The ordered mean difference as a portfolio performance measure

Roger J. Bowden
School of Economics and Finance, Victoria University of Wellington, P.O. Box 600, Wellington 6001, New Zealand
Accepted 17 July 2000

Abstract

The ordered mean difference (OMD) function is a running mean of the difference between returns on a given fund or security and a benchmark such as the market portfolio, ordered by values of the benchmark. The expectation of this function, conditional on the observed series of benchmark returns, can be used as a portfolio performance measure. Such a formulation is the financial counterpart of the equivalent variation of welfare economics, which in the present context can be regarded as the resource rental value of the manager’s market timing ability. The conditional ordered mean difference (COMD) is related to a number of recent approaches to the principles of performance measurement. It can be estimated either by using the OMD as a non-parametric estimate, or else by fitting the conditional expectation by least squares. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Benchmarking; Equivalent margin; Equivalent variation; Portfolio performance; Jensen’s alpha; Market timing; Ordered mean difference; State price deflator; Stochastic dominance

1. Introduction

Given a time series of returns on a designated fund and returns on a benchmark fund or portfolio with which it is to be compared, it is natural to base performance...
comparisons on the regression of the returns for the given fund on those for the benchmark. Indeed, this simple idea has a long history in performance measurement theory and practice, and is invoked particularly in the context of market timing, where one wishes to test whether the managers of the fund in question have superior market timing ability relative to a benchmark such as the market index. Thus, Merton (1981) showed that certain kinds of market timing can be regarded as equivalent to suitably chosen option strategies with puts or calls on the market index, the empirical counterpart of which is a nonlinear theoretical regression of the fund return on the market; for example, when the market index return rises, the return on the subject fund does at least as well, but when the market index falls, returns on the subject fund do not fall by as much. Dybvig and Ross (1985) showed that the consequences of ignoring such a nonlinearity could be disastrous — linear measures such as the familiar Jensen’s alpha could give quite the wrong answers where superior market timing existed. For further examples and discussion along these lines, see, e.g. Treynor and Mazuy (1966), Jensen (1969), Admati and Ross (1985), Admati et al. (1986), Henriksson and Merton (1981), Jagannathan and Korajczyk (1986), Grinblatt and Titman (1989), Glosten and Jagannathan (1994) and Chen and Knez (1996). In practical terms, one might typically fit a quadratic or piecewise linear regression of the subject fund on the market, and look for a significant positive coefficient on the quadratic term, or in the piecewise linear approach, for significantly different slope coefficients according to designated regions of market returns.

Investors differing in their risk profile might very well react quite differently to the same regression of fund returns on market returns. The more risk averse would presumably view favourably evidence of superiority in bad times, while those less risk averse would examine more carefully any evidence of superiority when times were good as well. Thus, a fitted regression by itself does not automatically constitute a formal performance assessment in terms of any investor welfare criterion. However, in the present paper, it is shown that such a role can be assigned, not to the regression itself, but to a construct derivable from the regression, namely the mean area between the regression curve or line and the 45\degree line. The result is called the conditional expected ordered mean difference, given the sequence of observations of the benchmark return.

The ordered mean difference (OMD) is a function whose value at any point $R$ is equal to the running mean difference between the fund return ($r$) and the benchmark return ($R$) up to that point, where the observations are first ordered by the benchmark return. The expected value at any point, given the market return, is the conditional ordered mean difference (COMD). Such functions may be computed and plotted from the time series of observations on fund and benchmark returns. It turns out that if the COMD is positive over the entire range of $R$ values, then the fund would be preferred to the benchmark by any risk averse investor, no matter what the specific utility function should be. The universal nature of this result is reminiscent of second order stochastic dominance (SSD), which involves
a running comparison of areas between the two empirical distribution functions of returns, for the fund and the benchmark. However, the OMD criterion is based on something more than a simple or marginal tabulation of the two returns, namely the regression of one return on the other, so that one is interested in knowing what happened to the fund at a time when the return on the benchmark was bad, or whether the fund outperformed the market when conditions turned out to be good for the latter. One can therefore have the situation where a fund is COMD-dominant over the market, without necessarily being stochastically dominant.

Recently, the problem of performance measurement has received added impetus from theoretical contributions that create a class of measures interpretable as expected differences between the fund and benchmark returns, the latter discounted by a stochastic discount factor derived from the “law of one price” principle of Hansen and Jagannathan (1991), but also implicit in the broader literature on financial general equilibrium. For contributions of this kind, see Glosten and Jagannathan (1994) and Chen and Knez (1996), the latter using the discount principle to construct an axiomatic approach to performance measures.

With regard to the present contribution, the theoretical underpinnings of the present COMD criterion lie in a formulation of the investor utility problem in an earlier paper (Bowden, 1992), in which the superiority of the fund over the benchmark is portrayed in terms of the equivalent margin, an idea borrowed from welfare economics, wherein the fund is judged to add value to the benchmark fund if the investor needs to be compensated to give up (i.e. hold zero of) the fund in addition to the benchmark fund, in his portfolio. The resulting measure turns out to be a variant of the stochastic discount family, indeed a natural version since it is related to the “state price deflator” construct of financial general equilibrium theory, which gives the value of a dollar in different states of the world. As originally formulated, the equivalent margin measure was not particularly operable, since it did require knowledge or specification of the investor’s utility function. But the virtue of the present COMD criterion is that if satisfied, the equivalent margin is positive for any risk averse utility function. The mean area underneath the theoretical regression of the fund on the market as benchmark can be taken as a measure of investor surplus from investing in the fund, an interesting parallel with the concept of consumer surplus in welfare economics. If the mean area is always positive, then so is the equivalent margin, no matter what risk averse utility function is used to evaluate it. The COMD criterion, therefore, can be interpreted as a dominance construct derived from the general family of expected discounted return difference measures.

The scheme of development in the paper is as follows. Section 2 is a brief exposition of the equivalent margin, as a theoretical rationale underlying the marginal utility weighted mean regression difference. The relationship of this to the discounted expected difference measures is outlined. Section 3 is the core of the paper. Here, the otherwise arbitrary utility functions are replaced with a set of utility generators representing put option payoff functions, to provide a sufficiency
condition for performance assessment that is independent of the observer’s utility function, provided only that he or she is risk averse. Section 4 briefly explores the relationship with classical stochastic dominance. Section 5 draws everything together empirically, showing that the sample OMD function has a theoretical justification as a non-parametric estimate of the COMD function. An alternative parametric procedure uses a prior fitting of a hypothesised theoretical regression to directly estimate the COMD, and the two approaches are compared. Some practical wrinkles and illustrative empirical calculations follow in Section 6. Section 7 offers some concluding remarks.

2. The equivalent margin

In a one-period model of portfolio selection, let \( R \) be the return on a benchmark portfolio. In studies on market timing, one would usually take this as the market portfolio, and would wish to examine whether the fund manager can consistently do better than the market over the ups and downs of the latter. However, the same methodology can apply even where \( R \) is simply the return on another fund and the two are to be compared. Suppose that one is to form a portfolio of initial value US$1 that combines \((1 - x)\) units of the benchmark portfolio with \( x \) units of the designated fund of return \( r \). If the agent has Von-Neumann/Morgenstern utility function \( U \), then the optimal portfolio proportions \( x, 1 - x \) solve the decision problem\(^1\) \[ \max_x E_{r,R} [U(xr + (1 - x)R)]. \]

Suppose now that a return tax \( t \) was levied to penalise the return \( r \) so that the net return will be \( r - t \), with \( t \) known in advance. Although much used in welfare economics, such notional taxes might sit a little oddly in a finance theoretic context. One can provisionally think of \( t \) as a resource rent margin generated by the special timing skills of the fund manager. If the natural preference is to go long in fund \( r \), i.e. \( x > 0 \) in the unconstrained decision problem, the imposition of a return tax (or resource rent tax) will penalise the holding of \( r \). By increasing the tax rate, one can ultimately drive the desired holding to zero. The higher the tax rate needed to do this, the more valuable the security or fund \( r \) is judged to be — the higher is the implied value of the manager’s resource rental that results from his special timing ability. The margin or tax that is necessary to drive the holding of the fund to zero is called the equivalent margin. The name is derived from the “equivalent variation” in economic project analysis, which measures the opportunity loss to the community if the project were withdrawn or not available.

\(^1\) On notation, subscripts beneath the expectation operator are sometimes used to denote explicitly which random variables are involved, and to briefly indicate conditioning. Also for brevity, we sometimes refer to the “security whose return is \( r \)” as just “security \( r \)”. 
Alternatively, suppose that the natural stance is to hold the fund short in conjunction with the benchmark portfolio, so that the unconstrained $x < 0$. The tax contribution to welfare is in this case positive; for the net return is $x(r - t)$, the tax component is $-xt$, and $x < 0$. A tax rate needed to drive portfolio holdings of $r$ to zero would have to be negative — the investor would actually have to be compensated to hold the designated fund. A negative value for the equivalent margin would therefore indicate that while $r$ may have value in conjunction with $R$, it is derived from going short, and it would be difficult to argue that $r$ is in any sense superior to $R$.

Formally, the equivalent margin $t_U$ corresponding to an arbitrary utility function $U$ is defined by:

$$t_U = \arg \sup_t \left\{ xy_{E_r} U(x(r - t) + (1 - x) R) = 0 \right\}.$$

In other words, for a fixed $t$, one proceeds to maximise $U$ with respect to $x$; then one adjusts $t$ so that the optimising $x$ is precisely zero. If the utility function $U$ is concave, the optimum for $x$ solves the first order conditions.

$$E[(r - R - t)U'[x(r - t) + (1 - x) R]] = 0.$$

Setting $t$ so that $x = 0$ yields,

$$t_U = \frac{E_{r,R}[(r - R)U'(R)]}{EU'(R)} = E_{r,R}[\pi(R)(r - R)]. \quad (1)$$

The equivalent margin is the expectation of a weighted sum of the return differences, with weights $\pi(R) = U'(R)/EU'(R)$ that are semi-positive and sum to unity. States of the world in which benchmark marginal utility is greatest receive highest weightings, so that this belongs to the marginal utility weighting class of portfolio measures.

2.1. Validational aspects

The equivalent margin measure can be related to a number of developments in the literature on fund performance and financial general equilibrium. Note that the term $\pi(R)$ can be regarded as a stochastic discount factor such that the implied performance indicator has the form $E[\pi(R)r - R]$. In the recent theory of financial general equilibrium (see Duffie, 1992), the weights $\pi(R) = U'(R)/EU'(R)$ can be taken as state price deflators where $R$ is linearly related to alternative states of investor wealth. In a complete market, the state deflators, which represent the value of money in different states of the world, are unique. In such terms, Eq. (1) can formally be written $t_U = E^Q(r - R)$, where $Q$ denotes a
revised probability measure (equivalent martingale measure) obtained by adjusting the natural probabilities by the state price deflators $\pi$. In such a world, investors would be risk neutral, and the natural measure of the welfare generated by $r$ relative to $R$ is simply the expectation of the difference $r - R$. Now, in a no-arbitrage equilibrium, this should be precisely zero. To the extent that $I_t$ as computed from expression (1) is not zero, this indicates either that security $r$ is not already spanned by those making up the market $R$; or that it is, but is not marketable so that potential arbitrages cannot be executed.

Recently, Chen and Knez (1996) have considered the case where markets are not complete. In this case, there is no uniquely suitable discount factor. However, they construct a family of measures (their $\lambda$ family) that have the general form of a discounted expectation as in Eq. (1), utilising earlier work on stochastic discount factors by Hansen and Jagannathan (1991). It follows from Theorem 2 of Chen and Knez that the equivalent margin is what they call an “admissible” performance measure, where one is to compare the performance of the fund with that of the market as a whole. Moreover, provided that titles to funds A and B are traded, the discount factor $\pi(R)$ generates an admissible performance measure $E[\pi(R)(r_A - r_B)]$ for bilateral comparisons between the two. In the Chen and Knez terminology, performance measures are admissible if they assign a performance measure of zero to a manager who possesses no special information, and cannot be “fooled” by simply repackaging or mixing funds or fund payoffs, together with other more technical requirements.

The special information aspect draws on the earlier work of Dybvig and Ross (1985). Suppose that returns $r_i$ on a set of securities are generated in terms of a single zero mean common factor $s$ and zero mean idiosyncratic influences $\varepsilon_i$: $$r_i = \mu_i + \beta_i s + \varepsilon_i; \quad s, \varepsilon_i \text{ are i.i.d. over time.}$$

The “signal” $s$ can be partially detected by just one of the market participants, a fund manager who observes a noise ($\eta$) distorted version as: $$v = s + \eta; \quad \eta \text{ i.i.d. over time.}$$

In such a case, Jensen’s alpha fails to capture the manager’s market timing ability, for as $v$ is observed the manager tends to switch in and out of stocks vs. the safe asset, and a quadratic relationship exists between the return on the generated portfolio and the market — the relevant theoretical regression $e(R)$ is nonlinear. This generates an unconditional $\chi^2$ distribution for the manager’s portfolio return and can easily generate a negative alpha as an artifice of the nonlinear theoretical regression on $R$. However, it can be shown (Bowden, 1992) that if the observer’s utility function is exponential, then the equivalent margin is directly proportional to the manager’s signal-to-noise ratio $\sigma_s^2/\sigma_\eta^2$, so that the source of the welfare gain is correctly captured.
In summary, the equivalent margin is well-motivated in terms of validational aspects derived from the literature on both financial general equilibrium and fund performance, as well as a degree of intuitive appeal from the argument used to establish its formula. The connection with the state price deflator theory further strengthens the case for the particular choice of weight as the ratio of the marginal utility in each state to the theoretical average marginal utility: if the fund of return \( r \) does well in conditions where the market return \( R \) is low and hence the marginal \( U'(R) \) is high, then this is a good fund to hold, at least in conjunction with the market, if not in total replacement.

2.2. The regression formulation

Let \( e(R) = E[r \mid R] \) denote the conditional expectation of \( r \) given the benchmark return \( R \). As a function of \( R \), \( e \) is the theoretical regression of \( r \) on \( R \). Thus,

\[
r = e(R) + e; \quad E[e \mid R = 0].
\]

Combining this with expression (1), yields:

\[
t_{U} = \frac{E_{R}[((e(R) - R)U'(R)]}{E_{R}U'(R)}.
\]

The notation \( E_{R} \) indicates that the expectation is to be taken over the marginal density of \( R \); in what follows, the analysis proceeds in the framework of the joint density of \( r \) and \( R \), and the expectational limiter \( E_{R} \) or \( E \) is frequently used to remind the reader of the precise domain of the expectational operator. It will be dropped where the domain is obvious\(^2\). In the form as defined, the equivalent variation is the marginal utility weighted sum of the theoretical regression differences. The condition \( e(R) \geq R \), all \( R \), is sufficient but not necessary for \( t_{U} \geq 0 \).

Operationally, however, the equivalent margin does depend upon the particular utility function employed to evaluate it, so that conclusions may depend on the observer and his or her preferences. Section 3 shows how this limitation can be removed.

\(^2\) Also on notation, in the interests of readability, we drop the square brackets associated with the expectational operator as in \( E[ \cdot ] \) where the argument is obvious, but retain the brackets where the argument is a compound expression — see for instance the numerator and denominator of expression (3).
3. Utility generators and the universal $t$

For a fixed number $P$, consider the following utility function\(^3\) defined on security returns $y$:

$$U_p(y) = -\max(0, P - y); \quad -\infty < y < \infty.$$  \(4\)

The graph of $U$, plotted in Fig. 1, corresponds to the end of period pay-off to the writer of a put option with strike price $P$, with the initial premium income disregarded. Apart from this, note that $U'_p(y) \geq 0$ for all\(^4\) $y$ and the function $U_p$ is concave.

Let $F(R)$ denote the distribution of the benchmark return $R$. In what follows, we shall allow $F$ to have jumps, e.g. to refer to empirical distribution functions, but of course, it remains continuous to the right, similarly, denoted by $G(r)$, the return distribution function for the subject fund or portfolio. Integrating by parts,

$$EU_p(R) = -\int_{-\infty}^{P} (P - R) dF(R)$$

$$= -\int_{-\infty}^{P} F(R) dR.$$

Similarly,

$$EU_p(r) = -\int_{-\infty}^{P} G(r) dr.$$

Recall that $r$ dominates $R$ by SSD if and only if for any number $P$, $\int_{-\infty}^{P} G(r) dr \leq \int_{-\infty}^{P} F(R) dR$ (see Rothschild and Stiglitz (1970)). Hence, $r$ dominates $R$ if and only if:

$$E_r U_p(r) > E_R U_p(R); \quad \text{all } P.$$  \(5\)

As the expectational limiters indicate, the SSD condition is cast in terms of the marginal distributions of $r$ and $R$, so that aspects of the joint distribution are not involved. The SSD property means that $r$ will be preferred to $R$ by investors with any risk averse utility function, i.e. $r \succeq_{SSD} R \Rightarrow E_r U(r) > E_R U(R)$, any concave $U$. In this sense, the utility functions $U_p$ could be said to span the set of arbitrary concave utility functions. We now show that a similar sort of property holds for the equivalent margin measure, although joint distributions are now involved in a more essential way.

---

\(^3\)The fact that strictly positive utility is not attained is immaterial, as $U_p$ is equivalent in decision theoretic terms to $a + \beta U_p$ for $\alpha$ arbitrary and any $\beta > 0$.

\(^4\)Strictly, $U'_p(P)$ does not exist but the point has measure zero and we simply define it as unity.
The equivalent margin \( t(P) \) associated with the utility function \( U_p \) is defined by:

\[
t(P) = \frac{E[(e(R) - R)U'_p(R)]}{EU'_p(R)},
\]

where,

\[
U'_p(R) = \begin{cases} 
1 & R \leq P \\
0 & \text{otherwise}
\end{cases}
\]

As \( EU'_p(R) = F(P) \), it follows that,

\[
t(P) = \frac{1}{F(P)} \int_{-\infty}^{P} (e(R) - R) \, dF(R). \tag{6}
\]

In what follows, function (6) will be called the COMD between \( r \) and \( R \). Note in particular that,

\[
t(\infty) = \lim_{P \to \infty} t(P) = E_e[e(R) - R] \\
= E_e[E_e[r | R] - R] \\
= \mu_e - \mu_R,
\]

where \( \mu_e = E(r) \) and \( \mu_R = E(R) \) are the marginal means. Theorem 1 shows that the equivalent margin \( t_U \) for an arbitrary concave twice differentiable utility function \( U \) can be expressed as a weighted average of the generating \( t_p \).

\[\text{Notationally, one ought to use } t_{U_p} \text{ instead of } t(P) \text{ to be consistent with general notation } t_U \text{ for a utility function } U. \text{ However, } t(P) \text{ will be required at a later point to indicate a function as well as a functional, so the notation } t(P) \text{ is adopted at the outset, and will be used only in connection with the utility generator indexed by } P. \text{ In such contexts, the notation } t_{U_p}(P) \text{ or } t_r(P) \text{ is later used to distinguish the utility generators associated with different returns } R \text{ and } r.\]
Theorem 1. Suppose that \( \lim_{P \to \infty} U'(P) = U'(\infty) \) exists and let \( U \) be a twice differentiable utility function of returns for which \( U''(y) \leq 0, -\infty < y < \infty \). Then, the associated equivalent margin \( t_U \) is given by:

\[
t_U = \frac{U'(\infty)}{EU'(R)} t(\infty) + \int_{-\infty}^{\infty} w(P) t(P) \, dP, \tag{7}
\]

where,

\[
w(P) = -\frac{U''(P)}{EU'(R)} F(P). \tag{8}
\]

The weights \( w(P) \) are such that \( w(P) \geq 0 \) and \( \int_{-\infty}^{\infty} w(P) \, dP \leq 1 \), with equality if \( U'(\infty) = 0 \).

Proof. From Eq. (6) above and changing \( P \) to \( R \), \( (e(R) - R) \, dF(R) = \, d[\mu(R)F(R)] \), whence,

\[
t_U \mu'(R) = E[(e(R) - R) U'(R)]
= \int_{-\infty}^{\infty} U'(R) d[\mu(R)F(R)].
\]

Integrating by parts (permissible for Stieltjes integrals) and dividing both sides by \( EU'(R) \), we get:

\[
t_U = \frac{U'(\infty) t(\infty)}{EU'(R)} - \frac{1}{EU'(R)} \int_{-\infty}^{\infty} F(R) U''(R) t(R) \, dR. \tag{9}
\]

Let \( w(R) = -U''(R)/EU'(R) \, F(R) \). Then,

\[
\int_{-\infty}^{\infty} w(R) \, dR = -\frac{1}{EU'(R)} \left\{ \left[ F(R) U'(R) \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} U'(R) \, dF(R) \right\}
= -\frac{1}{EU'(R)} \left( U'(\infty) - EU'(R) \right)
= 1 - \frac{U'(\infty)}{EU'(R)}.
\]

As \( U' \) is a decreasing function, \( 0 \leq U'(\infty) \leq EU'(R) \). Finally, changing the dummy variable of integration in Eq. (9) back to \( P \), we find that \( \int_{-\infty}^{\infty} w(P) \, dP \leq 1 \), with equality if \( U'(\infty) = 0 \). □
Corollary. If \( t(P) \geq 0 \) for \( -\infty < P < \infty \), then \( t_U \geq 0 \) for an otherwise arbitrary, twice differentiable risk averse utility function \( U \).

Several remarks may be useful at this point.
1. If \( U'(\infty) = 0 \) (meaning global strict concavity), then by the second mean value theorem of differential calculus, there must exist a value \( P^* \) such that \( t_U = t(P^*) \). There is therefore an option style utility function that has exactly the same equivalent margin as the given utility function, strengthening the interpretation of the \( U_p \) as a set of “utility generators” for arbitrary concave utility functions.

2. An alternative expression for the weight \( w(P) \) is,
\[
w(P) = \frac{U''(P)}{EU''(R)} F(P),
\]
where \( \xi = -EU''(R)/EU''(R) \) is the Rubinstein (1973) coefficient of absolute risk aversion. The pattern of weights as \( P \) varies will therefore depend on two factors. The first is the second derivative of the utility function at \( P \) relative to the average. Consider, for example, the constant relative risk aversion family \( U(R) = (1 + R)^\xi / \xi \); \( \xi < 0 \), with \( \xi \to 0 \) indicating the logarithmic utility function \( \log(1 + R) \). For such functions, the weights \( w(P) \) diminish with \( 1/(1 + P)^{\xi - 2} \), so that the log utility function has weights that diminish more slowly relative to other members of the family, which correspond to greater risk aversion. The second factor is the cumulative probability mass up to \( P \). If \( U'' \) is diminishing while \( F(P) \) is increasing, a trade-off between the two factors is therefore implied, and an interior maximum of \( w(P) \) will usually exist. It is generally apparent that in certain zones, notably the smallest and highest value of \( P \), \( t(P) < 0 \) over some small interval will not necessarily conflict with \( t_U \geq 0 \) for any realistic \( U \).

4. The COMD and stochastic dominance

If \( r \) dominates \( R \) by SSD, then \( t(P) \geq 0 \), for all \( P \). Hence, if \( t(P) < 0 \) for some \( P \) then \( r \) cannot dominate \( R \), so that a necessary test for SSD emerges as a by-product of a COMD analysis. Moreover, one can create a sufficiency test for SSD. The trick here is to reverse the benchmark, so that \( R \) is evaluated against \( r \). If generator equivalent margins are all negative in the framework, then \( r \text{ SSD } R \).

With \( r \) as benchmark, define \( \gamma(r) = E[R \mid r] \), the conditional expectation of \( R \), given \( r \). The function \( \gamma(r) \) is thus the theoretical regression of \( R \) on \( r \), as distinct from \( e(R) \), which is the theoretical regression of \( r \) on \( R \). In such terms, the equivalent margin with respect to \( U_p \) of \( R \) with respect to \( r \) as the benchmark is,
\[
t_r(P) = \frac{E_r[(\gamma(r) - r)U_p'(r)]}{EU_p'(r)}, \tag{10a}
\]
while,

\[ t_g(P) = \frac{E[(e(R) - R)U^*_p(R)]}{EU^*_p(R)}. \]  

(10b)

is the equivalent margin of \( r \) with \( R \) as the benchmark. Operationally, the \( t_g(P) \) is done in terms of the COMD of \( r - R \) ordered by \( R \) values, while \( t_s(P) \) is computed as the COMD of \( R - r \), using the \( r \)-values to order the observation pairs. The principal result is as follows.

**Theorem 2.** If the COMD with \( r \) as benchmark is semi-negative for every \( P \), then \( r \) dominates \( R \) by SSD:

\[ t_r(P) \leq 0 \text{ for all } P \Rightarrow r > R. \]

(11)

**Proof.** As \( U_p(r) \) is concave,

\[ U_p(r) \leq U_p(R) + (r - R)U^*_p(R). \]

Taking expectations over the joint distribution of \( r, R \) gives:

\[
E_r U_p(r) \leq E_R U_p(R) + E_r[(E_r[r|R] - R)U^*_p(R)] \\
= E_R U_p(R) + E_r[(e(R) - R)U^*_p(R)] \\
= E_R U_p(R) + t_s(P)EU^*_p(R). \]

(12a)

Similarly,

\[ U_p(R) \leq U_p(r) + (R - r)U^*_p(r). \]

Taking expectations with respect to \( r, R \) gives:

\[ E_R U_p(R) \leq E_r U_p(r) + t_r(P)EU^*_p(r). \]

(12b)

Combining Eqs. (12a) and (12b) yields the bounds:

\[-t_r(P)EU^*_p(r) \leq E_r U_p(r) - E_R U_p(R) \leq t_s(P)EU^*_p(R). \]

(13)

Hence, if \( t_r(P) \leq 0 \), all \( P \), it follows that \( E_r U_p(r) \geq E_R U_p(R) \), and hence from Eq. (5) of Section 3, \( r > \text{SSD} R. \)

The operational implication is that one plots both versions (alternative benchmarks) of the COMD. If there is some \( P \) for which \( t_g(P) < 0 \), then \( r \) cannot be SSD over \( R \). Or if \( t_r(P) < 0 \), all \( P \), then \( r \) is SSD over \( R \).
5. Operational matters

The COMD function can be estimated by either non-parametric or parametric techniques, and the following discussion divides along these lines.

5.1. Non parametric: the OMD

Given a series of observation pairs \((r_i, R_i); i = 1, 2, \ldots N\) on a fund or portfolio return \(r\) and a benchmark return \(R\) for the same year or period, the OMD is the sample function defined by:

\[
\hat{t}(P) = \frac{1}{n(P)} \sum_{R_i \leq P} (r_i - R_i); \quad n(P) = \{\#R_i \leq P\}. \tag{14}
\]

In other words, the pairs of observations are first ordered by values of the benchmark return, and as \(P\) increases, \(\hat{t}(P)\) represents the running mean difference between \(r\) and \(R\) over observations for which \(R\) is less than the given value \(P\). Operationally, \(t\) is usually defined at the respective sample points \(R_j\) and we write \(\hat{t}_j = \hat{t}(R_j)\) as the values at those sample points.

The COMD function evaluated at a sample point \(R\) may be viewed as the expectation of the OMD at the chosen point, given the sample values. To see this, note first that the generator equivalent margin \(t(P)\) at any point \(R = P\) can be interpreted as the expected value of the regression difference \(e(R) - R\) computed with respect to the truncated distribution function \(\phi_p(R)\), for which,

\[
d\phi_p(R) = \frac{dF(R)}{F(P)}; \quad R \leq P
\]

\[
= 0; \quad R > P.
\]

Thus,

\[
\int_{-\infty}^{\infty} (e(R) - R)d\phi_p(R) = \frac{1}{F(P)} \int_{-\infty}^{P} (e(R) - R)dF(R)
\]

\[
= t(P).
\]

Hence, the equivalent margin for utility generator \(U_p\) may be interpreted as the mean area between \(e(R)\) and \(R\), computed with respect to \(\phi_p\) rather than the original \(F\).

Suppose now that a random sample \(R_1, R_2, \ldots R_N\) is drawn from the original distribution \(F\) and only those observations, \(n(P)\) in number, for which \(R_i \leq P\) are retained. Then the sum,

\[
\frac{1}{n(P)} \sum_{R_i \leq P} (e(R_i) - R_i),
\]

.
is a mean estimate for the regression difference \( e(R) - R \) over the truncated distribution. From the central limit theorem with appropriate regularity conditions (e.g. boundedness on \( e(R); R \leq P \)) it follows that,

\[
\frac{1}{n(P)} \sum_{R_i \leq P} (e(R_i) - R_i) \xrightarrow{a.s.} t(P), \quad \text{as } N \to \infty.
\]  

(15)

For any given sample point \( R_j \), consider the OMD statistic:

\[
\hat{t}(R_j) = \frac{1}{n(R_j)} \sum_{R_i \leq R_j} (r_i - R_i)
\]

\[
= \frac{1}{n(R_j)} \sum_{R_i \leq R_j} (e(R_i) - R_i) + \frac{1}{n(R_j)} \sum_{R_i \leq R_j} e_i,
\]  

where \( r = E[r \mid R_i] + e_i \), with \( E[e_i \mid R_i] = 0 \).

The first term on the right hand side of Eq. (16) is the mean estimate for the regression difference over the truncated distribution \( \phi(R_j) \); in other words, the best estimate of \( t(R_j) \) given the sample observations \( R_1, R_2, \ldots, R_N \), if one actually knew the theoretical regression \( e(R) \) at each sample point. The second term arises because one does not know what the theoretical regression function actually is. Correspondingly, there are two sources of error in \( \hat{t}(R_j) \) as an estimate of \( t(R) \). The first is in the mean estimate ("mean error"), and the second arises from the regression residual ("residual error"). A similar decomposition arises in prediction based upon classical regression theory (e.g. Goldberger, 1964, pp. 168–169).

Conditional on the given sample observations, the residual error attached to \( \hat{t}_j \) has variance \( \sigma^2/n(R_j) \), where it will be recalled that \( \sigma^2 = \text{Var}(e_i) \) with i.i.d. residuals \( e_i \). This means that for the smallest observation as ranked by \( R \), the residual variance is \( \sigma^2 \), diminishing thereafter with \( n_j \). Thus, a degree of uncertainty attaches to the empirical OMD over the lowest \( R \) values. The next technique avoids this difficulty, at the cost of having to specify or estimate the conditional expectation \( e(R) \).

5.2. Parametric approach

This proceeds by specifying a functional form \( e(R; \theta) \) for the theoretical regression of \( r \) on \( R \), which may be either linear or nonlinear in the \( k \times 1 \) vector of parameters \( \theta \). The parameters are then estimated by linear or nonlinear least squares, an appropriate method since one is attempting to estimate the conditional expectation of the dependent variable, in this case \( r \), in terms of the independent variable, \( R \). As the true functional form is typically unknown, it is advisable to try
a range of alternatives. In most cases, one will try variants of polynomial curve fitting, either in terms of the monic polynomials 1, \( R \), \( R^2 \), \( R^3 \), etc. or more conveniently, in terms of suitably constructed orthogonal polynomials; Section 6 contains an example of the latter technique. In addition it is advisable to check for the preferred version that the equation disturbance is indeed thoroughly purged of \( R \) by a test for regressor/disturbance correlation such as Ramsey’s (1969) reset test.

As before, the function \( t(P) \) has to be estimated at each of the given sample observations \( R_1, R_2, \ldots, R_N \). One could contemplate working out the function by theoretical integration of the fitted function \( e(R; \theta) \) together with a fitted distribution function \( \hat{F}(P) \), but such a procedure has some hazards of its own, notably those arising from the unknown behaviour of the theoretical regression outside the observed sample range of \( R \) values, a well-known problem in smoothing data with polynomial regression. The suggested procedure is therefore to work out the values of \( t_j = t(R_j) \) at each sample point \( j = 1, 2, \ldots, N \); ending up with a set of estimated values that can be compared with those obtained from the non-parametric approach. Thus, one derives:

\[
\hat{t}_j = \frac{1}{j} \sum_{i=1}^{j} (\hat{e}_i - R_i); \quad \hat{e}_i = e(R_i, \hat{\theta}).
\]  

(17)

as the estimate of \( t(P) \) for \( P = R_j; j = 1, 2, \ldots, N \).

The estimated variance (OLS) or asymptotic variance (NLS) of \( t_j \), conditional on the given set of benchmark observations \( R_1, R_2, \ldots, R_N \), can be obtained in terms of the conditional covariance matrix of \( \hat{\theta} \). The simplest case is the fitting of a polynomial or other regression linear in the parameters by means of OLS. To illustrate, one might set up the quadratic regression:

\[
e(R, \theta) = h + \theta_1 R + \theta_2 R^2,
\]

where \( e_i \) are the equation disturbances — for simplicity we assume in what follows that these are serially uncorrelated and homoscedastic, though neither are essential for the general procedure to work. Let \( G \) be the data matrix of right hand variables, i.e. \( g_i = (1, R_i, R_i^2) \) is the \( i \)th row of \( G \). The estimated covariance matrix of \( \hat{\theta} \) is \( \hat{\Omega} = \hat{\sigma}^2 (G'G)^{-1} \), where \( \hat{\sigma}^2 \) is the usual OLS estimate of \( \sigma^2 \). Assemble the running means of the right hand variables up to and including observation \( j \) into a vector of running means for each of the right hand variables:

\[
\bar{g}_j = (1, \bar{R}_j, \bar{R}_j^2); \quad \bar{R}_j = \frac{1}{j} \sum_{i=1}^{j} R_i; \quad \bar{R}_j^2 = \frac{1}{j} \sum_{i=1}^{j} R_i^2.
\]

Then the estimated variance of \( \hat{t}_j \) is given by,

\[
\hat{\text{Var}}(\hat{t}_j) = \bar{g}_j' \hat{\Omega} \bar{g}_j.
\]  

(18)
The values of $t$ can then be plotted over $j = 1, 2, \ldots, N$, with the estimated variances used to set confidence bands or “tunnels” above and below. An immediate extension is to specify,

$$r_j = \sum_{k=0}^{K} \theta_k g_{ik} + \epsilon_j,$$

where the $k$th constructed regressor function $g_k(R_j; R)$ may utilise the entire set of available $R$ values to form the values $g_{ik}$ of each regressor $g_k$ at $R_j$. Section 6 contains an example where the regressors $g$ are constructed as a set of orthogonal polynomials over the given sample, a technique which is superior to ordinary polynomial curve fitting. Again, Eq. (18) gives the variances of the fitted $t_j$, conditional on the sample $R$ values. The proof of the variance formula in each case is just a special case of that for the nonlinear case, below, and will not be presented separately.

The procedure for nonlinear least squares, where the parameters $\theta$ in enter the regression in nonlinear fashion, is very similar to that just described. For the asymptotic variance, the data matrix $G$ is replaced by the $N \times k$ Jacobian matrix $\tilde{A}$ of the partial derivatives of $e(R_j, \theta)$ with respect to $\theta$. A derivation of the $t_j$ estimates and the resulting variance estimator is given in Appendix A.

6. Empirical illustration

The empirical work illustrates the COMD and related assessment techniques with the historical performance of three New Zealand funds using monthly returns over the period December 1988 to November 1995, with data kindly supplied by FPG Research. The funds are of significant size and are among the relatively few long established NZ funds. The market rate $R$ is taken as the monthly percent change in the NZSE top 40 gross accumulation index.

The respective Jensen’s alpha for funds A, B, C over the period, computed using the 30-day bank bill rate as the risk-free rate were, $-0.09\%$, $-0.03\%$, and $-0.02\%$, and all were statistically quite insignificant. Thus, there is little indication on the basis of Jensen’s alpha of managerial superiority over the market benchmark in any of the funds. None of the funds’ second order stochastically dominate the market, by either the classic cross over rule for the empirical distribution functions or the reversed $t$ procedure of Section 4 of the present paper, though funds B and C both come quite close.

Fig. 2a–c are fitted COMD functions for the three funds. In each case, the top diagram is a scatter plot of the fund return $r$ against the market return $R$, the middle diagram is the COMD function estimated by the non-parametric technique, while the bottom diagram is the parametric estimate with confidence bands of $\pm \text{STD } t_j$, i.e. one standard deviation.
The parametric estimate is prepared by a variant of polynomial curve fitting that uses Forsyth sample orthogonal polynomials (see Carnahan et al., 1969, Chap. 8.22). These simplify the task of choosing the order of the polynomial of best fit.

Fig. 2. (a) COMD: fund A against market benchmark. (b) COMD: fund B against market benchmark. (c) COMD: fund C against market benchmark.
The polynomials are designed so that the regression matrix of sums of squares and cross products $G'G$ is diagonal. One can then choose the order of the polynomial as that of the highest order Forsyth polynomial that remains significant, either in terms of ordinary Student’s $t$ or else more automatically in terms of simple
stepwise regression inclusion rules. Appendix B gives the formulas for the Forsyth polynomials, while Table B1 of the appendix lists the results from a stepwise regression procedure. In the case of fund A, Forsyth polynomials of degrees 1, 3 and 5 were stepped in and retained as significant. For fund B, only the polynomial
### Table B1
Stepwise regression results

#### Fund A

<table>
<thead>
<tr>
<th>Step number</th>
<th>Variable label</th>
<th>Status</th>
<th>F-value</th>
<th>df numerator</th>
<th>df denominator</th>
<th>F-probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G1</td>
<td>stepped in</td>
<td>725.3896</td>
<td>1</td>
<td>82</td>
<td>0.000000</td>
</tr>
<tr>
<td>2</td>
<td>G5</td>
<td>stepped in</td>
<td>7.3258</td>
<td>1</td>
<td>81</td>
<td>0.008287</td>
</tr>
<tr>
<td>3</td>
<td>G3</td>
<td>stepped in</td>
<td>4.5410</td>
<td>1</td>
<td>80</td>
<td>0.036163</td>
</tr>
</tbody>
</table>

**Summary for potential variables not entered into the reg. equation**

- G2 if entered: 0.0327, 1, 79, 0.857016
- G4 if entered: 0.0129, 1, 79, 0.909926
- G6 if entered: 2.4081, 1, 79, 0.124706
- G7 if entered: 0.3868, 1, 79, 0.535785

**End of stepping sequence**

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Estimated coefficient</th>
<th>Student’s t-ratio 80 df</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>0.83766</td>
<td>28.56</td>
</tr>
<tr>
<td>G3</td>
<td>0.58580 × 10^{-3}</td>
<td>2.131</td>
</tr>
<tr>
<td>G5</td>
<td>0.14281 × 10^{-4}</td>
<td>2.765</td>
</tr>
<tr>
<td>Constant</td>
<td>0.79498</td>
<td>4.714</td>
</tr>
</tbody>
</table>
Table B1 (continued)

Fund B

<table>
<thead>
<tr>
<th>Step number</th>
<th>Variable label</th>
<th>Status</th>
<th>$F$-value</th>
<th>df numerator</th>
<th>df denominator</th>
<th>$F$-probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G1</td>
<td>stepped in</td>
<td>501.2995</td>
<td>82</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Summary for potential variables not entered into the reg. equation

<table>
<thead>
<tr>
<th>Variable label</th>
<th>$F$-value</th>
<th>df numerator</th>
<th>df denominator</th>
<th>$F$-probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>G2</td>
<td>0.1932</td>
<td>1</td>
<td>81</td>
<td>0.661414</td>
</tr>
<tr>
<td>G3</td>
<td>1.9815</td>
<td>1</td>
<td>81</td>
<td>0.163057</td>
</tr>
<tr>
<td>G4</td>
<td>0.0395</td>
<td>1</td>
<td>81</td>
<td>0.842938</td>
</tr>
<tr>
<td>G5</td>
<td>0.4165</td>
<td>1</td>
<td>81</td>
<td>0.520522</td>
</tr>
<tr>
<td>G6</td>
<td>2.4766</td>
<td>1</td>
<td>81</td>
<td>0.119453</td>
</tr>
<tr>
<td>G7</td>
<td>0.3339</td>
<td>1</td>
<td>81</td>
<td>0.564946</td>
</tr>
</tbody>
</table>

End of stepping sequence

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Estimated coefficient</th>
<th>Student’s t-ratio 82 df</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>0.74262</td>
<td>22.39</td>
</tr>
<tr>
<td>Constant</td>
<td>0.83854</td>
<td>4.398</td>
</tr>
</tbody>
</table>

(continued on the next page)
Choice of the order of the Forsyth polynomials was executed using stepping in regression option in the SHAZAM econometric package (White, 1977 and updates).

Variables stepped in: Fund A — G1, G5, G3; fund B — G1; fund C — G1, G6.

Criteria for stepping in/out: if a variable becomes significant at the 5% level, it is included (probability values are based on the F values for each regression). If a variable becomes less significant than 5%, then it is deleted from the equation.

Fund A: $R^2 = 0.9119$, $\bar{R}^2 = 0.9086$; variance of the estimate $\sigma^2 = 2.3887$.
Fund B: $R^2 = 0.8594$, $\bar{R}^2 = 0.8577$; variance of the estimate $\sigma^2 = 3.054$.
Fund C: $R^2 = 0.5276$, $\bar{R}^2 = 0.5160$; variance of the estimate $\sigma^2 = 14.752$.

### Table B1 (continued)

#### Fund C

<table>
<thead>
<tr>
<th>Stepping sequence</th>
<th>Variable label</th>
<th>Status</th>
<th>F-value</th>
<th>df numerator</th>
<th>df denominator</th>
<th>F-probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G1</td>
<td>stepped in</td>
<td>80.1965</td>
<td>1</td>
<td>82</td>
<td>0.000000</td>
</tr>
<tr>
<td></td>
<td>G6</td>
<td>stepped in</td>
<td>5.6898</td>
<td>1</td>
<td>81</td>
<td>0.019398</td>
</tr>
</tbody>
</table>

**Summary for potential variables not entered into the reg. equation**

- G2 if entered: F-value 1.9529, df 1, 80, F-probability 0.166138
- G3 if entered: F-value 0.0011, df 1, 80, F-probability 0.973787
- G4 if entered: F-value 2.3265, df 1, 80, F-probability 0.131129
- G5 if entered: F-value 1.8825, df 1, 80, F-probability 0.173880
- G7 if entered: F-value 0.4060, df 1, 80, F-probability 0.525838

End of stepping sequence

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Estimated coefficient</th>
<th>Student's t-ratio</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>0.67122</td>
<td>9.208</td>
<td>81</td>
</tr>
<tr>
<td>G6</td>
<td>$4.44150 \times 10^{-5}$</td>
<td>2.385</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.4191</td>
<td>2.013</td>
<td></td>
</tr>
</tbody>
</table>

Variables chosen: G1, G5, G3
of degree 1 was stepped in, so that for this fund, a linear regression is a very good fit. Fund C stepped in orders 1 and 6. All this means that one has to go to a 5th-order polynomial to adequately represent fund A against the market benchmark, to order 1 for fund B, and to order 6 for fund C. There is no need to transform back to the natural or monic polynomials, as the various estimates can

Fig. 3. Equivalent margin as function of \( x_i \). A: Non-parametric (actual \( r \)). B: Non-parametric (estimated \( e[r] \)).
all be expressed directly in terms of the Forsyth polynomials.\textsuperscript{6} In all cases, Ramsey reset tests indicated no residual regressor–residual correlation at 5% significance levels for the resulting $\chi^2$ tests.

Both parametric and non-parametric techniques yield closely similar results for the COMD functions, in each case. Funds B and C are virtually COMD dominant over the market, failing only over the large or very small range of returns. Since the latter receives only a small empirical probability, it is evident that unless the degree of risk aversion was either very high or very low, investors would prefer these funds over the market. On the other hand, the performance of fund A relative to the market is much less convincing.

We can check the possible influence of the degree of investor risk aversion by choosing a family of utility functions and varying the risk aversion parameter(s). A convenient family is the constant relative risk aversion family mentioned in Section 3. We varied the parameter $\xi$ from 0 (which is effectively the log utility function) on down to $-5$, moving from mild to progressively more stringent risk aversion. For each value of $\xi$, the equivalent margin $t_\xi$ was calculated for the three funds, and the results plotted in Fig. 3 against $\xi$. It is clear that at least for the CRRA family, fund C generates most investor surplus relative to the market, while for lower degrees of risk aversion, the poor performance of fund A is self-evident.

As indicated in Section 2, the equivalent margin measures are primarily designed for testing against the market as benchmark, but they do provide a bilateral measure for comparing two funds if one is willing to accept the discount factor $\pi(R)$ as defined in terms of market returns. Thus, to compare fund B with fund C in terms of logarithmic utility one would compute:

$$
\sum_{i=1}^{N} \pi(R_i)(r_{B_i} - r_{C_i}); \quad \pi(R) = \frac{(1 + R_i)^{-1}}{\sum_{j=1}^{N} (1 + R_j)^{-1}}.
$$

\textsuperscript{6}While OLS is the correct method to use in estimating the conditional expectation function $e(R)$, an operational problem arises in determining the degree of the approximating polynomial. In most packages, stepping in is determined in terms of tests of significance that are based on independent spherical Normal distributions for the regression residuals. Financial returns data is often not so obliging, (e.g. Engle and Ng, 1993; Glosten et al., 1993), and indeed the core methodology of the present paper itself makes no such assumption. Thus, it is advisable to test for departures form Normality etc in the residuals and if thought necessary to adapt the stepping in algorithm somehow. In the present instance, none of the fund residuals resulting from the stepped in choices for $e(R)$ were serially correlated (DW statistics of 2.11, 2.00 and 1.91, respectively), but they did exhibit some positive skewness and kurtosis, also heteroscedascity, though only marginally significant in most standard tests except possibly for the ARCH test (e.g. chi sq. of 4.56 with 1 df for fund C). The general problem here will be to adapt the hypothesis testing of the OLS fitting for such departures (rather than adapt the estimation methodology, as OLS in itself remains correct). Thanks go to a referee for pointing out a potential problem here.
The resulting figure was \(-0.024\%\), virtually zero, so that for such an investor there would be little to choose between the funds, in spite of the fact that C does a little better relative to the market as benchmark than does B. This indicates if nothing else that a clear purpose is necessary for the proposed comparison. In the present paper, we have emphasised the role of the market as benchmark, simply because investors are so often concerned with the question of whether a passive holding of the market index would result in better returns than employing a fund manager.

7. Concluding remarks

The equivalent margin approach is anchored in the idea that the most natural way to judge performance is in comparison with a suitable chosen benchmark. The benchmark itself generally emerges from the purpose or context of the exercise. In the case of fund performance, it is natural to ask whether the fund in question generates investor surplus over and above what the investor could get simply by investing in the market index. In such a context, market timing of the fund manager is an important generator of excess returns, relative to the market. As remarked in Section 1 of the paper, conventional measures of performance are distorted in the presence of market timing skills. Market timing itself is most naturally assessed in terms of conditionality — given that the market return in any period, has the fund in question done any better? In focussing on the conditional expectation function the equivalent margin approach essentially operates in terms of the investor surplus as the area between the conditional expectation function and the 45 line. In its emphasis on areas, it is therefore reminiscent of the consumer surplus idea of welfare economics. The mathematical formulation is based on that insight, one which turns out also to have connections to the modern theory of financial general equilibrium.

The empirical work of the paper, while somewhat limited in scope, has shown that a clear picture of fund superiority relative to the market can emerge, in situations where classic stochastic dominance assessments yield more equivocal results. The latter approach disregards the conditionality that is inherent in market timing, and does not therefore utilise important information. Moreover, the equivalent margin results in dollar type measures for how much the investor can potentially gain from the use of the fund instead of the benchmark index or fund. Relative to the stochastic dominance approach, it can therefore be regarded in the light of a cardinal measure, rather then simply an ordinal one.

There is room for some technical improvements in the estimation theory associated with the equivalent margin. A current weakness is that low values of benchmark return are associated with only a few observations — the number of observation effectively builds up with higher values. This leads to a high sampling.
variability in such bands — in Fig. 2, the confidence limits look like a funnel rather than a tunnel. It might be possible to adopt a Bayesian approach whereby a set of artificial observations is added at the beginning for very low values of the benchmark return and set at an uninformative zero so far as the difference between fund and benchmark return is concerned. The same sort of effect might possibly be achieved by means of bootstrapping the early actual values. However, such developments must be left for further research.

Acknowledgements

The author is grateful to FPG Research, who kindly provided the data and access software. Thanks go also to Ross Tucker, for capable research assistance. Comments by three referees helped to improve the final version.

Appendix A

The following result establishes the asymptotic distribution of the parametric estimate of \( t \) for nonlinear least squares, assuming that the underlying regression model is as specified in the text (see Eq. (2), Section 2), and the functional form is \( e(R; \theta) \) correctly chosen.

Proposition A1 (NLS). Suppose that the theoretical regression \( e(R, \theta) \) is differentiable in \( \theta \) up to order 2, and let \( e(\theta) \) be the \( N \times 1 \) vector whose \( i \)th element is \( e(R_i, \theta) \). Let \( G(\theta) = \delta e / \delta \theta^t \) be the data Jacobian evaluated at some value \( \theta \). Considering the columns of \( G(\theta) \), define for each the running means up to and including a given observation \( R_j \) and for each such \( j \), assemble the results into a \( k \times 1 \) row vector \( \bar{g}_j(\theta) \). Suppose that the parameter space is compact, and the \( j \)th elements of \( G(\theta) \) are bounded, uniformly in \( R \), in some neighbourhood of the true value of \( \theta \). Then the nonlinear least squares estimate of \( t_j \) is consistent and asymptotically Normal:

\[
\sqrt{N}(\hat{t}_j - t_j) \overset{d}{\to} N(\theta, \alpha^2 \bar{g}_j \text{plim}[N(G'G)^{-1}] \bar{g}_j),
\]

where \( \bar{g}_j \) and \( G \) are evaluated at the true value of \( \theta \).

Proof. The demonstration essentially adapts the standard theory of nonlinear regression (e.g. Jennrich, 1969) to the present context. From Eq. (17) of the text,

\[
\hat{t}_j = \frac{1}{j} \begin{pmatrix} 1 \end{pmatrix} (e(\hat{\theta}) - R).
\]
where \( I_j \) is the unit vector of dimension \( j \). Hence,

\[
\sqrt{N} \left( \sum_{j} t_j \right) = \frac{1}{\sqrt{N}} \left( \mathbf{1}_j \cdot \mathbf{e} \right) \left( (\hat{\theta} - \theta) - e(\theta) \right).
\]

From the exact version of Taylor’s theorem, the right hand side is equal to:

\[
\frac{1}{\sqrt{N}} \left( \mathbf{1}_j \cdot \mathbf{e} \right) G(\theta^*) \sqrt{N} (\hat{\theta} - \theta),
\]

where \( \theta^* \) lies on the line joining \( \hat{\theta} \) and \( \theta \). Under the stated hypotheses, \( \hat{\theta} \) converges almost surely to the true \( \theta \), and

\[
\frac{1}{\sqrt{N}} \left( \mathbf{1}_j \cdot \mathbf{e} \right) G(\theta^*) = \tilde{g}_j(\theta^*),
\]

converges in probability to \( \tilde{g}(\theta) \). As a standard NLS result under the stated assumptions, \( \sqrt{N} (\hat{\theta} - \theta) \) converges in distribution to a \( N(0, \sigma^2 \text{plim} \left[ N(G'G)^{-1} \right]) \) variable, and the desired result follows.

\[\square\]

Appendix B

B.1. Forsyth orthogonal polynomials

Given a set of observations \( R_1, R_2, \ldots, R_N \) for the independent variable \( R \), a set \( g_0, g_1, g_2, \ldots, g_K \) of regressors are constructed which have the property that if \( G \) is the corresponding data matrix, of order \( N \times (K + 1) \), then \( G'G \) is orthogonal, in the sense that the off diagonal elements are all zero. The order \( K \) is chosen at will. The Forsyth polynomials are versions of the orthogonal polynomials that represent increasing powers of \( R \). They are constructed as follows:

For observations \( i = 1, 2, \ldots, N \), define: \( g_{i,-1} = 0 \) (not used among the regressors) and \( g_{i,0} = 1 \); for \( k = 1, 2, \ldots, K \):

\[
g_{i,k} = (R_i - \varphi_k) g_{i,k-1} - \rho_k g_{i,k-2},
\]

where,

\[
\varphi_k = \frac{\sum_{i=1}^{N} R_i g_{i,k-1}^2}{\sum_{i=1}^{N} g_{i,k-1}^2}; \quad \rho_k = \frac{\sum_{i=1}^{N} R_i g_{i,k-1} g_{i,k-2}}{\sum_{i=1}^{N} g_{i,k-2}^2}.
\]
It will be noted that the Forsyth polynomials are of increasing order in powers of \( R \).

**B.2. Choice of order**

Table B1 is the output from a stepwise regression of the respective funds A, B, C on the market that selects regressors — the Forsyth polynomials — on the basis of progressive contribution to the fitted sum of squares.

**References**


Bowden, R.J., 1992. Welfare economics and fund performance: the equivalent margin and associated measures. Mimeo, School of Banking and Finance, the University of New South Wales.


