Taylor Series Approximations to Expected Utility and Optimal Portfolio Choice*

Lorenzo Garlappi† Georgios Skoulakis‡

December 8, 2008

Abstract

This paper revisits the subject of Taylor series approximations to expected utility and investigate the applicability of the technique to optimal portfolio choice problems. We first provide conditions under which the approximate expected utility of 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@mccombs.utexas.edu or gskoulak@rhsmith.umd.edu.

†Finance Department, McCombs School of Business, University of Texas at Austin, Austin, TX 78712; Email: lorenzo.garlappi@mccombs.utexas.edu.

‡Finance Department, Robert H. Smith School of Business, University of Maryland, College Park, MD 20742; Email: 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.[1] 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[2] and Monte Carlo simulation,[3] 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 (2008)). 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.

The construction of a Taylor approximation to expected utility involves two steps: first, the utility of future wealth (or portfolio return) is approximated by a polynomial in the *deviations* of future realized return from a chosen *center of expansion*, with the coefficients depending on the derivatives of the utility function; second, the expectation of this polynomial approximation is computed. If the series expansion in the first step is converging uniformly, the second step reduces the computation of expected utility to the computation of moments of the distribution of deviations, which are simply "displaced" versions of the primitive asset returns. These moments can often be computed analytically or numerically via, for instance, quadrature methods (Tauchen and Hussey (1991), Judd (1998)) or Monte Carlo simulations (Glasserman (2004)). The main computational advantage of Taylor approximations is the fact that the computation of these moments can be kept separate from the control variable, providing a significant gain in computational efficiency, especially in dynamic portfolio choice problems (see Garlappi and Skoulakis (2008)).

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

[1] 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.

[2] See, for instance, Judd (1998) for a comprehensive textbook treatment and Tauchen and Hussey (1991) for applications to asset pricing models.

[3] See the textbooks of Judd (1998) and Glasserman (2004).

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 expected utility, the literature to date provides remarkably little guidance for a correct implementation of these techniques to the solution of portfolio choice problems.

The purpose of this paper is to provide a set of general conditions under which Taylor series can be used as a sound computational tool for both the evaluation of expected utility of a given portfolio and the solution to a portfolio choice problem. Given a utility function, the convergence of a Taylor approximation to expected utility depends on the properties of the distribution of future wealth (or asset returns). Intuitively, the "smaller" the deviations of the shocks to wealth from the chosen center of the Taylor expansion, the more likely that the series converges. In this paper, we provide precise conditions to make this intuition rigorous. Our results are derived for the class of HARA utility functions and are valid for any distribution of the shocks to asset returns, provided that it has a bounded support, a rather mild requirement for returns over a finite time period.

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 to convergence to the exact expected utility. Moreover, by choosing a sufficiently high 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.

The key to a successful implementation a Taylor approximation to expected utility 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 a 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. Through numerical examples, we illustrate that Taylor expansions applied to symmetrically distributed returns provide more accurate approximations than expansions applied to skewed return distributions. This happens because using symmetric distributions allows to minimize the maximum deviations from the center of expansion. The use of such transformations can help resolve the difficulties associated with applying Taylor series approximation to problems with skewed returns.

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. Under the mild assumption of bounded asset returns, we determine sufficient 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 before, 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, instead, they should be avoided. We finally show that, when solving a portfolio choice problem with skewed asset returns, the accuracy of the Taylor approximation can be improved if one takes as primitive in the expansion not the actual (skewed) returns but, instead, a suitable nonlinear transformation designed to reduce the degree of asymmetry and shrink the support of the transformed variables. This change-of-variable approach relates to, and can be viewed as a 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)). Note that loglinearization amounts to a first-order Taylor expansion with respect to the logarithm of the variable of interest. The approach in this paper, instead, 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, Taylor approximations applied to the transformed asset returns converge faster and are more accurate.

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

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

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

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

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

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

Because $\widetilde{\mathcal{U}}(d_p) = U(R_p)$ and $\widetilde{\mathcal{U}}(0) = U(c_p)$, 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

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

where

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

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

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

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

[5] 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 $U(W_1) = W_0^{1-\gamma}\left[1/(1-\gamma)\left(R_p + \gamma/\alpha W_0\right)^{1-\gamma}\right]$, where the quantity in square bracket is of the HARA form (1).

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 $\underline{R}_p > -\frac{\gamma}{\alpha}$.*

Note that, under Assumption 2.1, the utility function (2) is well-defined. The assumption of compact support on the portfolio return might seem restrictive and at odds with common modeling practice that incorporates distributions with infinite support, such as normal. However, our results are still relevant for two reasons. First, the return distribution for a specific investment horizon, e.g., annual, can reasonably be assumed to have bounded support. Second, even under the assumption of infinite support for asset returns, one can obtain an arbitrarily close approximation by first truncating the return distribution in the extreme tails and then applying the Taylor approximation.

The following proposition 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.

Proposition 2.2 *Let R_p satisfy Assumption 2.1, c_p be the center of expansion of the Taylor approximation to expected 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 $\widetilde{\mathcal{U}}_M(d_p)$ converges to $\widetilde{\mathcal{U}}(d_p)$ uniformly in d_p and the approximate expected utility $\mathbb{E}[\widetilde{\mathcal{U}}_M(d_p)]$ converges to the exact expected utility $\mathbb{E}[\widetilde{\mathcal{U}}(d_p)]$, as $M \to \infty$.*

It is important to stress that the conditions of the proposition are *sufficient* but not necessary. To illustrate the point, consider the case of exponential utility, $U(R_p) = -e^{-\alpha R_p}$, and assume that the portfolio return R_p is normally distributed with mean μ_p and variance σ_p^2. The expected utility then is $\mathbb{E}[U(R_p)] = -e^{-\alpha\mu_p + \frac{1}{2}\alpha^2\sigma_p^2}$. If we choose μ_p as the center of expansion, the M-th order Taylor approximation to the expected utility is given by $\mathbb{E}[U_M(R_p)] = \sum_{m=0}^{M} \frac{1}{m!}(-1)^{m+1}\alpha^m e^{-\alpha\mu_p}\mathbb{E}[d_p^m]$. Given the normality assumption, it follows that the odd moments of d_p are zero and the even moments of d_p are given by

$$\mathbb{E}[d_p^{2k}] = \sigma_p^{2k}(2k-1)!!$$

where $(2k-1)!! = 1 \cdot 3 \cdots (2k-1)$. Hence, $\frac{1}{(2k)!}\mathbb{E}[d_p^{2k}] = \frac{1}{k!}\frac{1}{2^k}\sigma_p^{2k}$ and therefore

$$\mathbb{E}[U_{2K}(R_p)] = -e^{-\alpha\mu_p}\sum_{k=0}^{K}\frac{1}{k!}\left[\frac{\alpha^2\sigma_p^2}{2}\right]^k \to -e^{-\alpha\mu_p + \frac{1}{2}\alpha^2\sigma_p^2} = \mathbb{E}[U(R_p)],$$

as $K \to \infty$, despite the fact that R_p does not satisfy Assumption 2.1.

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 $c_p > \frac{1}{2}\left(\overline{R}_p - \frac{\gamma}{\alpha}\right)$, then there exists a bound $D_p < c_p + \frac{\gamma}{\alpha}$ so that Proposition 2.2 applies. Moreover, $c_p \geq \frac{1}{2}\left(\underline{R}_p + \overline{R}_p\right)$ implies $c_p > \frac{1}{2}\left(\overline{R}_p - \frac{\gamma}{\alpha}\right)$.*

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 $\frac{1}{2}\left(\underline{R}_p + \overline{R}_p\right)$ equals the expected portfolio return μ_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 μ_p is larger than the midpoint $\frac{1}{2}\left(\underline{R}_p + \overline{R}_p\right)$. In these two cases, selecting the mean μ_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 μ_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, then the portfolio return takes values outside the radius of convergence with positive probability and, as the next proposition shows, the Taylor approximation to expected utility does not converge to the exact expected utility.[6]

Proposition 2.4 *Let R_p satisfy Assumption 2.1, c_p be the center of expansion of the Taylor approximation to expected utility (2), and $d_p = R_p - c_p$ be the associated portfolio return deviation. If $\mathbb{P}\left[|d_p| \geq L\right] = q > 0$, 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

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

[6] Technically, the sufficient condition for divergence in Proposition 2.4 is *not* exactly the complement of the sufficient condition for convergence in Proposition 2.2. In fact, the sufficient condition for convergence is that $\mathbb{P}[|d_p| \leq D_p] = 1$ for some $D_p < c_p + \frac{\gamma}{\alpha}$ while the sufficient condition for divergence is that $\mathbb{P}[|d_p| \geq L] > 0$ for some $L > c_p + \frac{\gamma}{\alpha}$. If $\mathbb{P}[|d_p| = c_p + \frac{\gamma}{\alpha}] > 0$ and $\mathbb{P}[|d_p| > c_p + \frac{\gamma}{\alpha}] = 0$, then it is not clear whether the Taylor series converges or diverges.

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.[7]

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 $K = 1, 2, \ldots$,*

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

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

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

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 $K = 1, 2 \ldots$. 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 $K = 1, \ldots, L$. 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.

[7] A similar condition is derived in Jondeau and Rockinger (2006).

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 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 expected utility (2), the above intuition suggests working with the following transformation

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

To proceed with the Taylor expansion, we decompose the transformed variable r_p as $r_p = \kappa_r + \delta_r$, where κ_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 δ_r does not necessarily have mean zero. From equation (2), the utility and its derivatives, defined in terms of the deviation δ_r, are[8]

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

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

and the M-th order Taylor approximation is

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

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 δ_r. It follows that, if $|\delta_r|$ is bounded by *any* positive real number Δ, then the approximate expected utility based on a Taylor series with respect to r_p converges to the exact expected utility.[9]

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

$$R_p = A + Be^{Cy}. \tag{14}$$

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

[8] We introduce the notation $\widehat{\ }$, as opposed to $\widetilde{\ }$, to differentiate between the linear decomposition in 2.1 and the present nonlinear decomposition.

[9] The assumption of bounded support for δ_r, or equivalently r_p, is *sufficient* but not necessary for convergence of the approximate expected utility. To see why, consider the case of power utility ($\alpha = \infty$) with $r_p = \log(R_p)$ being normally distributed. Then, following the argument made in the discussion after Proposition 2.2, one can show that the approximate expected utility $\mathbb{E}[U_{2K}(r_p)]$ converges monotonically from above to the exact expected utility $\mathbb{E}[U(r_p)]$, as $K \to \infty$.

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 κ_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 δ_y is

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

and the M-th order Taylor approximation with respect to δ_y is

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

Note that $\widehat{\mathcal{U}}^{(m)}(0)$ 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 (2008) 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 δ_y with $|\delta_y| < \Delta_p^*$ where*

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

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

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 (μ_p, σ_p, ξ_p) of the mean, standard deviation, and skewness of R_p, we can find A, B, and C such that

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

Solving the system for A, B, and C yields

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

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

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

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 (2008)), 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 rather computationally efficient, an aspect that becomes evident in dynamic portfolio choice applications.

We consider the problem of assessing the expected utility $E[U(R_p)]$, where $U(\cdot)$ belongs to the CRRA class with risk aversion γ (special case of the HARA class in (2), when $\alpha \to \infty$). We assume that the portfolio gross return $R_p = \exp(r_p)$ is log-normally distributed. Note that this assumption implies that the portfolio return R_p is positively skewed. We consider two Taylor approximations and present the results in Table 1. The first, 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 $r_p = \kappa_r + \delta_r$. We choose the center of expansion c_p to be the midpoint of the support of R_p truncated at its 0.5 and 99.5 percentiles, and κ_r equal to the mean of r_p.

Because utility is an ordinal concept and unique only up to positive affine transformations, 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 rankings. High values of the correlation, i.e., close to 1, imply better quality of the approximation.

Table 1 reports the Spearman correlations between the ranks of 11 portfolios based on exact and approximate expected utility calculations. Each portfolio return distribution is determined by randomly drawing the mean μ_R and the standard deviation σ_R of R_p from the intervals $[\underline{\mu}, \overline{\mu}]$ and $[\underline{\sigma}, \overline{\sigma}]$. We consider two different calibrations: one for quarterly equity returns and one for annual equity returns. The intervals used for the moments are chosen to roughly match the corresponding frequency. Specifically, for the case of quarterly returns, we use $[\underline{\mu}, \overline{\mu}] = [1.03, 1.06]$ and $[\underline{\sigma}, \overline{\sigma}] = [0.10, 0.15]$, while for the case of annual returns, we use $[\underline{\mu}, \overline{\mu}] = [1.10, 1.20]$ and $[\underline{\sigma}, \overline{\sigma}] = [0.20, 0.30]$. For different values of the coefficient of risk aversion γ and order M of the Taylor expansion, the table reports the average Spearman correlation across 10,000 simulation repetitions.

Overall, the most striking result emerging from the table is the clear superior performance of the Taylor series expansion based on the log return decomposition. All the correlation coefficients for this decomposition are larger than those obtained by using the gross return decomposition. This should not be surprising, in light of our earlier discussion. Because gross returns, in this experiment, are lognormal (and hence positively skewed), the log returns are normal and hence symmetric. This reduces the magnitude of the potential errors incurred in the implementation of the Taylor approximation. Naturally, all else being equal, the approximation is deteriorating as the curvature of the utility function increases (i.e., it is worse for higher γ), and as the horizon of the portfolio return increases (i.e., it is worse for annual data). 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. This issue is not present in the case of log return decomposition because the distribution of log returns is normal and hence symmetric. Finally, the low level of correlation for low values of M indicates that the quality of low order expansions, shown by Hlawitschka (1994) in the case of CRRA preferences with risk aversion between 0.1 and 0.9, easily deteriorates when one considers higher, and more plausible, levels of risk aversion.

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 (2008)). 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 $R_{g,n}$, $n = 1, \ldots, N$. The vector of *excess* returns is denoted by $\mathbf{R} = (R_1, \ldots, R_N)'$ where $R_n = R_{g,n} - R_f$. The allocation to the risky assets is denoted by $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)'$, the portfolio return is $R_p(\boldsymbol{\omega}) = R_f + \boldsymbol{\omega}'\mathbf{R}$ and the resulting HARA utility is given by

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

We assume throughout this section that the set $\boldsymbol{\Omega}$ of admissible portfolio allocation vectors is a bounded, closed, and convex subset of $\mathbb{R}^N$. Our objective is to solve the portfolio choice problem

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

where

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

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 $n = 1, \ldots, N$, the support of the distribution of the excess return R_n is the closed interval $\left[\underline{R}_n, \overline{R}_n\right]$. 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 ϵ 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 $\mathbf{c} = (c_1, \ldots, c_N)'$ is a vector of deterministic constants to be used as the center of the Taylor expansion and $\mathbf{d} = (d_1, \ldots, d_N)'$ is the corresponding vector of stochastic deviations. Naturally, we assume that $c_n \in \left[\underline{R}_n, \overline{R}_n\right]$. We then define the utility $\widetilde{\mathcal{U}}(\mathbf{d}; \boldsymbol{\omega})$ in terms of the deviation vector $\mathbf{d}$ as follows

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

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

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

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

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

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

$$\widetilde{\boldsymbol{\omega}}_M = \arg\max_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} \widetilde{V}_M(\boldsymbol{\omega}). \tag{27}$$

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

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

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

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

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

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

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

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

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

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

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 allocation to its exact counterpart, when no short selling is allowed. In the case of the risk-free asset, no short selling simply means no borrowing.

Proposition 3.3 (No short selling) *Let* $\mathbf{R}$ *satisfy Assumption 3.1. Suppose that no short selling is allowed, i.e.,* $0 \leq \omega_n \leq 1$ *and* $0 \leq \sum_{n=1}^{N} \omega_n \leq 1$*. If for all* $n = 1, \ldots, N$ *the excess return deviations* $d_n = R_n - c_n$ *satisfy the condition* $|d_n| \leq D_n$ *with* $D_n < D_n^* = R_f + c_n + \frac{\gamma}{\alpha}$*, then the approximate expected utility* $\widetilde{V}_M(\boldsymbol{\omega})$ *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* $\widetilde{\boldsymbol{\omega}}_M$ *converges to the exact optimal weight* $\boldsymbol{\omega}^*$*, as* $M \to \infty$*.*

Note that the convergence of $\widetilde{\boldsymbol{\omega}}_M$ to $\boldsymbol{\omega}$ in the above proposition holds with respect to any generic norm $\|\cdot\|$ on $\mathbb{R}^N$. One immediate implication of the convergence of portfolio allocations $\widetilde{\boldsymbol{\omega}}_M$ is that the certainty equivalent loss associated with $\widetilde{\boldsymbol{\omega}}_M$ converges to zero, as $M \to \infty$.

If no short selling is allowed and Assumption 3.1 holds, the utility function (21) is well-defined. The following corollary is the analogue to Corollary 2.3 and provides a practical way of selecting the center of expansion in order to guarantee the convergence in Proposition 3.3.

Corollary 3.4 *Let* $\mathbf{R}$ *satisfy Assumption 3.1. If* $c_n > \frac{1}{2}\left[\overline{R}_n - \left(R_f + \frac{\gamma}{\alpha}\right)\right]$*,* $n = 1, \ldots, N$*, then there exist bounds* $D_n < c_n + \frac{\gamma}{\alpha}$ *so that Proposition 3.3 applies. Moreover,* $c_n \geq \frac{1}{2}\left(\underline{R}_n + \overline{R}_n\right)$ *implies* $c_n > \frac{1}{2}\left[\overline{R}_n - \left(R_f + \frac{\gamma}{\alpha}\right)\right]$*, for* $n = 1, \ldots, N$*.*

The second part of the corollary confirms that, by selecting the midpoint of the distribution of each asset returns, we can guarantee convergence of the approximate optimal portfolio weight vector $\widetilde{\boldsymbol{\omega}}_M$ to the exact optimal portfolio weight vector $\boldsymbol{\omega}^*$.

The following proposition generalizes Proposition 3.3 to an arbitrary (bounded, closed, and convex) set $\boldsymbol{\Omega}$ of admissible portfolio allocations.

Proposition 3.5 *If the excess return deviations* $d_n = R_n - c_n$*,* $n = 1, \ldots, N$*, satisfy the condition* $|d_n| \leq D$ *where* $D < D^*$ *and*

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

then the approximate expected utility $\widetilde{V}_M(\boldsymbol{\omega})$ *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* $\widetilde{\boldsymbol{\omega}}_M$ *converges to the exact optimal weight* $\boldsymbol{\omega}^*$*, as* $M \to \infty$*.*

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 $D^* > 0$.

Remark 3.6 *Note that if there are no short selling constraints, and if the centers of expansion are identical for all assets, i.e.,* $c_n = c$ *for all* $n = 1, \ldots, N$*, then* D^* *in (31) simplifies to* $D^* = R_f + \frac{\gamma}{\alpha} + c$*. This bound coincides with the bound in Proposition 3.3 under the restriction that the centers of expansion are identical across assets.*

The following corollary provides sufficient conditions under which Proposition 3.5 is applicable. The conditions are not optimal but they are relatively simple and easy to check in practical applications.

Corollary 3.7 *Define* $\underline{\mathbf{R}} = (\underline{R}_1, \ldots, \underline{R}_N)'$, $\overline{\mathbf{R}} = \left(\overline{R}_1, \ldots, \overline{R}_N\right)'$, $\underline{R} = \min\left\{\underline{R}_n : n = 1, \ldots, N\right\}$, *and* $\overline{R} = \max\left\{\overline{R}_n : n = 1, \ldots, N\right\}$*. If the condition*

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

is satisfied, then, under the choice $c_n = \frac{\underline{R}_n + \overline{R}_n}{2}$, $n = 1, \ldots, N$*, the condition* $|d_n| \leq D$, $n = 1, \ldots, N$ *in Proposition 3.5 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 $\boldsymbol{\Omega}$ 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.[10] 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 $\mathbf{y} = (y_1, \ldots, y_N)'$ through the following nonlinear transformation

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

[10] 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.

The excess return R_n is then given by $R_n = R_f \left(A_n + B_n e^{C_n y_n}\right)$. Note that Assumption 3.1 implies that y_n also has compact support, which we denote by $[\underline{y}_n, \overline{y}_n]$, for $n = 1, \ldots, N$. 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 μ_n, σ_n, ξ_n represent the mean, standard deviation, and skewness of the excess return R_n, $n = 1, \ldots, N$, the appropriate values of A_n, B_n, and C_n are then determined by

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

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

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

Let κ_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 $\kappa_n \in \left[\underline{y}_n, \overline{y}_n\right]$, $n = 1, \ldots, N$. Since the goal of the transformation is to symmetrize the distribution, in empirical applications it makes sense to set κ_n equal to the midpoint of $[\underline{y}_n, \overline{y}_n]$ in order to minimize the magnitude of the deviation δ_n. Note that the deviation δ_n does not necessarily have mean zero. The portfolio return $R_p(\boldsymbol{\omega})$ is then expressed as

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

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

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

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

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

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

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

The approximate optimal portfolio allocation, denoted by $\widehat{\boldsymbol{\omega}}_M$, is obtained as the solution to the following maximization problem

$$\widehat{\boldsymbol{\omega}}_M = \arg\max_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} \widehat{V}_M(\boldsymbol{\omega}). \tag{41}$$

The following proposition, which is the analogue to Proposition 3.5, 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 $\mathbf{A} = (A_1, \ldots, A_N)'$, $L_n = B_n e^{C_n \kappa_n}$, $n = 1, \ldots, N$, and $\mathbf{L} = (L_1, \ldots, L_N)'$.

Proposition 3.8 *For all $\boldsymbol{\omega} \in \boldsymbol{\Omega}$, let*

$$\lambda_0(\boldsymbol{\omega}) = 1 + \frac{\gamma}{\alpha R_f} + \boldsymbol{\omega}'\mathbf{A}, \ \lambda_n(\boldsymbol{\omega}) = \omega_n L_n, \ n = 1, \ldots, N, \tag{42}$$

$$\lambda^-(\boldsymbol{\omega}) = \sum_{n=1}^{N} \min(\lambda_n(\boldsymbol{\omega}), 0), \ \lambda^+(\boldsymbol{\omega}) = \sum_{n=1}^{N} \max(\lambda_n(\boldsymbol{\omega}), 0), \tag{43}$$

and

$$\boldsymbol{\Lambda}(\boldsymbol{\omega}) = \left(\lambda_0(\boldsymbol{\omega}), \lambda^-(\boldsymbol{\omega}), \lambda^+(\boldsymbol{\omega})\right). \tag{44}$$

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

$$\Delta^* = \min_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} \mathcal{I}^{-1}(0; \boldsymbol{\Lambda}(\boldsymbol{\omega})), \tag{45}$$

where $\mathcal{I}(\cdot; \boldsymbol{\Lambda})$ is defined in Lemma A.9, then the approximate expected utility $\widehat{V}_M(\boldsymbol{\omega})$ converges to the exact utility $V(\boldsymbol{\omega})$, uniformly with respect to the allocation $\boldsymbol{\omega}$. Moreover, the sequence of the approximate optimal portfolio weights $\widehat{\boldsymbol{\omega}}_M$ converges to the exact optimal portfolio weight $\boldsymbol{\omega}^*$, as $M \to \infty$.

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 ξ_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 3.9 *Assume that no short selling is allowed, i.e.,*

$$\boldsymbol{\Omega} = \left\{\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N) : 0 \leq \omega_n \leq 1, \ n = 1, \ldots, N, \ 0 \leq 1 - \boldsymbol{\omega}'\mathbf{1} \leq 1\right\}. \tag{46}$$

Consider the transformed variable $\mathbf{y} = (y_1, \ldots, y_N)$ introduced in (33) with $B_n > 0$, for $n = 1, \ldots, N$. Then, the bound Δ^* in Proposition 3.8 becomes

$$\Delta^* = \min_{\boldsymbol{\omega} \in \boldsymbol{\Omega}} \mathcal{E}^{-1}\left(-\frac{1 + \frac{\gamma}{\alpha R_f} + \boldsymbol{\omega}'\mathbf{A}}{\boldsymbol{\omega}'\mathbf{L}}\right). \tag{47}$$

where $\mathcal{E}(\cdot)$ is the function

$$\mathcal{E}(x) = \min\{e^{ax}\cos(bx) : \ a^2 + b^2 \leq 1\}, \quad x \in \mathbb{R}, \tag{48}$$

defined in Lemma A.6.

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 $\mathbf{r} = (r_1, \ldots, r_N)$, where $\frac{R_{g,n}}{R_f} = e^{r_n}$, $n = 1, \ldots, N$. In this case, $\alpha \to \infty$ and $A_n = -1$, $n = 1, \ldots, N$ and hence since $\mathcal{E}$ is strictly decreasing we have

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

where the last equality follows from equation (A7). This establishes the following

Corollary 3.10 *Assume that no short selling is allowed. Then, for CRRA utility, i.e., when $\alpha \to \infty$, and $\mathbf{y}$ equal to the log excess return vector $\mathbf{r}$, the bound Δ^* in Proposition 3.8 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 $[\underline{r}_n, \overline{r}_n]$. Then, the condition in Corollary 3.10 is satisfied if (i) $\overline{r}_n - \underline{r}_n < \pi$ and (ii) we select the midpoint of the support as the center of expansion: $\kappa_r = \frac{\underline{r}_n + \overline{r}_n}{2}$. Note that the condition $\overline{r}_n - \underline{r}_n < \pi$ is satisfied for a large set of values of $\underline{r}_n$ and $\overline{r}_n$ that imply a rather wide support for the primitive asset return distribution. For instance, for $\underline{r}_n = -1.65$ and $\overline{r}_n = 1.45$ we have $\overline{r}_n - \underline{r}_n = 3.1 < \pi$. 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 $E[U(R_p)]$, where $U(\cdot)$ belongs to the CRRA class with risk aversion γ. We assume that the investment opportunity set consists of a risk-free asset and $N = 3$ risky assets. The annualized gross risk-free rate is set equal to 1.05. The joint distribution of the vector of log excess returns is assumed to be multivariate normal. The parameters of the joint distribution of the risky asset is determined as follows. The means and standard deviations of the risky asset gross returns are drawn at random from the intervals

$[\underline{\mu}, \overline{\mu}]$ and $[\underline{\sigma}, \overline{\sigma}]$, respectively. The correlation matrix of the log excess returns 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. We consider two cases calibrated to quarterly and annual equity returns. The intervals $[\underline{\mu}, \overline{\mu}]$ and $[\underline{\sigma}, \overline{\sigma}]$ are selected to roughly match the corresponding frequency. In the case of quarterly returns, we use $[\underline{\mu}, \overline{\mu}] = [1.03, 1.08]$ and $[\underline{\sigma}, \overline{\sigma}] = [0.10, 0.20]$, while in the case of annual returns, we use $[\underline{\mu}, \overline{\mu}] = [1.10, 1.30]$ and $[\underline{\sigma}, \overline{\sigma}] = [0.20, 0.40]$. We consider two different types of Taylor approximation. The first, labeled "Gross return decomposition", is based on the decomposition of the gross risky asset return $R_n = c_n + d_n$, $n = 1, 2, 3$, while the second, labeled "Log return decomposition", is based on the decomposition of the log excess risky asset return $r_n = \kappa_n + \delta_n$, $n = 1, 2, 3$. We set c_n equal to the midpoint of the support of R_n truncated at its 0.5 and 99.5 percentiles, and κ_n equal to the mean of r_n.

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 6 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 γ 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 superior performance of the Taylor series expansion based on the log return decomposition. The certainty equivalent loss for the log return decomposition are uniformly lower than those based on gross return decomposition. It is remarkable that, for annual equity return parameters and $\gamma = 10$, and when the order of the Taylor expansion is $M = 2$ and $M = 4$, the log return decomposition results in a certainty equivalent loss of 4.48 and 0.16 basis points while the gross return decomposition results in a loss of 455.51 and 53.23 basis points, respectively.

Because log returns are normal in this example, these findings are not surprising in light of the discussion in subsection 2.3. All else being equal, the approximation is deteriorating as the curvature of the utility function increases and for lower frequency data. Note finally that the certainty equivalent losses are monotonically decreasing in the order of expansion M for both the gross and log return decomposition. However, this need not be the case in general. In unreported results we find that, using the mean of the distribution of R_n as a center of expansion c_n may lead to *divergence* of the approximation. 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 3.3 and Corollary 3.4.

4 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, $n = 0, 1, \ldots$ 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., $\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \le 1\}$.

Theorem A.1 [*Theorem 9.20 in Apostol (1974)*] *Given a power series* $\sum_{n=0}^{\infty} a_n (z - z_0)^n$, *define* $\lambda = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$, $r = \frac{1}{\lambda}$, *(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* $c_k = \sum_{n=0}^{\infty} a_n b_k(n)$, $k = 0, 1, \ldots$ *and the numbers* $b_k(n)$ *are defined by the equation* $g(z)^n = \left(\sum_{k=0}^{\infty} b_k z^k\right)^n = \sum_{k=0}^{\infty} b_k(n) z^k$.

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* $f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)(a)}}{n!}(z - z_0)^n$ *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 *Define the functions*

$$\mathcal{R}_1(w) = \frac{\sqrt{1+w^2}}{w} \arctan(w), \quad w \in (0, \infty), \tag{A1}$$

$$\mathcal{R}_2(w) = \frac{\sqrt{1+w^2}}{w} \left(\pi - \arctan(w)\right), \quad w \in (0, \infty). \tag{A2}$$

Then, the following properties hold: (i) $\mathcal{R}_1$ is strictly increasing, $\lim_{w\to 0}\mathcal{R}_1(w) = 1$, *and* $\lim_{w\to\infty}\mathcal{R}_1(w) = \frac{\pi}{2}$ *and (ii) $\mathcal{R}_2$ is strictly decreasing,* $\lim_{w\to 0}\mathcal{R}_2(w) = \infty$, *and* $\lim_{w\to\infty}\mathcal{R}_2(w) = \frac{\pi}{2}$. *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 $\left(1, \frac{\pi}{2}\right)$ and $\left(\frac{\pi}{2}, \infty\right)$, respectively. Moreover, $x\mathcal{S}_2(x) \to \pi$, as $x \to \infty$.*

Proof. Using L'Hospital's rule, we obtain $\lim_{w\to 0}\mathcal{R}_1(w) = \lim_{w\to 0}\frac{\arctan(w)}{w} = \lim_{w\to 0}\frac{1}{1+w^2} = 1$. Moreover, $\lim_{w\to\infty}\mathcal{R}_1(w) = \lim_{w\to\infty}\arctan(w) = \frac{\pi}{2}$. Similarly, the corresponding limits for $\mathcal{R}_2$ are $\lim_{w\to 0}\mathcal{R}_2(w) = \infty$ and $\lim_{w\to\infty}\mathcal{R}_2(w) = \frac{\pi}{2}$. 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 $x_n \to \infty$ as $n \to \infty$, and define $w_n = \mathcal{S}_2(x_n)$. Since $\mathcal{R}_2(w)$ is strictly decreasing and converges to ∞ as $w \to 0$, we have $w_n \to 0$. Then $x_n\mathcal{S}_2(x_n) = \mathcal{R}_2(w_n)w_n = \frac{\sqrt{1+w_n^2}}{w_n}\left(\pi - \arctan(w_n)\right) w_n \to \pi$ since $\lim_{w\to 0}\arctan(w) = 0$. This completes the proof. ■

Lemma A.6 *Define the function*

$$\mathcal{T}(a,b;x) = e^{ax}\cos(bx), \;\; a^2+b^2 \le 1, \;\; x \in \mathbb{R}, \tag{A3}$$

and let

$$\begin{aligned} \mathcal{M}(x) &= \max\{\mathcal{T}(a,b;x) : \;\; a^2+b^2 \le 1\}, \quad x \in \mathbb{R}, && \text{(A4)} \\ \mathcal{E}(x) &= \min\{\mathcal{T}(a,b;x) : \;\; a^2+b^2 \le 1\}, \quad x \in \mathbb{R}. && \text{(A5)} \end{aligned}$$

Then,

$$\mathcal{M}(x) = e^{|x|}, \quad x \in \mathbb{R}, \tag{A6}$$

and

$$\mathcal{E}(x) = \begin{cases} \exp\left(-|x|\right), \quad \text{if } |x| \le 1, \\[2ex] \exp\left(-\frac{1}{\sqrt{1+[\mathcal{S}_1(|x|)]^2}}|x|\right)\cos\left(\frac{\mathcal{S}_1(|x|)}{\sqrt{1+[\mathcal{S}_1(|x|)]^2}}|x|\right), \quad \text{if } 1 < |x| < \frac{\pi}{2}, \\[2ex] 0, \quad \text{if } |x| = \frac{\pi}{2}, \\[2ex] \exp\left(\frac{1}{\sqrt{1+[\mathcal{S}_2(|x|)]^2}}|x|\right)\cos\left(\frac{\mathcal{S}_2(|x|)}{\sqrt{1+[\mathcal{S}_2(|x|)]^2}}|x|\right), \quad \text{if } |x| > \frac{\pi}{2}, \end{cases} \tag{A7}$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ are the functions defined in Lemma A.5. Furthermore, the function $\mathcal{M}(x)$ is strictly increasing in $|x|$ and $\mathcal{M}(x) \to \infty$, as $|x| \to \infty$, while the function $\mathcal{E}(x)$ is strictly decreasing in $|x|$, and $\mathcal{E}(x) \to -\infty$, as $|x| \to \infty$.

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,

$$\mathcal{E}(x) = \min\{\tau(a;x) : -1 \leq a \leq 1\}, \tag{A8}$$

$$\mathcal{M}(x) = \max\{\tau(a;x) : -1 \leq a \leq 1\}, \tag{A9}$$

where

$$\tau(a;x) = e^{ax}\cos\left(x\sqrt{1-a^2}\right). \tag{A10}$$

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

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 τ with respect to a is given by

$$\tau'(a;x) = xe^{ax}\left[\cos\left(\sqrt{1-a^2}x\right) + \frac{a}{\sqrt{1-a^2}}\sin\left(\sqrt{1-a^2}x\right)\right], \; -1 < a < 1. \tag{A11}$$

First, consider $0 < x < \frac{\pi}{2}$ (cases (i) and (ii)). Then, we have $\cos\left(\sqrt{1-a^2}x\right) \geq 0$ for all $a \in [-1, 1]$. Recalling that $\tau(a;x) = e^{ax}\cos\left(\sqrt{1-a^2}x\right)$, 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 $x \in \left(0, \frac{\pi}{2}\right)$ and so $x\frac{w}{\sqrt{1+w^2}} \in \left(0, \frac{\pi}{2}\right)$ which implies $k = 0$ since $\arctan(w) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. 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 in Lemma A.5. 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 of $\mathcal{R}_1$ defined in Lemma A.5. Therefore, $\underline{a}(x) = -\frac{1}{\sqrt{1+[\mathcal{S}_1(x)]^2}}$ for $x \in \left(1, \frac{\pi}{2}\right)$.

Next consider $x > \frac{\pi}{2}$ (case (iii)). Note that $\cos\left(\sqrt{1-a^2}x\right)$ takes negative values as a ranges from -1 to 1. Recalling that $\tau(a;x) = e^{ax}\cos\left(\sqrt{1-a^2}x\right)$ and $\tau(-a;x) = e^{-ax}\cos\left(\sqrt{1-a^2}x\right)$, 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 $x = \frac{\sqrt{1+w^2}}{w}\left(-\arctan(w) + k\pi\right)$ where k is an integer.

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

Next, let $x \geq \frac{3\pi}{2}$. In this case, k should be a positive integer. Define the function $\mathcal{J}(a;x) = \cos\left(\sqrt{1-a^2}x\right)$ 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 $\mathcal{J}'(a;x) = \frac{ax}{\sqrt{1-a^2}}\sin\left(\sqrt{1-a^2}x\right)$. For $a \in (a_1, 1)$, we have $\sqrt{1-a^2}x < \pi$ which implies $\mathcal{J}'(a;x) > 0$, i.e., $\mathcal{J}(\cdot;x)$ is strictly increasing in $(a_1, 1)$. Moreover, $\mathcal{J}(a;x) \to 1$ as $a \to 1$. 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 $\sqrt{1-a^2}x < \pi$ 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 $\arctan(w) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. 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 $x \in \left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$ and $x \in \left(\frac{3\pi}{2}, \infty\right)$, 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 in Lemma A.5. 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 in Lemma A.5. 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 $(0,1)$, $\left(1, \frac{\pi}{2}\right)$, and $\left(\frac{\pi}{2}, \infty\right)$. Clearly $\mathcal{E}(x)$ is strictly decreasing over $(0,1)$. Next, consider the interval $\left(1,\frac{\pi}{2}\right)$. Define the function

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

Since $\cos\left(\arctan(w)\right)=\frac{1}{\sqrt{1+w^2}}$, we have $\mathcal{Q}_1(w)=-\frac{\arctan(w)}{w}-\frac{1}{2}\log\left(1+w^2\right)$. 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 $\mathcal{E}\left(\mathcal{R}_1(w)\right)>\mathcal{E}\left(\mathcal{R}_1(w')\right)\Leftrightarrow\mathcal{E}\left(x\right)>\mathcal{E}\left(x'\right)$. Hence, $\mathcal{E}\left(\cdot\right)$ is strictly decreasing over $\left(1,\frac{\pi}{2}\right)$. To deal with the interval $\left(\frac{\pi}{2},\infty\right)$, define the function

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

Since $\cos\left(\pi-\arctan(w)\right)=-\frac{1}{\sqrt{1+w^2}}$, we have $\mathcal{Q}_2(w)=\frac{\pi-\arctan(w)}{w}-\frac{1}{2}\log\left(1+w^2\right)$. 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 $-\mathcal{E}\left(\mathcal{R}_2(w')\right)>-\mathcal{E}\left(\mathcal{R}_2(w)\right)\Leftrightarrow\mathcal{E}\left(x'\right)<\mathcal{E}\left(x\right)$. Hence, $\mathcal{E}\left(\cdot\right)$ is strictly decreasing over $\left(\frac{\pi}{2},\infty\right)$. Finally, note that as $|x|\to\infty$, Lemma A.5 shows that $|x|\mathcal{S}_2(|x|)\to\pi$ and hence $\mathcal{S}_2(|x|)\to 0$. This limits imply that $\cos\left(\frac{\mathcal{S}_2(|x|)}{\sqrt{1+[\mathcal{S}_2(|x|)]^2}}|x|\right)\to -1$ and therefore $\mathcal{E}(x)\to-\infty$ as $|x|\to\infty$. 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 $\mathbf{x}=(x_1,\ldots,x_N)\in\mathbb{R}^N$ and $\mathbf{m}=(m_1,\ldots,m_N)\in\mathbb{N}^N$, where $\mathbb{N}$ is the set of nonnegative integers, define $\mathbf{x}^{\mathbf{m}}=x_1^{m_1}\cdots x_N^{m_N}$, $\mathbf{m}!=m_1!\cdots m_N!$, $|\mathbf{m}|=m_1+\cdots+m_N$. Furthermore, for a function G defined on $\mathbb{R}^N$ denote its partial derivatives by $G^{(\mathbf{m})}(\mathbf{x})=\frac{\partial^{m_1+\cdots+m_N}}{\partial x_1^{m_1}\cdots\partial x_N^{m_N}}G(x)$.

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

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

and

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

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

$$G(\mathbf{x})=\begin{cases}\frac{1}{1-\gamma}(\lambda_0+\lambda_1x_1+\cdots+\lambda_Nx_N)^{1-\gamma} & \textit{if}\ \ \gamma\neq 1,\\ \\ \log(\lambda_0+\lambda_1x_1+\cdots+\lambda_Nx_N) & \textit{if}\ \ \gamma=1,\end{cases}\tag{A12}$$

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

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

converges uniformly to $G(\mathbf{x})$*, as* $M \to \infty$*, 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

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

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

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

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

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

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 $\gamma \ne 1$, 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

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

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 λ_0, λ_1, and γ be real constants with $\gamma > 0$, $\lambda_1 \neq 0$, $\lambda_0 + \lambda_1 > 0$ and define the function $g(\cdot)$ by*

$$g(x) = \begin{cases} \frac{1}{1-\gamma}(\lambda_0 + \lambda_1 e^x)^{1-\gamma} & \text{if } \; \gamma \neq 1, \\ \\ \log(\lambda_0 + \lambda_1 e^x) & \text{if } \; \gamma = 1. \end{cases} \tag{A13}$$

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 $h(z) = \frac{\lambda_1 e^z}{\lambda_0 + \lambda_1 e^z}$, $z \in \lambda$. 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. ■

Next, we generalize the above result to the N-dimensional case.

Lemma A.9 *Let $\gamma > 0$, $\lambda_0 \in \mathbb{R}$, $\lambda^- \leq 0$, and $\lambda^+ \geq 0$ where at least one of λ^- or λ^+ is different from 0, and assume that $\lambda = \lambda_0 + \lambda^- + \lambda^+ > 0$. Denote $\mathbf{\Lambda} = (\lambda_0, \lambda^-, \lambda^+)$ and define the function*

$$\mathcal{I}(\Delta; \mathbf{\Lambda}) = \lambda_0 + \lambda^+\mathcal{E}(\Delta) + \lambda^-\mathcal{M}(\Delta), \ \ \Delta \in [0, \infty), \tag{A14}$$

where the functions $\mathcal{E}$ and $\mathcal{M}$ are defined in Lemma A.6. The function $\mathcal{I}(\cdot; \mathbf{\Lambda})$ satisfies the following properties: (i) $\mathcal{I}(0; \mathbf{\Lambda}) = \lambda > 0$, (ii) $\mathcal{I}(\cdot; \mathbf{\Lambda})$ is strictly decreasing, and (iii) $\mathcal{I}(\Delta; \mathbf{\Lambda}) \to -\infty$ as $\Delta \to \infty$. 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 $\lim_{\Delta\to\infty} \mathcal{E}(\Delta) = -\infty$, and $\mathcal{M}(\cdot)$ is strictly increasing with $\lim_{\Delta\to\infty} \mathcal{M}(\Delta) = \infty$. 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 $\lim_{\Delta\to\infty} \mathcal{I}(\Delta; \mathbf{\Lambda}) = -\infty$. ■

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

$$\lambda^- = \sum_{n=1}^{N} \min(\lambda_n, 0), \ \ \lambda^+ = \sum_{n=1}^{N} \max(\lambda_n, 0), \tag{A15}$$

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

$$G(\mathbf{x}) = \begin{cases} \frac{1}{1-\gamma}(\lambda_0 + \lambda_1 e^{x_1} + \cdots + \lambda_N e^{x_N})^{1-\gamma} & \text{if } \ \gamma \neq 1, \\ \\ \log(\lambda_0 + \lambda_1 e^{x_1} + \cdots + \lambda_N e^{x_N}) & \text{if } \ \gamma = 1, \end{cases} \tag{A16}$$

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

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

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

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

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

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

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 $\mathbf{x} \in \mathbb{R}^N$ with $|x_n| \leq \Delta$, for $n = 1, \ldots, N$. 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 $F_1(t; \mathbf{x}) = \log(\lambda_0 + \lambda_1 e^{tx_1} + \cdots + \lambda_N e^{tx_N})$. 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

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

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

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

with domain $\mathbb{D}_H = \{z \in \mathbb{C} : \ \lambda_0 + \lambda_1 e^{zx_1} + \cdots + \lambda_N e^{zx_N} \neq 0\}$. 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

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

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 $\lambda_0 + \lambda_1 e^{zx_1} + \cdots + \lambda_N e^{zx_N} = 0$ 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 $f : \mathbf{X} \to \mathbb{R}$ 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 $f(\mathbf{x}_M) \to f(\mathbf{x}^*)$, as $M \to \infty$, then $\| \mathbf{x}_M - \mathbf{x}^* \| \to 0$, as $M \to \infty$, 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 $\mathbf{x}_{L(M)} \to \tilde{\mathbf{x}} \in \mathbf{Q}_\varepsilon$, as $M \to \infty$. By the continuity of f, it follows that $f(\mathbf{x}_{L(M)}) \to f(\tilde{\mathbf{x}})$, as $M \to \infty$. But $f(\mathbf{x}_{L(M)})$ is a subsequence of $f(\mathbf{x}_M)$ and so $f(\mathbf{x}_{L(M)}) \to f(\mathbf{x}^*)$, as $M \to \infty$. 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 $M \to \infty$, is established. ■

Lemma A.12 *Let* $\mathbf{X}$ *be a bounded, closed, and convex subset of* $\mathbb{R}^N$ *and* $f : \mathbf{X} \to \mathbb{R}$ *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* $M = 1, 2, \ldots$*, let* $f_M : \mathbf{X} \to \mathbb{R}$ *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* $\mathbf{x}_M^* \to \mathbf{x}^*$*, as* $M \to \infty$*.*

Proof. First we show that, as $M \to \infty$, $f(\mathbf{x}_M^*) \to f(\mathbf{x}^*)$. 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 $f_M(\mathbf{x}_M^*) - f(\mathbf{x}_M^*) \to 0$. Therefore, it suffices to show that $f_M(\mathbf{x}_M^*) - f(\mathbf{x}^*) \to 0$, as $M \to \infty$. Recall that, by assumption, $f_M(\mathbf{x}) \to f(\mathbf{x})$ uniformly in $\mathbf{x} \in \mathbf{X}$. Let $\varepsilon > 0$. Then, there exists M_ε 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

$$f(\mathbf{x}^*) - f_M(\mathbf{x}_M^*) \leq \varepsilon. \tag{A17}$$

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

$$-\varepsilon \leq f(\mathbf{x}^*) - f_M(\mathbf{x}_M^*). \tag{A18}$$

Therefore, combining (A17) and (A18), we obtain $-\varepsilon \leq f(\mathbf{x}^*) - f_M(\mathbf{x}_M^*) \leq \varepsilon$, for all $M \geq M_\varepsilon$ which establishes that $f(\mathbf{x}_M^*) \to f(\mathbf{x}^*)$, as $M \to \infty$. 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 $M \to \infty$, 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 $M \to \infty$,

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

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 ζ, 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 $c_p \geq \frac{1}{2}\left(\underline{R}_p + \overline{R}_p\right)$ implies the condition $c_p > \frac{1}{2}\left(\overline{R}_p - \frac{\gamma}{\alpha}\right)$. ■

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

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

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

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

It follows that

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

since $\ell > 1$. Hence, the sequence τ_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 Δ_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 $\boldsymbol{\omega} \in \boldsymbol{\Omega}$ and consider a sequence $\boldsymbol{\omega}_M \in \boldsymbol{\Omega}$ with $\boldsymbol{\omega}_M \to \boldsymbol{\omega}$, as $M \to \infty$. Then $V(\boldsymbol{\omega}_M) = \mathbb{E}\left[U(R_f + \boldsymbol{\omega}_M'\mathbf{R})\right]$. Since, by assumption 3.1, the support of $\mathbf{R}$, denoted by $\mathbf{S_R}$, is compact and $\boldsymbol{\omega}_M \to \boldsymbol{\omega}$, it follows that $R_f + \boldsymbol{\omega}_M'\mathbf{R}$ lies on a compact set, for all $\mathbf{R} \in \mathbf{S_R}$ and all $M = 1, 2, \ldots$ The continuity of U then implies that there exists a bound $K > 0$ such that $|U(R_f + \boldsymbol{\omega}_M'\mathbf{R})| \leq K$ for all $\mathbf{R} \in \mathbf{S_R}$ and all $M = 1, 2, \ldots$ Hence, we can apply the dominated convergence theorem to pass the limit inside the expectation and obtain $\lim_M V(\boldsymbol{\omega}_M) = \lim_M \mathbb{E}\left[U(R_f + \boldsymbol{\omega}_M'\mathbf{R})\right] = \mathbb{E}\left[\lim_M U(R_f + \boldsymbol{\omega}_M'\mathbf{R})\right] = \mathbb{E}\left[U(R_f + \boldsymbol{\omega}'\mathbf{R})\right] = V(\boldsymbol{\omega})$. 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

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

for all $\mathbf{R} \in \mathbf{S_R}$ 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. First, note that there exists $0 < \theta < 1$ such that $D_n \leq \theta(R_f + c_n + \frac{\gamma}{\alpha})$ for all $n = 1, \ldots, N$. It follows that

$$|d_p(\boldsymbol{\omega})| = |\boldsymbol{\omega}'\mathbf{d}| \leq \sum_{n=1}^{N}\omega_n|d_n| \leq \sum_{n=1}^{N}\omega_n D_n \leq \theta\sum_{n=1}^{N}\omega_n\left(R_f + c_n + \frac{\gamma}{\alpha}\right) = \theta\left(c_p(\boldsymbol{\omega}) + \frac{\gamma}{\alpha}\right).$$

Hence, $|d_p(\boldsymbol{\omega})| \leq D$ where $D = \theta\left(c_p(\boldsymbol{\omega}) + \frac{\gamma}{\alpha}\right) < c_p(\boldsymbol{\omega}) + \frac{\gamma}{\alpha}$ uniformly in $\boldsymbol{\omega}$ and $\mathbf{d}$. The uniform convergence of $\widetilde{V}_M(\cdot)$ to $V(\cdot)$ now follows from Proposition 2.2. The convergence of $\widetilde{\boldsymbol{\omega}}_M$ to $\boldsymbol{\omega}^*$ then follows from Lemma A.12. ■

Proof of Corollary 3.4. The proof follows the steps in the proof of Corollary 2.3. ■

Proof of Proposition 3.5. According to equation (24), we have

$$\widetilde{\mathcal{U}}(\mathbf{d};\boldsymbol{\omega}) = \frac{1}{1-\gamma}\left(\left(R_f + \frac{\gamma}{\alpha} + \boldsymbol{\omega}'\mathbf{c}\right) + \sum_{n=1}^{N}\omega_n d_n\right)^{1-\gamma}. \tag{B19}$$

Let $\lambda_0(\boldsymbol{\omega}) = R_f + \frac{\gamma}{\alpha} + \boldsymbol{\omega}'\mathbf{c}$ and $\lambda_n(\boldsymbol{\omega}) = \omega_n$, $n = 1, \ldots, N$. Since, by assumption, $c_n \in [\underline{R}_n, \overline{R}_n]$, $n = 1, \ldots, N$, 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 $G = \widetilde{\mathcal{U}}(\cdot;\boldsymbol{\omega})$ and $\mathbf{x} = \mathbf{d}$, we obtain that the approximate utility $\widetilde{\mathcal{U}}_M(\mathbf{d};\boldsymbol{\omega})$ converges to the exact utility $\widetilde{\mathcal{U}}(\mathbf{d};\boldsymbol{\omega})$ uniformly in $\mathbf{d} \in [-D, D]^N$ with $0 < D < D^*(\boldsymbol{\omega})$ where $D^*(\boldsymbol{\omega}) = \frac{\lambda_0(\boldsymbol{\omega})}{\sum_{n=1}^{N}|\omega_n|} = \frac{R_f + \frac{\gamma}{\alpha} + \boldsymbol{\omega}'\mathbf{c}}{|\boldsymbol{\omega}|'\mathbf{1}}$. It follows that the $\widetilde{\mathcal{U}}_M(\mathbf{d};\boldsymbol{\omega})$ converges $\widetilde{\mathcal{U}}(\mathbf{d};\boldsymbol{\omega})$ uniformly in $\mathbf{d} \in [-D, D]^N$ and $\boldsymbol{\omega} \in \boldsymbol{\Omega}$, with $0 < D < D^*$ where D^* is defined by equation (31). This, in turn, implies that the approximate expected utility $\widetilde{V}_M(\boldsymbol{\omega})$ converges to the exact expected utility $\widetilde{V}(\boldsymbol{\omega})$, uniformly in $\boldsymbol{\omega} \in \boldsymbol{\Omega}$, as $M \to \infty$. The convergence of $\widetilde{\boldsymbol{\omega}}_M$ to $\boldsymbol{\omega}^*$ then follows from Lemma A.12. ■

Proof of Corollary 3.7. Note that

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

Therefore, it follows from condition (32) that Proposition 3.5 is applicable. ■

Proof of Proposition 3.8. According to (38), we obtain

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

where $\mathbf{A} = (A_1, \ldots, A_N)'$ and $L_n = B_n e^{C_n\kappa_n}$, $n = 1, \ldots, N$. Let $\lambda_0(\boldsymbol{\omega}) = 1 + \frac{\gamma}{\alpha R_f} + \boldsymbol{\omega}'\mathbf{A}$, $\lambda_n(\boldsymbol{\omega}) = \omega_n L_n$, $n = 1, \ldots, N$, $\lambda^-(\boldsymbol{\omega}) = \sum_{n=1}^{N}\min(\lambda_n(\boldsymbol{\omega}), 0)$, $\lambda^+(\boldsymbol{\omega}) = \sum_{n=1}^{N}\max(\lambda_n(\boldsymbol{\omega}), 0)$, and $\boldsymbol{\Lambda}(\boldsymbol{\omega}) = (\lambda_0(\boldsymbol{\omega}), \lambda^-(\boldsymbol{\omega}), \lambda^+(\boldsymbol{\omega}))$. Note that $\lambda_0(\boldsymbol{\omega}) + \lambda_1(\boldsymbol{\omega}) + \cdots + \lambda_N(\boldsymbol{\omega})$ equals $\frac{\gamma}{\alpha}$ plus the portfolio return realization for $\delta_n = 0$ or equivalently $y_n = \kappa_n$, $n = 1, \ldots, N$. Since $\kappa_n \in \left[\underline{y}_n, \overline{y}_n\right]$, Assumption 3.1 implies that $\sum_{n=0}^{N}\lambda_n(\boldsymbol{\omega}) \geq \epsilon > 0$. Hence, for a given $\boldsymbol{\omega} \in \boldsymbol{\Omega}$, applying Lemma A.10 with $G = \widehat{\mathcal{U}}$ and $x_n = C_n\delta_n$, $n = 1, \ldots, N$, we obtain that the approximate utility $\widehat{\mathcal{U}}_M(\boldsymbol{\delta};\boldsymbol{\omega})$ converges to the exact utility $\widehat{\mathcal{U}}(\boldsymbol{\delta};\boldsymbol{\omega})$, uniformly in $\boldsymbol{\delta} \in [-\Delta, \Delta]^N$, where

$0 < \Delta < \Delta^*(\mathbf{\Lambda}(\boldsymbol{\omega})) = \mathcal{I}^{-1}(0; \mathbf{\Lambda}(\boldsymbol{\omega}))$. Note that $\mathcal{I}^{-1}(0; \mathbf{\Lambda}(\boldsymbol{\omega}))$ is well-defined since, according to Lemma A.9, $\mathcal{I}^{-1}(\cdot; \mathbf{\Lambda}(\boldsymbol{\omega}))$ 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 $\widehat{\mathcal{U}}_M(\boldsymbol{\delta}; \boldsymbol{\omega})$ converges to $\widehat{\mathcal{U}}(\boldsymbol{\delta}; \boldsymbol{\omega})$ uniformly in $\boldsymbol{\delta} \in [-\Delta, \Delta]^N$ and $\boldsymbol{\omega} \in \mathbf{\Omega}$, with $0 < \Delta < \Delta^*$ where Δ^* is defined by equation (45). This, in turn, implies that the approximate expected utility $\widehat{V}_M(\boldsymbol{\omega})$ converges to the exact expected utility $\widehat{V}(\boldsymbol{\omega})$, uniformly in $\boldsymbol{\omega} \in \mathbf{\Omega}$, as $M \to \infty$. The convergence of $\widehat{\boldsymbol{\omega}}_M$ to $\boldsymbol{\omega}^*$ then follows from Lemma A.12. ■

Proof of Corollary 3.9. Under the stated assumptions and using the notation in Proposition 3.8, we have $\lambda_n(\boldsymbol{\omega}) = \omega_n L_n \geq 0$, $n = 1, \ldots, N$. 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 ranks of 11 portfolios based on exact and approximate expected utility calculations. The utility function is CRRA with risk aversion γ. The portfolio gross returns $R_p = \exp(r_p)$ are assumed to be lognormally distributed. The mean μ_R and the standard deviation σ_R of R_p are drawn at random from the intervals $[\underline{\mu}, \overline{\mu}]$ and $[\underline{\sigma}, \overline{\sigma}]$. We consider two cases calibrated to quarterly and annual equity returns. The intervals are selected to roughly match the corresponding frequency. In the case of quarterly returns, we use $[\underline{\mu}, \overline{\mu}] = [1.03, 1.06]$ and $[\underline{\sigma}, \overline{\sigma}] = [0.10, 0.15]$, while in the case of annual returns, we use $[\underline{\mu}, \overline{\mu}] = [1.10, 1.20]$ and $[\underline{\sigma}, \overline{\sigma}] = [0.20, 0.30]$. The reported Spearman correlations are averages across 10,000 simulation repetitions. 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 $r_p = \kappa_r + \delta_r$. We set c_p equal to the midpoint of the support of R_p truncated at its 0.5 and 99.5 percentiles, and κ_r equal to the mean of r_p.

	Order of approximation			
	$M = 2$	$M = 4$	$M = 6$	$M = 8$
	Quarterly returns			
	Gross return decomposition			
$\gamma = 2$	0.9861	0.9992	0.9999	1.0000
$\gamma = 4$	0.9500	0.9955	0.9995	1.0000
$\gamma = 6$	0.8916	0.9878	0.9982	0.9999
$\gamma = 8$	0.8226	0.9767	0.9955	0.9995
$\gamma = 10$	0.7478	0.9622	0.9914	0.9985
	Log return decomposition			
$\gamma = 2$	0.9998	1.0000	1.0000	1.0000
$\gamma = 4$	0.9965	0.9999	1.0000	1.0000
$\gamma = 6$	0.9874	0.9991	0.9999	1.0000
$\gamma = 8$	0.9750	0.9971	0.9996	1.0000
$\gamma = 10$	0.9607	0.9939	0.9990	0.9998
	Annual returns			
	Gross return decomposition			
$\gamma = 2$	0.9502	0.9933	0.9990	0.9996
$\gamma = 4$	0.8173	0.9646	0.9922	0.9988
$\gamma = 6$	0.6327	0.9042	0.9735	0.9986
$\gamma = 8$	0.4461	0.8143	0.9383	0.9895
$\gamma = 10$	0.2848	0.7060	0.8866	0.9718
	Log return decomposition			
$\gamma = 2$	0.9994	1.0000	1.0000	1.0000
$\gamma = 4$	0.9898	0.9990	0.9999	1.0000
$\gamma = 6$	0.9653	0.9938	0.9989	0.9998
$\gamma = 8$	0.9305	0.9820	0.9951	0.9987
$\gamma = 10$	0.8919	0.9646	0.9874	0.9954

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 γ. There is one risk-free asset with annualized gross rate equal to 1.05 and 3 risky assets with log excess returns following a joint normal distribution. The parameters of the log excess return multivariate distribution are drawn at random in a fashion detailed in subsection 3.3. The reported certainty equivalent losses are stated in annualized basis points and are computed as the medians across 1,000 simulation repetitions. We consider two Taylor approximations. The first is based on the decomposition of the gross risky asset 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 $r_n = \kappa_n + \delta_n$, $n = 1, 2, 3$. We set c_n equal to the midpoint of the support of R_n truncated at its 0.5 and 99.5 percentiles, and κ_n equal to the mean of r_n.

	Order of approximation			
	$M = 2$	$M = 4$	$M = 6$	$M = 8$
	Quarterly returns			
	Gross return decomposition			
$\gamma = 2$	1.04	0.00	0.00	0.00
$\gamma = 4$	4.97	0.02	0.00	0.00
$\gamma = 6$	15.50	0.11	0.00	0.00
$\gamma = 8$	40.37	0.40	0.00	0.00
$\gamma = 10$	109.27	0.81	0.01	0.00
	Log return decomposition			
$\gamma = 2$	0.02	0.00	0.00	0.00
$\gamma = 4$	0.07	0.00	0.00	0.00
$\gamma = 6$	0.17	0.00	0.00	0.00
$\gamma = 8$	0.36	0.00	0.00	0.00
$\gamma = 10$	0.47	0.00	0.00	0.00
	Annual returns			
	Gross return decomposition			
$\gamma = 2$	6.14	0.08	0.00	0.00
$\gamma = 4$	37.34	0.96	0.03	0.00
$\gamma = 6$	109.76	4.27	0.23	0.01
$\gamma = 8$	242.30	15.93	1.02	0.06
$\gamma = 10$	455.51	53.23	3.05	0.19
	Log return decomposition			
$\gamma = 2$	0.12	0.00	0.00	0.00
$\gamma = 4$	0.47	0.01	0.00	0.00
$\gamma = 6$	1.24	0.03	0.00	0.00
$\gamma = 8$	2.33	0.08	0.00	0.00
$\gamma = 10$	4.48	0.16	0.00	0.00

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," *The 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.

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, 2008, "A State Variable Decomposition Approach for Solving Portfolio Choice Problems," Working Paper, University of Texas at Austin.

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," *The American Economic Review*, 84, 713–71.

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.

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," *The 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," *The American Economic Review*, 66, 904–910.

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.

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," *The American Economic Review*, 62, 354–371.

Wu, L., 2003, "Jumps and Dynamic Asset Allocation," *Review of Quantitative Finance and Accounting*, 20, 207–243.