\mathbf{m}!} G^{(\mathbf{m})}(\mathbf{0}_N) \mathbf{x}^{\mathbf{m}} \right] t^m$$

and so

$$F_{\gamma, M}(1; \mathbf{x}) = G_M(\mathbf{x}).$$

We seek to show that the Taylor series of  $F_\gamma(t; \mathbf{x})$  with respect to  $t$  centered at 0 converges to  $F_\gamma(t; \mathbf{x})$  absolutely and uniformly in  $t \in [-1, 1]$  and  $\mathbf{x} \in [-D, D]^N$ . We claim that it suffices to show the result only for the case  $\gamma = 1$ . Note that when  $\gamma \neq 1$ , we have  $F_\gamma(t; \mathbf{x}) = \frac{1}{1-\gamma} \exp((1-\gamma)F_1(t; \mathbf{x}))$ . Given that the Taylor series expansion of the exponential function converges over the entire real line, the claim follows from Theorem A.3 which provides conditions on the convergence of the Taylor series expansion for a composition of two functions. Next, we focus on the case  $\gamma = 1$  for which  $F_1(t; \mathbf{x}) = \log(\lambda_0 + \lambda_1(tx_1) + \cdots + \lambda_N(tx_N))$ . To show absolute and uniform convergence with respect to  $t$  over  $[-1, 1]$  and  $\mathbf{x}$ , it suffices to show

that the radius of convergence is greater than 1 uniformly in  $\mathbf{x}$ . According to Theorem A.2, the Taylor series of  $F_1(t; \mathbf{x})$  and

$$F_1'(t; \mathbf{x}) = \left[ \sum_{n=1}^N \lambda_n x_n \right] \frac{1}{\lambda_0 + \lambda_1(tx_1) + \cdots + \lambda_N(tx_N)}$$

have the same radius of convergence. Next, define the function  $H$  of the complex variable  $z$  by

$$H(z) = \frac{1}{\lambda_0 + \lambda_1(zx_1) + \cdots + \lambda_N(zx_N)}$$

with domain  $\mathbb{D}_H = \{z \in \mathbb{C} : \lambda_0 + \lambda_1(zx_1) + \cdots + \lambda_N(zx_N) \neq 0\}$ . The function  $H$  is defined as the inverse of an analytic function and, hence, is analytic on  $\mathbb{D}_H$ . Therefore, according to Theorem A.4, the radius of convergence of the Taylor series of  $H$  around 0 is equal to the radius of the largest open ball centered at 0 on which the function  $H$  is defined. Next, we show that, for  $\mathbf{x} \in [-D, D]^N$ , the equation  $L(z) = \lambda_0 + \lambda_1(zx_1) + \cdots + \lambda_N(zx_N) = 0$  has no roots on the closed unit disk  $\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leq 1\}$ . Let  $z = a + ib$  with  $|z| \leq 1$ . Then  $L(z) = 0$  implies  $\lambda_0 + \lambda_1(ax_1) + \cdots + \lambda_N(ax_N) = 0$ . This implies that the function  $G$  is not defined at  $a\mathbf{x}$ . But this is a contradiction since  $|a| \leq |z| \leq 1$  and so  $a\mathbf{x} \in [-D, D]^N$ . Therefore,  $L(z) \neq 0$  for all  $z \in \overline{\mathbb{D}}$  and so the closed unit disk  $\overline{\mathbb{D}}$  is a subset of the domain  $\mathbb{D}_H$ , uniformly in  $\mathbf{x} \in \overline{\mathbb{D}}$ . Hence, the radius of convergence of the Taylor series of  $H$  centered at 0 is greater than 1, uniformly in  $\mathbf{x} \in \overline{\mathbb{D}}$ . The same is true for the real functions  $F_1'(t; \mathbf{x})$  and  $F_1(t; \mathbf{x})$  and the convergence of the Taylor series is absolute and uniform in  $t$  over  $[-1, 1]$  and  $\mathbf{x} \in [-D, D]^N$  according to Theorem A.1. Therefore, it follows that  $F_{\gamma, M}(t; \mathbf{x}) = \sum_{m=0}^M \frac{1}{m!} F_{\gamma}^{(m)}(0; \mathbf{x}) t^m$  converges to  $F_{\gamma}(t; \mathbf{x})$  absolutely and uniformly in  $t$  in  $[-1, 1]$  and  $\mathbf{x} \in [-D, D]^N$ . Hence,  $G_M(\mathbf{x}) = F_{\gamma, M}(1; \mathbf{x})$  converges to  $F_{\gamma}(1; \mathbf{x}) = G(\mathbf{x})$  uniformly in  $\mathbf{x} \in [-D, D]^N$ . This completes the proof. ■

**Lemma A.8** *Let  $\lambda_0$ ,  $\lambda_1$ , and  $\gamma$  be real constants with  $\gamma > 0$ ,  $\lambda_1 \neq 0$ ,  $\lambda_0 + \lambda_1 > 0$  and define the function  $g(\cdot)$  by*

$$g(x) = \begin{cases} \frac{1}{1-\gamma}(\lambda_0 + \lambda_1 e^x)^{1-\gamma} & \text{if } \gamma \neq 1, \\ \log(\lambda_0 + \lambda_1 e^x) & \text{if } \gamma = 1. \end{cases}$$

*Then, the Taylor series approximation  $g_M(x) = \sum_{m=0}^M \frac{1}{m!} g^{(m)}(0) x^m$  converges to  $g(x)$  for all  $x$  with  $|x| < \Delta^*$  where  $\Delta^* = \sqrt{[\log(\lambda_0/\lambda_1)]^2 + \pi^2}$  if  $\lambda_0/\lambda_1 > 0$  and  $\Delta^* = |\log(-\lambda_0/\lambda_1)|$  if  $\lambda_0/\lambda_1 < 0$ . The convergence is uniform on any interval  $[-\Delta, \Delta]$  where  $0 < \Delta < \Delta^*$ .*

**Proof.** First, we claim that it suffices to show the result only for the case  $\gamma = 1$ . Note that when  $\gamma \neq 1$ , we have  $g(x) = \frac{1}{1-\gamma} \exp((1-\gamma) \log(\lambda_0 + \lambda_1 e^x))$ . Given that the Taylor expansion of the exponential function converges over the entire real line, the claim follows from Theorem A.3 which provides conditions on the convergence of the Taylor series approximation for a composition of two functions. Next, we focus on the case  $\gamma = 1$  for which  $g(x) = \log(\lambda_0 + \lambda_1 e^x)$ . According to Theorem A.2, the Taylor series of  $g$  and  $g'$  have the same radius of convergence. Note that  $g'(x) = \frac{\lambda_1 e^x}{\lambda_0 + \lambda_1 e^x}$  and define the function  $h(\cdot)$  on the

complex plane by  $h(z) = \frac{\lambda_1 e^z}{\lambda_0 + \lambda_1 e^z}$ ,  $z \in \mathbb{C}$ . Since  $h$  is defined as a ratio of analytic functions, it is analytic on its domain  $\mathbb{D}_h = \{z \in \mathbb{C} : \lambda_0 + \lambda_1 e^z \neq 0\}$ . Therefore, according to Theorem A.4, the radius of convergence of the Taylor series of  $h$  around the origin is equal to the radius of the largest open ball centered at the origin on which the function  $h$  is defined. Therefore, to find the radius of convergence we need to find the roots of equation  $\frac{\lambda_0}{\lambda_1} + e^z = 0$ . Let  $z = a + ib$  which implies  $e^z = e^a(\cos(b) + i \sin(b))$ . Then,  $\frac{\lambda_0}{\lambda_1} + e^z = 0 \Leftrightarrow e^a \cos(b) = -\frac{\lambda_0}{\lambda_1}$  and  $\sin(b) = 0$ . This is equivalent to  $a = \log(\lambda_0/\lambda_1)$  and  $b = (2n + 1)\pi$ ,  $n = 0, \pm 1, \dots$  if  $\lambda_0/\lambda_1 > 0$  and  $a = \log(-\lambda_0/\lambda_1)$  and  $b = 2n\pi$ ,  $n = 0, \pm 1, \dots$  if  $\lambda_0/\lambda_1 < 0$ . Thus, the roots of the equation  $\frac{\lambda_0}{\lambda_1} + e^z = 0$  that are closest to the origin are  $z_0 = \log(\lambda_0/\lambda_1) \pm i\pi$  if  $\lambda_0/\lambda_1 > 0$  and  $z_0 = \log(-\lambda_0/\lambda_1) + i0$  if  $\lambda_0/\lambda_1 < 0$ . Therefore, the largest open ball centered at the origin that is a subset of the domain  $\mathbb{D}_h$  has radius equal to  $|z_0| = \sqrt{[\log(\lambda_0/\lambda_1)]^2 + \pi^2}$  if  $\lambda_0/\lambda_1 > 0$  and  $|z_0| = |\log(-\lambda_0/\lambda_1)|$  if  $\lambda_0/\lambda_1 < 0$ . This is the radius of convergence of the Taylor series of the function  $h$  on the complex plane and the function  $g'$  on the real line. This completes the proof. ■

The following lemma generalizes the above result to the  $N$ -dimensional case.

**Lemma A.9** *The function  $\mathcal{I}(\cdot; \mathbf{\Lambda})$  defined by (50) satisfies the following properties: (i)  $\mathcal{I}(0; \mathbf{\Lambda}) = \lambda > 0$ , (ii)  $\mathcal{I}(\cdot; \mathbf{\Lambda})$  is strictly decreasing, and (iii)  $\mathcal{I}(\Delta; \mathbf{\Lambda}) \rightarrow -\infty$  as  $\Delta \rightarrow \infty$ . Hence, the inverse of  $\mathcal{I}(\cdot; \mathbf{\Lambda})$  is well-defined on  $(-\infty, \lambda)$ .*

**Proof.** Note that, by assumption,  $\lambda^+ > 0$  or  $\lambda^- < 0$ . From Lemma A.6, we have that  $\mathcal{E}(0) = \mathcal{M}(0) = 1$ ,  $\mathcal{E}(\cdot)$  is strictly decreasing with  $\lim_{\Delta \rightarrow \infty} \mathcal{E}(\Delta) = -\infty$ , and  $\mathcal{M}(\cdot)$  is strictly increasing with  $\lim_{\Delta \rightarrow \infty} \mathcal{M}(\Delta) = \infty$ . Hence,  $\mathcal{I}(0; \mathbf{\Lambda}) = \lambda_0 + \lambda^+ \mathcal{E}(0) + \lambda^- \mathcal{M}(0) = \lambda_0 + \lambda^+ + \lambda^- = \lambda > 0$ ,  $\mathcal{I}(\cdot; \mathbf{\Lambda})$  is strictly decreasing, and  $\lim_{\Delta \rightarrow \infty} \mathcal{I}(\Delta; \mathbf{\Lambda}) = -\infty$ . ■

**Lemma A.10** *Let  $\boldsymbol{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_N)' \in \mathbb{R}^{N+1}$  where  $\lambda = \lambda_0 + \lambda_1 + \dots + \lambda_N > 0$  and  $\lambda_n \neq 0$  for at least one  $n \in \{1, \dots, N\}$ . Define the quantities*

$$\lambda^- = \sum_{n=1}^N \min(\lambda_n, 0), \quad \lambda^+ = \sum_{n=1}^N \max(\lambda_n, 0),$$

and denote  $\mathbf{\Lambda} = (\lambda_0, \lambda^-, \lambda^+)$ . Furthermore, define the function  $G(\cdot)$  on  $\mathbb{R}^N$  by

$$G(\mathbf{x}) = \begin{cases} \frac{1}{1-\gamma} (\lambda_0 + \lambda_1 e^{x_1} + \dots + \lambda_N e^{x_N})^{1-\gamma} & \text{if } \gamma \neq 1, \\ \log(\lambda_0 + \lambda_1 e^{x_1} + \dots + \lambda_N e^{x_N}) & \text{if } \gamma = 1, \end{cases}$$

where  $\mathbf{x} = (x_1, \dots, x_N) \in \mathbb{R}^N$ . Then, the Taylor series approximation

$$G_M(\mathbf{x}) = \sum_{|\mathbf{m}| \leq M} \frac{1}{\mathbf{m}!} G^{(\mathbf{m})}(\mathbf{0}_N) \mathbf{x}^{\mathbf{m}}$$

converges uniformly to  $G(\mathbf{x})$  for all  $\mathbf{x} \in [-\Delta, \Delta]^N$ , where  $0 < \Delta < \Delta^*(\mathbf{\Lambda}) = \mathcal{I}^{-1}(0; \mathbf{\Lambda})$  and  $\mathcal{I}(\cdot; \mathbf{\Lambda})$  is defined in Lemma A.9.

**Proof.** By Lemma A.9,  $\mathcal{I}(\cdot; \mathbf{\Lambda})$  is strictly decreasing and hence we have  $\mathcal{I}(\Delta; \mathbf{\Lambda}) > \mathcal{I}(\Delta^*(\mathbf{\Lambda}); \mathbf{\Lambda}) = 0$  for all  $0 < \Delta < \Delta^*(\mathbf{\Lambda})$ . Let  $\mathbf{x} \in [-\Delta, \Delta]^N$ . Define the function  $F_\gamma(t; \mathbf{x})$  for  $t \in [-1, 1]$  by

$$\begin{aligned} F_\gamma(t; \mathbf{x}) &= G(tx_1, \dots, tx_N) \\ &= \begin{cases} \frac{1}{1-\gamma}(\lambda_0 + \lambda_1 e^{tx_1} + \dots + \lambda_N e^{tx_N})^{1-\gamma}, & \text{if } \gamma \neq 1, \\ \log(\lambda_0 + \lambda_1 e^{tx_1} + \dots + \lambda_N e^{tx_N}), & \text{if } \gamma = 1. \end{cases} \end{aligned}$$

It follows that the  $m$ -th order derivative of  $F_\gamma(t; \mathbf{x})$  with respect to  $t$  is given

$$F_\gamma^{(m)}(t; \mathbf{x}) = \sum_{|\mathbf{m}|=m} \frac{m!}{\mathbf{m}!} G^{(\mathbf{m})}(tx_1, \dots, tx_N) x_1^{m_1} \dots x_N^{m_N}$$

and the  $M$ -th order Taylor series of  $F_\gamma(\cdot; \mathbf{x})$  centered at 0 is

$$F_{\gamma, M}(t; \mathbf{x}) = \sum_{m=0}^M \frac{1}{m!} F_\gamma^{(m)}(0; \mathbf{x}) t^m = \sum_{m=0}^M \left[ \sum_{|\mathbf{m}|=m} \frac{1}{\mathbf{m}!} G^{(\mathbf{m})}(\mathbf{0}_N) \mathbf{x}^{\mathbf{m}} \right] t^m$$

and so

$$F_{\gamma, M}(1; \mathbf{x}) = G_M(\mathbf{x}).$$

We seek to show that the Taylor series of  $F_\gamma(t; \mathbf{x})$  with respect to  $t$  centered at 0 converges to  $F_\gamma(t; \mathbf{x})$  absolutely and uniformly in  $t \in [-1, 1]$  and  $\mathbf{x} \in \mathbb{R}^N$  with  $|x_n| \leq \Delta$ , for  $n = 1, \dots, N$ . We claim that it suffices to show the result only for the case  $\gamma = 1$ . Note that when  $\gamma \neq 1$ , we have  $F_\gamma(t; \mathbf{x}) = \frac{1}{1-\gamma} \exp((1-\gamma)F_1(t; \mathbf{x}))$ . Given that the Taylor series expansion of the exponential function converges over the entire real line, the claim follows from Theorem A.3 which provides conditions on the convergence of the Taylor series expansion for a composition of two functions. Next, we focus on the case  $\gamma = 1$  for which  $F_1(t; \mathbf{x}) = \log(\lambda_0 + \lambda_1 e^{tx_1} + \dots + \lambda_N e^{tx_N})$ . To show absolute and uniform convergence with respect to  $t$  over  $[-1, 1]$  and  $\mathbf{x}$ , it suffices to show that the radius of convergence is greater than 1 uniformly in  $\mathbf{x}$ . According to Theorem A.2, the Taylor series of  $F_1(t; \mathbf{x})$  and

$$F_1'(t; \mathbf{x}) = \sum_{n=1}^N \lambda_n x_n \frac{e^{tx_n}}{\lambda_0 + \lambda_1 e^{tx_1} + \dots + \lambda_N e^{tx_N}}$$

have the same radius of convergence. Next, for  $n = 1, \dots, N$ , define the function  $H_n$  of the complex variable  $z$  by

$$H_n(z) = \frac{e^{zx_n}}{\lambda_0 + \lambda_1 e^{zx_1} + \dots + \lambda_N e^{zx_N}}.$$

with domain  $\mathbb{D}_H = \{z \in \mathbb{C} : \lambda_0 + \lambda_1 e^{zx_1} + \dots + \lambda_N e^{zx_N} \neq 0\}$ . The function  $H_n$  is defined as a ratio of analytic functions and, hence, is analytic on  $\mathbb{D}_H$ . Therefore, according to Theorem A.4, the radius of convergence of the Taylor series of  $H_n$  around 0 is equal to the radius of the largest open ball centered at 0 on which the function  $H_n$  is defined. Letting  $z = a + ib$  we obtain

$$\lambda_0 + \sum_{n=1}^N \lambda_n e^{zx_n} = \left[ \lambda_0 + \sum_{n=1}^N \lambda_n e^{ax_n} \cos(bx_n) \right] + i \left[ \sum_{n=1}^N \lambda_n e^{ax_n} \sin(bx_n) \right].$$

If  $\lambda_n \geq 0$ , then  $\lambda_n e^{ax_n} \cos(bx_n) \geq \lambda_n \mathcal{E}(x_n) \geq \lambda_n \mathcal{E}(\Delta)$ , and if  $\lambda_n < 0$ , then  $\lambda_n e^{ax_n} \cos(bx_n) \geq \lambda_n \mathcal{M}(x_n) \geq \lambda_n \mathcal{M}(\Delta)$ . It follows that  $\lambda_0 + \sum_{n=1}^N \lambda_n e^{ax_n} \cos(bx_n) \geq \lambda_0 + \lambda^+ \mathcal{E}(\Delta) + \lambda^- \mathcal{M}(\Delta) = \mathcal{I}(\Delta; \mathbf{\Lambda}) > 0$ , for  $0 < \Delta < \Delta^*(\mathbf{\Lambda})$ . This implies that the equation  $\lambda_0 + \lambda_1 e^{zx_1} + \dots + \lambda_N e^{zx_N} = 0$  does not have any roots on the closed unit disk  $\overline{\mathbb{D}} = \{z : |z| \leq 1\}$  and so the closed unit disk  $\overline{\mathbb{D}}$  is a subset on the domain  $\mathbb{D}_H$ , uniformly  $\mathbf{x} \in [-\Delta, \Delta]^N$ . Hence, the radius of convergence of the Taylor series of  $H_n$  centered at 0 is greater than 1, uniformly in  $\mathbf{x}$ . The same is true for the real functions  $F'_1(t; \mathbf{x})$  and  $F_1(t; \mathbf{x})$  and the convergence of the Taylor series is absolute and uniform in  $t$  over  $[-1, 1]$  and  $\mathbf{x} \in [-\Delta, \Delta]^N$  according to Theorem A.1. Therefore, it follows that  $F_{\gamma, M}(t; \mathbf{x}) = \sum_{m=0}^M \frac{1}{m!} F_{\gamma}^{(m)}(0; \mathbf{x}) t^m$  converges to  $F_{\gamma}(t; \mathbf{x})$  absolutely and uniformly in  $t$  in  $[-1, 1]$  and  $\mathbf{x} \in [-\Delta, \Delta]^N$ . Hence,  $G_M(\mathbf{x}) = F_{\gamma, M}(1; \mathbf{x})$  converges to  $F_{\gamma}(1; \mathbf{x}) = G(\mathbf{x})$  uniformly in  $\mathbf{x} \in [-\Delta, \Delta]^N$ . ■

**Lemma A.11** *Let  $\mathbf{X}$  be a bounded, closed, and convex subset of  $\mathbb{R}^N$  and  $f : \mathbf{X} \rightarrow \mathbb{R}$  be a continuous and strictly concave function on  $\mathbf{X}$ . Let  $\mathbf{x}^* \in \mathbf{X}$  denote the unique maximizer of  $f$  over  $\mathbf{X}$ , i.e.,  $\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ . If the sequence  $\mathbf{x}_M \in \mathbf{X}$  satisfies  $f(\mathbf{x}_M) \rightarrow f(\mathbf{x}^*)$ , as  $M \rightarrow \infty$ , then  $\|\mathbf{x}_M - \mathbf{x}^*\| \rightarrow 0$ , as  $M \rightarrow \infty$ , for any norm  $\|\cdot\|$  on  $\mathbb{R}^N$ .*

**Proof.** Suppose that  $\mathbf{x}_M$  does not converge to  $\mathbf{x}^*$ . Then, for a given  $\varepsilon > 0$ , there exists a subsequence  $\mathbf{x}_{K(M)}$  such that  $\|\mathbf{x}_{K(M)} - \mathbf{x}^*\| \geq \varepsilon$ ,  $M = 1, 2, \dots$ . Define the subset  $\mathbf{Q}_{\varepsilon}$  of  $\mathbf{X}$  by  $\mathbf{Q}_{\varepsilon} = \{\mathbf{x} \in \mathbf{X} : \|\mathbf{x} - \mathbf{x}^*\| \geq \varepsilon\}$ . Clearly,  $\mathbf{Q}_{\varepsilon}$  is closed and bounded and therefore compact. Hence, there exists a subsequence of  $\mathbf{x}_{K(M)}$ , denoted by  $\mathbf{x}_{L(M)}$ , such that  $\mathbf{x}_{L(M)} \rightarrow \tilde{\mathbf{x}} \in \mathbf{Q}_{\varepsilon}$ , as  $M \rightarrow \infty$ . By the continuity of  $f$ , it follows that  $f(\mathbf{x}_{L(M)}) \rightarrow f(\tilde{\mathbf{x}})$ , as  $M \rightarrow \infty$ . But  $f(\mathbf{x}_{L(M)})$  is a subsequence of  $f(\mathbf{x}_M)$  and so  $f(\mathbf{x}_{L(M)}) \rightarrow f(\mathbf{x}^*)$ , as  $M \rightarrow \infty$ . This is a contradiction, since  $\mathbf{x}^*$  is the unique maximizer of  $f$  on  $\mathbf{X}$  and so  $f(\mathbf{x}^*) > f(\tilde{\mathbf{x}})$ . Therefore, the convergence of  $\mathbf{x}_M$  to  $\mathbf{x}^*$ , as  $M \rightarrow \infty$ , is established. ■

**Lemma A.12** *Let  $\mathbf{X}$  be a bounded, closed, and convex subset of  $\mathbb{R}^N$  and  $f : \mathbf{X} \rightarrow \mathbb{R}$  be a continuous and strictly concave function on  $\mathbf{X}$ . Let  $\mathbf{x}^* \in \mathbf{X}$  denote the unique global maximizer of  $f$  over  $\mathbf{X}$ , i.e.,  $\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathbf{X}} f(\mathbf{x})$ . Moreover, for  $M = 1, 2, \dots$ , let  $f_M : \mathbf{X} \rightarrow \mathbb{R}$  be a sequence of continuous functions. Let  $\mathbf{x}_M^* \in \mathbf{X}$  denote a global maximizer of  $f_M$  on  $\mathbf{X}$ , i.e.,  $\mathbf{x}_M^* \in \arg \max_{\mathbf{x} \in \mathbf{X}} f_M(\mathbf{x})$ . If  $f_M(\cdot)$  converges to  $f(\cdot)$  uniformly in  $\mathbf{x} \in \mathbf{X}$ , then  $\mathbf{x}_M^* \rightarrow \mathbf{x}^*$ , as  $M \rightarrow \infty$ .*

**Proof.** First we show that, as  $M \rightarrow \infty$ ,  $f(\mathbf{x}_M^*) \rightarrow f(\mathbf{x}^*)$ . Note that  $f(\mathbf{x}_M^*) - f(\mathbf{x}^*) = [f(\mathbf{x}_M^*) - f_M(\mathbf{x}_M^*)] + [f_M(\mathbf{x}_M^*) - f(\mathbf{x}^*)]$ . The uniform convergence of  $f_M(\cdot)$  to  $f(\cdot)$  implies that  $f_M(\mathbf{x}_M^*) - f(\mathbf{x}^*) \rightarrow 0$ . Therefore, it suffices to show that  $f(\mathbf{x}_M^*) - f(\mathbf{x}^*) \rightarrow 0$ , as  $M \rightarrow \infty$ . Recall that, by assumption,  $f_M(\mathbf{x}) \rightarrow f(\mathbf{x})$  uniformly in  $\mathbf{x} \in \mathbf{X}$ . Let  $\varepsilon > 0$ . Then, there exists  $M_{\varepsilon}$  such that

$$\max_{\mathbf{x}} |f_M(\mathbf{x}) - f(\mathbf{x})| \leq \varepsilon, \forall M \geq M_{\varepsilon} \Rightarrow -\varepsilon \leq f_M(\mathbf{x}) - f(\mathbf{x}) \leq \varepsilon, \forall \mathbf{x} \in \mathbf{X}, \forall M \geq M_{\varepsilon}.$$

Since  $\mathbf{x}_M^*$  is a global maximizer of  $f_M$ , we have that, for all  $\mathbf{x} \in \mathbf{X}$  and  $M \geq M_\varepsilon$ ,

$$f_M(\mathbf{x}_M^*) + \varepsilon \geq f_M(\mathbf{x}) + \varepsilon \geq f(\mathbf{x}).$$

Hence, since  $\mathbf{x}^*$  is the global maximizer of  $f$ , we also have that, for all  $M \geq M_\varepsilon$ ,  $f_M(\mathbf{x}_M^*) + \varepsilon \geq f(\mathbf{x}^*)$  which implies

$$f(\mathbf{x}^*) - f_M(\mathbf{x}_M^*) \leq \varepsilon. \quad (\text{A1})$$

Moreover, for all  $\mathbf{x} \in \mathbf{X}$  and  $M \geq M_\varepsilon$ , we have

$$f(\mathbf{x}^*) + \varepsilon \geq f(\mathbf{x}) + \varepsilon \geq f_M(\mathbf{x}).$$

It follows that, for all  $M \geq M_\varepsilon$ ,  $f(\mathbf{x}^*) + \varepsilon \geq f_M(\mathbf{x}_M^*)$  which implies

$$-\varepsilon \leq f(\mathbf{x}^*) - f_M(\mathbf{x}_M^*). \quad (\text{A2})$$

Therefore, combining (A1) and (A2), we obtain  $-\varepsilon \leq f(\mathbf{x}^*) - f_M(\mathbf{x}_M^*) \leq \varepsilon$ , for all  $M \geq M_\varepsilon$  which establishes that  $f(\mathbf{x}_M^*) \rightarrow f(\mathbf{x}^*)$ , as  $M \rightarrow \infty$ . By assumption, the function  $f$  is continuous, strictly concave, and is uniquely maximized at  $\mathbf{x}^*$ . Therefore, applying Lemma A.11 yields that  $\mathbf{x}_M^*$  converges to  $\mathbf{x}^*$ , as  $M \rightarrow \infty$ , with respect to any norm  $\|\cdot\|$  on  $\mathbb{R}^N$ . ■

## B Proofs

**Proof of Proposition 2.2.** We claim that it suffices to show the result only for the case  $\gamma = 1$ . To see this, note that when  $\gamma \neq 1$ , we have

$$u(d_p) = \frac{1}{1-\gamma} \left( c_p + \frac{\gamma}{\alpha} + d_p \right)^{1-\gamma} = \frac{1}{1-\gamma} \exp \left( (1-\gamma) \log \left( c_p + \frac{\gamma}{\alpha} + d_p \right) \right).$$

Given that the Taylor expansion of the exponential function converges over the entire real line, the claim follows from Theorem A.3 which provides conditions on the convergence of the Taylor series approximation for a composition of two functions. Next, we focus on the case  $\gamma = 1$  for which  $u(d_p) = \log \left( c_p + \frac{\gamma}{\alpha} + d_p \right)$ . According to Theorem A.2, the Taylor series of  $u$  and  $u'$  have the same radius of convergence. Note that  $u'(d_p) = \frac{1}{c_p + \frac{\gamma}{\alpha} + d_p}$  and define the function  $h(\cdot)$  on the complex plane by  $h(z) = \frac{1}{c_p + \frac{\gamma}{\alpha} + z}$ ,  $z \in \mathbb{C}$ . Since  $h$  is defined as the reciprocal of an analytic function, it is analytic on its domain  $\mathbb{D}_h = \{z \in \mathbb{C} : c_p + \frac{\gamma}{\alpha} + z \neq 0\}$ . Therefore, according to Theorem A.4, the radius of convergence of the Taylor series of  $h$  around the origin is equal to the radius of the largest open ball centered at the origin on which the function  $h$  is defined. The only root of the equation  $c_p + \frac{\gamma}{\alpha} + z = 0$  is  $z_0 = -\left(c_p + \frac{\gamma}{\alpha}\right)$  and therefore the radius of convergence is  $|z_0| = c_p + \frac{\gamma}{\alpha}$ . Therefore, the Taylor approximation of the function  $u(d_p)$  also has radius of convergence equal to  $c_p + \frac{\gamma}{\alpha}$ . According to Theorem A.1, the convergence is uniform when  $|d_p| \leq D_p$  where  $0 < D_p < c_p + \frac{\gamma}{\alpha}$ . Finally, note that, as  $M \rightarrow \infty$ ,

$$|\mathbb{E}[u_M(d_p)] - \mathbb{E}[u(d_p)]| \leq \mathbb{E}[|u_M(d_p) - u(d_p)|] \leq \max_{|d_p| \leq D_p} |u_M(d_p) - u(d_p)| \rightarrow 0.$$

This completes the proof. ■

**Proof of Corollary 2.3.** Define

$$\zeta = \min \left\{ \underline{R}_p + \frac{\gamma}{\alpha}, 2c_p - \left( \overline{R}_p - \frac{\gamma}{\alpha} \right) \right\} > 0$$

and choose  $D_p = c_p + \frac{\gamma}{\alpha} - \zeta < c_p + \frac{\gamma}{\alpha}$ . From the definition of  $\zeta$ , we have  $\underline{R}_p \geq \zeta - \frac{\gamma}{\alpha}$  and  $\overline{R}_p \leq 2c_p - \zeta + \frac{\gamma}{\alpha}$ . Therefore,

$$d_p = R_p - c_p \geq \underline{R}_p - c_p \geq \zeta - \frac{\gamma}{\alpha} - c_p = -D_p$$

and

$$d_p = R_p - c_p \leq \overline{R}_p - c_p \leq c_p + \frac{\gamma}{\alpha} - \zeta = D_p$$

implying that the condition  $|d_p| \leq D_p$  holds. Finally, from Assumption 2.1, we have  $\underline{R}_p > -\frac{\gamma}{\alpha}$  and hence the condition  $c_p \geq \frac{1}{2}(\underline{R}_p + \overline{R}_p)$  implies the condition  $c_p > \frac{1}{2}(\overline{R}_p - \frac{\gamma}{\alpha})$ . ■

**Proof of Proposition 2.4.** The  $m$ -th term in the approximate expected utility (6) is given by

$$\tau_m = \frac{1}{m!} \tilde{\mathcal{U}}^{(m)}(0) \mathbb{E} [d_p^m] = \frac{(-\gamma)_{m-1}}{m!} \left( c_p + \frac{\gamma}{\alpha} \right)^{1-\gamma-m} \mathbb{E} [d_p^m].$$

Letting  $\ell = \frac{L}{c_p + \frac{\gamma}{\alpha}} > 1$ , we observe that

$$\mathbb{E} [d_p^{2m}] \geq \mathbb{E} [d_p^{2m} \mathbf{1}_{\{|d_p| \geq L\}}] \geq L^{2m} \mathbb{P} [|d_p| \geq L] = \ell^{2m} \left( c_p + \frac{\gamma}{\alpha} \right)^{2m} q.$$

It follows that

$$|\tau_{2m}| \geq \frac{\gamma(\gamma+1)\cdots(\gamma+2m-2)}{(2m)!} \left( c_p + \frac{\gamma}{\alpha} \right)^{1-\gamma} \ell^{2m} q \rightarrow \infty$$

since  $\ell > 1$ . Hence, the sequence  $\tau_m$  does not converge to zero and, as a result, the series representing the approximate expected utility in (6) is not convergent. ■

**Proof of Proposition 2.6.** The expression for the radius of convergence as well as the uniform convergence on any interval  $[-\Delta_p, \Delta_p]$  where  $0 < \Delta_p < \Delta_p^*$  and  $\Delta_p^*$  is defined by (17), are immediate consequences of Lemma A.8 upon using  $\lambda_0 = \gamma/\alpha + A$ ,  $\lambda_1 = Be^{C\kappa_y}$  and  $x = C\delta_y$ . ■

**Proof of Proposition 3.2.** To show continuity of  $V$ , fix  $\omega \in \Omega$  and consider a sequence  $\omega_M \in \Omega$  with  $\omega_M \rightarrow \omega$ , as  $M \rightarrow \infty$ . Then  $V(\omega_M) = \mathbb{E} [U(R_f + \omega'_M \mathbf{R})]$ . Since, by assumption 3.1, the support of  $\mathbf{R}$ , denoted by  $\mathbf{S}$ , is compact and  $\omega_M \rightarrow \omega$ , it follows that  $R_f + \omega'_M \mathbf{R}$  lies on a compact set, for all  $\mathbf{R} \in \mathbf{S}$  and all  $M = 1, 2, \dots$ . The continuity of  $U$  then implies that there exists a bound  $K > 0$  such that  $|U(R_f + \omega'_M \mathbf{R})| \leq K$  for all  $\mathbf{R} \in \mathbf{S}$  and all  $M = 1, 2, \dots$ . Hence, we can apply the dominated convergence theorem to pass the limit inside the expectation and obtain  $\lim_{M \rightarrow \infty} V(\omega_M) = \lim_{M \rightarrow \infty} \mathbb{E} [U(R_f + \omega'_M \mathbf{R})] = \mathbb{E} [\lim_{M \rightarrow \infty} U(R_f + \omega'_M \mathbf{R})] = \mathbb{E} [U(R_f + \omega' \mathbf{R})] = V(\omega)$ . Hence, continuity of  $V$  is established.

To show strict concavity of  $V$ , let  $\lambda \in (0, 1)$  and  $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2 \in \boldsymbol{\Omega}$  with  $\boldsymbol{\omega}_1 \neq \boldsymbol{\omega}_2$ . Then, by the no-redundancy condition, it follows that  $\mathbb{P}[\mathbf{R} \in \mathbf{H}] > 0$  where  $\mathbf{H} = \{\mathbf{x} \in \mathbb{R}^N : (\boldsymbol{\omega}_1 - \boldsymbol{\omega}_2)' \mathbf{x} \neq 0\}$ . Strict concavity of  $U$  implies that

$$U(\lambda(R_f + \boldsymbol{\omega}'_1 \mathbf{R}) + (1 - \lambda)(R_f + \boldsymbol{\omega}'_2 \mathbf{R})) \geq \lambda U(R_f + \boldsymbol{\omega}'_1 \mathbf{R}) + (1 - \lambda)U(R_f + \boldsymbol{\omega}'_2 \mathbf{R})$$

for all  $\mathbf{R} \in \mathbf{S}$  while the inequality is strict for  $\mathbf{R} \in \mathbf{H}$ . Since  $\mathbb{P}[\mathbf{R} \in \mathbf{H}] > 0$ , taking expectations yields  $V(\lambda \boldsymbol{\omega}_1 + (1 - \lambda) \boldsymbol{\omega}_2) > \lambda V(\boldsymbol{\omega}_1) + (1 - \lambda)V(\boldsymbol{\omega}_2)$  and so strict concavity of  $V$  is established. ■

**Proof of Proposition 3.3.** According to equation (24), we have

$$\tilde{U}(\mathbf{d}; \boldsymbol{\omega}) = \frac{1}{1 - \gamma} \left( \left( R_f + \frac{\gamma}{\alpha} + \boldsymbol{\omega}' \mathbf{c} \right) + \sum_{n=1}^N \omega_n d_n \right)^{1 - \gamma}.$$

Let  $\lambda_0(\boldsymbol{\omega}) = R_f + \frac{\gamma}{\alpha} + \boldsymbol{\omega}' \mathbf{c}$  and  $\lambda_n(\boldsymbol{\omega}) = \omega_n$ ,  $n = 1, \dots, N$ . Since, by assumption,  $c_n \in [\underline{R}_n, \bar{R}_n]$ ,  $n = 1, \dots, N$ , then Assumption 3.1 implies that  $\lambda_0(\boldsymbol{\omega}) \geq \epsilon > 0$ . Hence, for a given portfolio allocation  $\boldsymbol{\omega}$ , applying Lemma A.7 with  $G = \tilde{U}(\cdot; \boldsymbol{\omega})$  and  $\mathbf{x} = \mathbf{d}$ , we obtain that the approximate utility  $\tilde{U}_M(\mathbf{d}; \boldsymbol{\omega})$  converges to the exact utility  $\tilde{U}(\mathbf{d}; \boldsymbol{\omega})$  uniformly in  $\mathbf{d} \in [-D, D]^N$  with  $0 < D < D^*(\boldsymbol{\omega})$  where  $D^*(\boldsymbol{\omega}) = \frac{\lambda_0(\boldsymbol{\omega})}{\sum_{n=1}^N |\omega_n|} = \frac{R_f + \frac{\gamma}{\alpha} + \boldsymbol{\omega}' \mathbf{c}}{|\boldsymbol{\omega}' \mathbf{1}_N|}$ . It follows that  $\tilde{U}_M(\mathbf{d}; \boldsymbol{\omega})$  converges to  $\tilde{U}(\mathbf{d}; \boldsymbol{\omega})$  uniformly in  $\mathbf{d} \in [-D, D]^N$ , with  $0 < D < D^*$ , where  $D^*$  is defined by equation (31). This, in turn, implies that the approximate expected utility  $\tilde{V}_M(\boldsymbol{\omega})$  converges to the exact expected utility  $V(\boldsymbol{\omega})$ , pointwise in  $\boldsymbol{\omega}$ , as  $M \rightarrow \infty$ . To show that the convergence is uniform in  $\boldsymbol{\omega} \in \boldsymbol{\Omega}$  we proceed as follows. Observe that

$$\begin{aligned} \left| \tilde{V}_M(\boldsymbol{\omega}) - V(\boldsymbol{\omega}) \right| &= \left| \sum_{m=M+1}^{\infty} \frac{1}{m!} \tilde{u}^{(m)}(0) \mathbb{E}[(\boldsymbol{\omega}' \mathbf{d})^m] \right| \\ &\leq \sum_{m=M+1}^{\infty} \frac{1}{m!} \left| \tilde{u}^{(m)}(0) \right| \cdot (|\boldsymbol{\omega}' \mathbf{1}_N|)^m \cdot D^m, \end{aligned}$$

and

$$\begin{aligned} \left| \tilde{u}^{(m)}(0) \right| \cdot (|\boldsymbol{\omega}' \mathbf{1}_N|)^m &\leq |(-\gamma)_{m-1}| \cdot \left[ \left( c_p(\boldsymbol{\omega}) + \frac{\gamma}{\alpha} \right) \right]^{1 - \gamma - m} \cdot (|\boldsymbol{\omega}' \mathbf{1}_N|)^m \\ &= \left( R_f + \boldsymbol{\omega}' \mathbf{c} + \frac{\gamma}{\alpha} \right)^{-\gamma - 1} \cdot |(-\gamma)_{m-1}| \cdot \left( \frac{|\boldsymbol{\omega}' \mathbf{1}_N|}{R_f + \boldsymbol{\omega}' \mathbf{c} + \frac{\gamma}{\alpha}} \right)^m \\ &\leq \epsilon^{-\gamma - 1} \cdot |(-\gamma)_{m-1}| \cdot \left( \frac{1}{D^*} \right)^m \equiv \xi_m, \end{aligned}$$

where the last inequality follows from Assumption 3.1. Hence, for all  $\boldsymbol{\omega} \in \boldsymbol{\Omega}$ ,

$$\left| \tilde{V}_M(\boldsymbol{\omega}) - V(\boldsymbol{\omega}) \right| \leq \sum_{m=M+1}^{\infty} v_m,$$

where

$$v_m = \frac{1}{m!} \cdot \xi_m \cdot D^m.$$

It follows that

$$\frac{v_{m+1}}{v_m} = \frac{1}{m+1} \cdot \frac{\xi_{m+1}}{\xi_m} \cdot D = \frac{1}{m+1} \cdot \frac{|\gamma + m - 1|}{D^*} \cdot D \rightarrow \frac{D}{D^*} < 1,$$

as  $m \rightarrow \infty$ . Hence, by the ratio test, the series  $\sum_{m=M+1}^{\infty} v_m$  converges, which, in turn, yields that, as  $M \rightarrow \infty$ ,

$$\max_{\boldsymbol{\omega} \in \Omega} \left| \tilde{V}_M(\boldsymbol{\omega}) - V(\boldsymbol{\omega}) \right| \leq \sum_{m=M+1}^{\infty} v_m \rightarrow 0,$$

and, hence,  $\tilde{V}_M(\boldsymbol{\omega})$  converges to  $V(\boldsymbol{\omega})$ , uniformly in  $\boldsymbol{\omega}$ . The convergence of  $\tilde{\boldsymbol{\omega}}_M$  to  $\boldsymbol{\omega}^*$  then follows from Lemma A.12. ■

**Proof of Corollary 3.4.** Note that

$$|d_n| = |R_n - c_n| = \left| R_n - \frac{R_n + \bar{R}_n}{2} \right| \leq \frac{\bar{R}_n - R_n}{2} \leq \frac{\bar{R} - \underline{R}}{2} = D.$$

Therefore, it follows from condition (32) that Proposition 3.3 is applicable. ■

**Proof of Proposition 3.5.** According to (38), we obtain

$$\hat{\mathcal{U}}(\boldsymbol{\delta}; \boldsymbol{\omega}) = \frac{1}{1-\gamma} R_f^{1-\gamma} \left[ \left( \frac{1+\gamma}{\alpha R_f} + \boldsymbol{\omega}' \mathbf{A} \right) + \sum_{n=1}^N (\omega_n L_n) e^{C_n \delta_n} \right]^{1-\gamma},$$

where  $\mathbf{A} = (A_1, \dots, A_N)'$  and  $L_n = B_n e^{C_n \kappa_n}$ ,  $n = 1, \dots, N$ . Let  $\lambda_0(\boldsymbol{\omega}) = 1 + \frac{\gamma}{\alpha R_f} + \boldsymbol{\omega}' \mathbf{A}$ ,  $\lambda_n(\boldsymbol{\omega}) = \omega_n L_n$ ,  $n = 1, \dots, N$ ,  $\lambda^-(\boldsymbol{\omega}) = \sum_{n=1}^N \min(\lambda_n(\boldsymbol{\omega}), 0)$ ,  $\lambda^+(\boldsymbol{\omega}) = \sum_{n=1}^N \max(\lambda_n(\boldsymbol{\omega}), 0)$ , and  $\boldsymbol{\Lambda}(\boldsymbol{\omega}) = (\lambda_0(\boldsymbol{\omega}), \lambda^-(\boldsymbol{\omega}), \lambda^+(\boldsymbol{\omega}))$ . Note that  $\lambda_0(\boldsymbol{\omega}) + \lambda_1(\boldsymbol{\omega}) + \dots + \lambda_N(\boldsymbol{\omega})$  equals  $\frac{\gamma}{\alpha}$  plus the portfolio return realization for  $\delta_n = 0$  or equivalently  $y_n = \kappa_n$ ,  $n = 1, \dots, N$ . Since  $\kappa_n \in [\underline{y}_n, \bar{y}_n]$ , Assumption 3.1 implies that  $\sum_{n=0}^N \lambda_n(\boldsymbol{\omega}) \geq \epsilon > 0$ . Hence, for a given  $\boldsymbol{\omega} \in \Omega$ , applying Lemma A.10 with  $G = \hat{\mathcal{U}}$  and  $x_n = C_n \delta_n$ ,  $n = 1, \dots, N$ , we obtain that the approximate utility  $\hat{\mathcal{U}}_M(\boldsymbol{\delta}; \boldsymbol{\omega})$  converges to the exact utility  $\hat{\mathcal{U}}(\boldsymbol{\delta}; \boldsymbol{\omega})$ , uniformly in  $\boldsymbol{\delta} \in [-\Delta, \Delta]^N$ , where  $0 < \Delta < \Delta^*(\boldsymbol{\Lambda}(\boldsymbol{\omega})) = \mathcal{I}^{-1}(0; \boldsymbol{\Lambda}(\boldsymbol{\omega}))$ . Note that  $\mathcal{I}^{-1}(0; \boldsymbol{\Lambda}(\boldsymbol{\omega}))$  is well-defined since, according to Lemma A.9,  $\mathcal{I}^{-1}(\cdot; \boldsymbol{\Lambda}(\boldsymbol{\omega}))$  is defined on  $(-\infty, \lambda(\boldsymbol{\omega}))$  where  $\lambda(\boldsymbol{\omega}) = \lambda_0(\boldsymbol{\omega}) + \lambda^-(\boldsymbol{\omega}) + \lambda^+(\boldsymbol{\omega}) = \sum_{n=0}^N \lambda_n(\boldsymbol{\omega}) \geq \epsilon > 0$ . It follows that  $\hat{\mathcal{U}}_M(\boldsymbol{\delta}; \boldsymbol{\omega})$  converges to  $\hat{\mathcal{U}}(\boldsymbol{\delta}; \boldsymbol{\omega})$  uniformly in  $\boldsymbol{\delta} \in [-\Delta, \Delta]^N$ , with  $0 < \Delta < \Delta^*$  where  $\Delta^*$  is defined by equation (54). This, in turn, implies that the approximate expected utility  $\hat{V}_M(\boldsymbol{\omega})$  converges to the exact expected utility  $V(\boldsymbol{\omega})$ , for all  $\boldsymbol{\omega} \in \Omega$ , as  $M \rightarrow \infty$ . Finally, adapting the argument in the proof of Lemma A.12 to the case of a finite choice set  $\Omega$  yields that  $V(\hat{\boldsymbol{\omega}}_M) \rightarrow V(\boldsymbol{\omega}^*)$ , i.e., the certainty equivalent loss associated with  $\hat{\boldsymbol{\omega}}_M$  converges to 0, as  $M \rightarrow \infty$ . ■

**Proof of Corollary 3.6.** Under the stated assumptions and using the notation in Proposition 3.5, we have  $\lambda_n(\boldsymbol{\omega}) = \omega_n L_n \geq 0$ ,  $n = 1, \dots, N$ . Hence,  $\lambda^-(\boldsymbol{\omega}) = 0$ ,  $\lambda^+(\boldsymbol{\omega}) = \boldsymbol{\omega}' \mathbf{L}$ , and

$\mathcal{I}(\Delta; \mathbf{\Lambda}(\boldsymbol{\omega})) = \left(1 + \frac{\gamma}{\alpha R_f} + \boldsymbol{\omega}' \mathbf{A}\right) + (\boldsymbol{\omega}' \mathbf{L}) \mathcal{E}(\Delta)$ . It follows that

$$\mathcal{I}^{-1}(0; \mathbf{\Lambda}(\boldsymbol{\omega})) = \mathcal{E}^{-1} \left( -\frac{1 + \frac{\gamma}{\alpha R_f} + \boldsymbol{\omega}' \mathbf{A}}{\boldsymbol{\omega}' \mathbf{L}} \right)$$

and, therefore, the proof is complete. ■

## C Tables

Table 1: Correlation between rankings based on exact and approximate expected utility

The table reports the Spearman correlations between the exact and approximate expected utilities based on the rankings of 11 portfolios. The utility function is CRRA with risk aversion  $\gamma$ . The log portfolio returns  $r_p$  are assumed to follow either a normal distribution or a mixture of normal distributions. The Taylor approximate expected utility calculations are based on the Gauss-Hermite quadrature approximation when the distribution of  $r_p$  is assumed to be normal. When the distribution of  $r_p$  is assumed to be a mixture of normal distributions, we use the same approximation for each component of the mixture. In all cases, we use five nodes in the quadrature approximation. The parameters that determine the distribution of  $r_p$  are selected at random to roughly correspond to quarterly equity returns. The exact selection procedure is detailed in subsection 2.3. We consider two Taylor approximations. The first is based on the decomposition of the gross portfolio return  $R_p = c_p + d_p$  while the second is based on the decomposition of the log portfolio return  $r_p = \kappa_r + \delta_r$ . When  $r_p$  is assumed to follow a normal distribution, we set  $c_p$  equal to the midpoint of the support of the quadrature-based approximation to the distribution of  $R_p$ , and  $\kappa_r$  equal to the mean of  $r_p$ . When  $r_p$  is assumed to follow a mixture of normal distributions, we apply the same procedure to each component of the mixture. The reported Spearman correlations are the medians across 10,000 simulation repetitions.

	Order of approximation				
	$M = 2$	$M = 4$	$M = 6$	$M = 8$	$M = 10$
Normal distribution					
Gross return decomposition					
$\gamma = 2$	0.98	1.00	1.00	1.00	1.00
$\gamma = 4$	0.91	0.99	1.00	1.00	1.00
$\gamma = 6$	0.79	0.98	1.00	1.00	1.00
$\gamma = 8$	0.64	0.95	0.99	1.00	1.00
$\gamma = 10$	0.48	0.93	0.99	1.00	1.00
Log return decomposition					
$\gamma = 2$	1.00	1.00	1.00	1.00	1.00
$\gamma = 4$	1.00	1.00	1.00	1.00	1.00
$\gamma = 6$	0.99	1.00	1.00	1.00	1.00
$\gamma = 8$	0.97	1.00	1.00	1.00	1.00
$\gamma = 10$	0.95	0.99	1.00	1.00	1.00
Mixture of normal distributions					
Gross return decomposition					
$\gamma = 2$	0.84	0.98	1.00	1.00	1.00
$\gamma = 4$	0.33	0.90	0.97	0.99	1.00
$\gamma = 6$	-0.15	0.70	0.92	0.97	0.99
$\gamma = 8$	-0.45	0.38	0.80	0.93	0.96
$\gamma = 10$	-0.59	0.04	0.61	0.85	0.93
Log return decomposition					
$\gamma = 2$	1.00	1.00	1.00	1.00	1.00
$\gamma = 4$	0.99	1.00	1.00	1.00	1.00
$\gamma = 6$	0.95	0.99	1.00	1.00	1.00
$\gamma = 8$	0.91	0.98	0.99	1.00	1.00
$\gamma = 10$	0.85	0.95	0.98	0.99	1.00

Table 2: Certainty equivalent loss from Taylor approximation

The table reports the certainty equivalent loss incurred by the approximate solution based on Taylor approximation using the quadrature-based solution as benchmark. The utility function is CRRA with risk aversion  $\gamma$ . There is one risk-free asset and 3 risky assets. The log excess returns are assumed to jointly follow either a normal distribution or a mixture of normal distributions. When log excess returns follow a normal distribution, the Taylor approximate expected utility calculations are based on the Gauss-Hermite quadrature approximation to the normal distribution using five nodes in each dimension. When log excess returns are assumed to follow a mixture of normal distributions, we apply the same approximation for each component of the mixture. The parameters of the log excess return multivariate distribution are selected at random as roughly correspond to quarterly equity returns. The exact selection procedure is detailed in subsection 3.3. We consider two Taylor approximations. The first is based on the decomposition of the risky asset excess return  $R_n = c_n + d_n$ ,  $n = 1, 2, 3$ , while the second is based on the decomposition of the log excess risky asset return  $r_n = \kappa_n + \delta_n$ ,  $n = 1, 2, 3$ . When log excess returns are assumed to follow a normal distribution, we set  $c_n$  equal to the midpoint of the support of the quadrature-based approximation to the distribution of  $R_n$ , and  $\kappa_n$  equal to the mean of  $r_n$ . When log excess returns are assumed to follow a mixture of normal distributions, we apply the same procedure to each component of the mixture. The reported certainty equivalent losses are stated in annualized basis points and are computed as the medians across 1,000 simulation repetitions.

	Order of approximation			
	$M = 2$	$M = 4$	$M = 6$	$M = 8$
Normal distribution				
	Excess return decomposition			
$\gamma = 2$	1.1	0.0	0.0	0.0
$\gamma = 4$	3.6	0.0	0.0	0.0
$\gamma = 6$	7.6	0.1	0.0	0.0
$\gamma = 8$	14.9	0.2	0.0	0.0
$\gamma = 10$	29.2	0.5	0.0	0.0
	Log excess return decomposition			
$\gamma = 2$	0.0	0.0	0.0	0.0
$\gamma = 4$	0.0	0.0	0.0	0.0
$\gamma = 6$	0.0	0.0	0.0	0.0
$\gamma = 8$	0.1	0.0	0.0	0.0
$\gamma = 10$	0.2	0.0	0.0	0.0
Mixture of normal distributions				
	Excess return decomposition			
$\gamma = 2$	2.2	0.0	0.0	0.0
$\gamma = 4$	7.2	0.2	0.0	0.0
$\gamma = 6$	27.7	1.6	0.2	0.0
$\gamma = 8$	148.7	87.1	5.7	0.3
$\gamma = 10$	386.7	307.9	10.7	0.4
	Log excess return decomposition			
$\gamma = 2$	0.0	0.0	0.0	0.0
$\gamma = 4$	0.1	0.0	0.0	0.0
$\gamma = 6$	0.4	0.0	0.0	0.0
$\gamma = 8$	6.4	0.0	0.0	0.0
$\gamma = 10$	11.8	0.0	0.0	0.0

## References

- Apostol, T. M., 1974, *Mathematical Analysis*, Addison-Wesley.
- Arrow, K. J., 1964, “The Role of Securities in the Optimal Allocation of Risk-Bearing,” *Quarterly Journal of Economics*, 31, 91–96.
- Borch, C., 1974, “The Rationale of the Mean-Standard Deviation Analysis: Comment,” *American Economic Review*, 64, 428–430.
- Box, G. E. P., and D. R. Cox, 1964, “An Analysis of Transformations,” *Journal of the Royal Statistical Society. Series B*, 26, 211–252.
- Brandt, M. W., A. Goyal, P. Santa-Clara, and J. R. Stroud, 2005, “A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning about Return Predictability,” *Review of Financial Studies*, 18, 831–873.
- Campbell, J. Y., Y. L. Chan, and L. M. Viceira, 2003, “A Multivariate Model of Strategic Asset Allocation,” *Journal of Financial Economics*, 67, 41–80.
- Campbell, J. Y., and L. M. Viceira, 1999, “Consumption and Portfolio Decisions when Expected Returns are Time Varying,” *Quarterly Journal of Economics*, 114, 433–495.
- Campbell, J. Y., and L. M. Viceira, 2002, *Strategic Asset Allocation*, Oxford University Press, New York.
- Draper, N. R., and D. R. Cox, 1969, “On Distributions and their Transformation to Normality,” *Journal of the Royal Statistical Society. Series B*, 31, 472–476.
- Feldstein, M. S., 1969, “Mean-Variance Analysis in the Theory of Liquidity Preference and Portfolio Selection,” *Review of Economic Studies*, 36, 5–12.
- Garlappi, L., and G. Skoulakis, 2010, “Solving Consumption and Portfolio Choice Problems: The State Variable Decomposition Method,” *Review of Financial Studies*, 23, 3346–3400.
- Glasserman, P., 2004, *Monte Carlo Methods in Financial Engineering*, Springer-Verlag.
- Hinkley, D. V., 1975, “On Power Transformations to Symmetry,” *Biometrika*, 62, 101–111.
- Hlawitschka, W., 1994, “The Empirical Nature of Taylor-Series Approximations to Expected Utility,” *American Economic Review*, 84, 713–719.
- Jondeau, E., and M. Rockinger, 2006, “Optimal Portfolio Allocation under Higher Moments,” *European Financial Management*, 12, 29–55.
- Judd, K., 1998, *Numerical Methods in Economics*, MIT Press.

- Jurczenko, E., and B. Maillet, 1996, "Theoretical Foundations of Asset Allocation and Pricing Models with Higher-order Moments," in Emmanuel Jurczenko, and Bertrand Maillet (ed.), *Multi-moment Asset Allocation and Pricing Models*. pp. 1–36, John Wiley & Sons.
- Koo, H., 1999, "Consumption and portfolio selection with labor income: A discrete-time approach," *Mathematical Methods of Operations Research*, 50, 219–243.
- Kroll, Y., H. Levy, and H. M. Markowitz, 1984, "Mean-Variance Versus Direct Utility Maximization," *Journal of Finance*, 39, 47–75.
- Levy, H., and H. M. Markowitz, 1979, "Approximating Expected Utility by a Function of Mean and Variance," *American Economic Review*, 69, 308–317.
- Loistl, O., 1976, "The Erroneous Approximation of Expected Utility by Means of a Taylor's Series Expansion: Analytic and Computational Results," *American Economic Review*, 66, 904–910.
- Marcet, A., and K. Singleton, 1999, "Equilibrium Asset Prices and Savings of Heterogeneous Agents in the Presence of Incomplete Markets and Portfolio Constraints," *Macroeconomic Dynamics*, 3, 243–277.
- Markowitz, H. M., 1991, "Foundations of Portfolio Theory," *Journal of Finance*, 46, 469–477.
- Pratt, J. W., 1964, "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122–136.
- Pulley, L. B., 1981, "A General Mean-Variance Approximation to Expected Utility for Short Holding Periods," *Journal of Financial and Quantitative Analysis*, 16, 361–373.
- Pulley, L. B., 1983, "Mean-Variance Approximations to Expected Logarithmic Utility," *Operations Research*, 31, 685–696.
- Rockafellar, R. T., 1970, *Convex Analysis*, Princeton University Press.
- Royden, H. L., 1988, *Real Analysis*, Prentice Hall, 3rd edn.
- Rust, J., 1997, "Using Randomization to Break the Curse of Dimensionality," *Econometrica*, 65, 487–516.
- Sakia, R. M., 1992, "The Box-Cox Transformation Technique: A Review," *The Statistician*, 41, 169–178.
- Samuelson, P. A., 1970, "The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Moments," *Review of Economic Studies*, 37, 537–542.
- Savits, T. H., 2006, "Some Statistical Applications of Faà di Bruno," *Journal of Multivariate Analysis*, 97, 2131–2140.

- Tauchen, G., and R. Hussey, 1991, "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59, 371–396.
- Tsiang, S. C., 1972, "The Rationale of the Mean-Standard Deviation Analysis, Skewness Preference, and the Demand for Money," *American Economic Review*, 62, 354–371.
- Wu, L., 2003, "Jumps and Dynamic Asset Allocation," *Review of Quantitative Finance and Accounting*, 20, 207–243.