Compatibility of Expected Utility and μ / σ Approaches to Risk for a Class of Non Location-Scale Distributions

By Gerry Boyle and Denis Conniffe

Abstract

Proofs of compatibility of the expected utility and μ / σ approaches to incorporating uncertainty in decision making exist for at least some utility functions and location-scale distributions. But there are severe constraints and it is desirable to investigate compatibility more widely. We do so for the class of distributions that are transformable to location-scale form by concave transformation and where the utility functions remain concave under transformation. The class is important, containing distributions such as the lognormal and Pareto, usually considered more appropriate for modelling income or wealth than those in the location-scale family.

1. INTRODUCTION

As is well known, two frequently employed approaches to incorporating uncertainty about some key variable – often income or wealth - in analyses of the comparative statics of optimal decision-making are μ / σ analysis and maximisation of expected utility. The former approach takes uncertainty as representable by the standard deviation of the variable. The decision maker is assumed to proceed by constrained maximisation of some function of the mean and standard deviation, which is increasing in the mean, decreasing in the standard deviation (the monotonicity conditions) and quasiconcave in μ and σ. The expected utility approach commences from a utility function which is monotonically increasing and concave in the variable. Uncertainty is embodied in a probability distribution for the variable and the decision maker evaluates outcomes in terms of expected utility. Elementary texts often remark that the approaches are equivalent for either a quadratic utility function or normality of distribution, although the latter case is really greatly constrained by being conditional on the existence of the expected utility. For other situations, the consistency of the two approaches has been debated at considerable length in the literature, which is so extensive that only publications directly relevant to the theme of this paper will be mentioned.

Meyer (1987) and Sinn (1989) showed that equivalence of expected utility maximisation and μ / σ analysis could be extended from the normal distribution to the location-scale family of distributions although again this is actually conditional on the existence of expected utilities. This is one of the key issues in discussing equivalence. Some of the most frequently advocated utility functions u(x) do not have expectations under normality, or some other location-scale distributions, but do under plausible distributions that are not in the location-scale family. For example,

$$u(x) = \log x \quad \text{or} \quad u(x) = x^{\lambda}, \text{ with } 0 < \lambda < 1,$$

do not have expectations under normality. Nor have they expectations under other location-scale distributions where x can be ≤ 0, such as the Gumbel. There are location-scale distributions for which x must be positive, such as the two parameter uniform and the two parameter exponential, but these are surely implausible as models for income or wealth.

However, the expectations do exist for several two parameter distributions that are not of location-scale form and that are plausible as models for income or wealth. The Pareto distribution was one of the first advocated (for example, Arnold, 1983) for that purpose[1], but it is not in the location-scale family. The log normal has been frequently employed for income modelling, justified by both theoretical (for example, Aitchison and Brown, 1957) and empirical findings and, of course, it is not of location-scale type. Other non location-scale two parameter distributions and various multiparameter distributions have also been proposed (for example, Bandourian, McDonald and Jurley, 2003). Meyer's insight that if an originally multiparameter non location-scale distribution is employed in circumstances where all parameters except the location and scale parameters are held constant it effectively becomes a location-scale distribution, does establish equivalence for some applied problems. But this situation, termed the 'location scale condition', must be of limited occurrence.

So it is important to investigate if the equivalence between the expected utility function and $\boldsymbol{\mu}/\boldsymbol{\sigma}$ approaches holds for the distributions for which the expectations exist, but that are not in the location-scale family. The approach in this paper is to investigate two parameter distributions of x that are not location-scale, but where the distribution of y = h(x) is location scale, where h(x) is a monotonically increasing and concave function of x. For example, if x is log normal, y = log x is normal and if x is Pareto, y = log x follows a two parameter exponential and this, like the normal, is a location-scale distribution. If we also require that our utility functions are concave when expressed as functions of y, we can show that equivalence will hold in the sense that $E(u(x)) = V(\mu, \sigma)$ satisfies the monotonicity conditions throughout (σ, μ) space and is quasiconcave in μ and σ in a region of that space. In our proofs we will exploit Meyer's (1987) results for the location-scale case, although we will need to broaden his proof somewhat. The location-scale family is defined by the density

[1] Nowadays it is recognised that the Pareto is unsuitable as a general income distribution, because it does not fit well at low incomes, although it does fit well to a population of higher income groups. But that might be the relevant population in contexts such as investment.

$$\frac{1}{\alpha} f\left(\frac{x-\beta}{\alpha}\right),$$

where β is the location and α the scale parameter. The mean and standard deviation are $\mu = \beta + \alpha k_1$ and $\sigma = \alpha k_2$ respectively, where k_1 and k_2 are constants. Meyer assumed β and α equal to μ and σ, which would be true for a normal, but not for some other location-scale distributions, and the lower and upper variable bounds independent of parameters. These assumptions are not necessary and would, for example, prevent us from employing the two parameter exponential, where the lower bound depends on the location parameter. Our somewhat more general proof is given in Appendix A.

2. CONCAVE TRANSFORMATIONS TO LOCATION-SCALE DENSITIES

As already mentioned, y = h(x) is monotonically increasing and concave and y has a location-scale distribution. The case y = x corresponds to x location-scale. Suppose the inverse function is $x = h^{-1}(y) = g(y)$. Clearly, any number of distributions of x with the required property can be found by obtaining the distributions of $x = h^{-1}(y) = g(y)$ with h(x) any appropriate function and y following any of the location-scale distributions. But whether these distributions of x are good fits for income or wealth distributions is another matter. Some might be, even if they have not appeared in the literature, but probably many would not be. So, although much of the paper maintains the generality of y = h(x), our examples will use y = log x, where we know appropriate distributions exist.

Suppose the utility function u(x) =u(g(y)) is also concave in y. Examples implying familiar utility functions are:

$$u(x) = \log x - b(\log x)^2,\ x > 1, \quad \rightarrow \quad u(y) = y - by^2, \quad y < 1/2b$$

$$u(x) = (\log x)^{\gamma}, \quad x > 1, \quad \rightarrow \quad u(y) = y^{\gamma}, \quad 0 < \gamma < 1$$

$$u(x) = A - \lambda e^{-\lambda \log x}, \quad x > 0, \quad \rightarrow \quad u(y) = 1 - e^{-\lambda y}, \quad \lambda > 0$$

$$u(x) = \log\log x, \quad x > 1, \quad \rightarrow \quad u(y) = \log y.$$

Since y = h(x) has a location-scale distribution, we know that the expectation of u(y), $V(\mu^*, \sigma^*)$, has[2] the required properties for monotonicity and concavity with respect to the mean μ^* and standard deviation σ^* of the variable y. Of course, the expectation of u(x) with respect to the distribution of x must also be $V(\mu^*, \sigma^*)$, since it amounts to just a change of variable in integration. However, we are interested in its properties with respect to μ and σ and not μ^* and σ^*. We will investigate these in the next two sections.

Since there are evidently concave utility functions of x that will not be concave in terms of y, we are restricting the class of utility functions somewhat. Some may be totally excluded, while others may require restriction of the range of the variable or parameters. For example,

$$u(x) = x^{\alpha} = e^{\alpha \log x}, \quad x > 0, \qquad 0 < \alpha < 1,$$

is concave in x, but with y = log x,

$$u(y) = e^{\alpha y}$$

is not concave in y as its second derivative is positive. However,

$$u(x) = 1 - e^{-\lambda x}, \quad \text{x > 0.}$$

with y = log x would give

$$u(y) = 1 - e^{-\lambda e^{y}} \quad .$$

and then

$$\frac{\partial^2 u}{\partial y^2} = \lambda e^{y} e^{-\lambda e^{y}} \left(1 - \lambda e^{y}\right).$$

which is negative if the bracketed term is. So concavity requires $\lambda > 1$, or y > - log λ.

3. THE MONOTONICITY CONDITIONS

Since $V(\mu^*, \sigma^*)$ satisfies monotonicity conditions we know that

$$\frac{\partial V}{\partial \mu^*} \quad \text{and} \quad \frac{\partial V}{\partial \sigma^*}$$

---

[2] Assuming V exists, of course. $E(\log \log x) = E(\log y)$ will not exist if y is normal, that is, if x is log normal, but it will if y is two parameter exponential, that is, if x is Pareto.

are positive and negative respectively. Now

$$\frac{\partial V}{\partial \mu} = \frac{\partial V}{\partial \mu^*}\frac{\partial \mu^*}{\partial \mu} + \frac{\partial V}{\partial \sigma^*}\frac{\partial \sigma^*}{\partial \mu}$$

and

$$\frac{\partial V}{\partial \sigma} = \frac{\partial V}{\partial \mu^*}\frac{\partial \mu^*}{\partial \sigma} + \frac{\partial V}{\partial \sigma^*}\frac{\partial \sigma^*}{\partial \sigma}.$$

So if

$$\frac{\partial \mu^*}{\partial \mu} \quad \text{and} \quad \frac{\partial \sigma^*}{\partial \sigma} \tag{1}$$

are positive, as is plausible from y = h(x) being a monotonically increasing function of x, and

$$\frac{\partial \mu^*}{\partial \sigma} \quad \text{and} \quad \frac{\partial \sigma^*}{\partial \mu} \tag{2}$$

are negative, as is plausible from y = h(x) being concave in x, the first derivatives of V with respect to μ and σ will be positive and negative respectively. The detailed proof is given in Appendix B. Of course, (1) and (2) imply the slope of an indifference curve

$$S = \frac{d\mu}{d\sigma} = -\frac{\partial V}{\partial \sigma} / \frac{\partial V}{\partial \mu}$$

is positive.

For a first example, we take the case of x log normally distributed, so that y = log x is normally distributed and the utility function as u(x) = log x. This might seem a rather trivial utility function to choose, since then u(y) = y, which is on the limit of concavity, while $V(\mu^*, \sigma^*) = \mu^*$ and the slope of the indifference curve in (σ^*, μ^*) space is zero. However, the case is important, both historically and because many textbooks on investment or portfolio theory (for example, Elton and Gruber, 1995, p.234) state that expected utility and $\boldsymbol{\mu} / \boldsymbol{\sigma}$ analysis are equivalent for a log utility function given log normality. The claim is based on results in Elton and Gruber (1974), but the case had previously been analysed by Feldstein (1969) and commented on by others.

Statistical textbooks write the parameters of the log normal as the mean and variance of the log of the variable, that is, of μ^* and σ^*. The mean and variance of the lognormal are

$$\mu = e^{\mu^* + \frac{1}{2}\sigma^{*2}}$$

and

$$\sigma^2 = e^{2\mu^* + \sigma^{*2}}\left(e^{\sigma^{*2}} - 1\right).$$

Expressing μ^* in terms of μ and σ gives

$$\mu^* = \log\mu - \frac{1}{2}\log\left(1 + \frac{\sigma^2}{\mu^2}\right) \tag{3}$$

and

$$\sigma^{*2} = \log\left(1 + \frac{\sigma^2}{\mu^2}\right). \tag{4}$$

Since μ^* is also $V(\mu,\sigma)$, differentiation of (3), as performed by Elton and Gruber (1974), gives

$$\frac{\partial V}{\partial \mu} = \frac{\mu^2 + 2\sigma^2}{\mu(\mu^2 + \sigma^2)} \quad \text{and} \quad \frac{\partial V}{\partial \sigma} = \frac{-\sigma}{\mu^2 + \sigma^2},$$

which are positive and negative respectively. Utility is increased by increasing μ for fixed σ or decreasing σ for fixed μ. So if the set of possible values in (σ,μ) space is bounded by a concave frontier the expected utility maximum will lie upon that frontier. The issue of precisely where on the frontier depends on how the slope of the indifference curve

$$S = \frac{d\mu}{d\sigma} = \frac{\mu\sigma}{\mu^2 + 2\sigma^2} \tag{5}$$

changes along the curve and this will be returned to in the next Section. Of course, the signs of the derivatives of V with respect to μ and σ were already guaranteed by the general proof in Appendix B. In fact, for every utility function u(x) = u(g(y)), concave in y as well as x, with y = h(x) following a location-scale distribution, expected utility is compatible with $\boldsymbol{\mu}/\boldsymbol{\sigma}$ analysis in this sense of the maximum occurring on the frontier.

There are obviously very many possibilities, but keeping to lognormal x, one interesting case is

$$u(x) = A - \lambda x^{-\lambda} = A - \lambda e^{-\lambda \log x} = A - \lambda e^{-\lambda y}.$$

where $\lambda > -1$. In terms of y, this is the very frequently employed constant absolute risk aversion utility function. From the well known moment generating function of the normal distribution

$$V(\mu^*,\sigma^*) = A - \lambda e^{-\lambda\mu^* + \frac{1}{2}\lambda^2\sigma^{*2}}.$$

It may be worth remembering that $S^* = d\mu^* / d\sigma^* = \lambda\sigma^*$, which is not a function of μ^* in accordance with constant absolute risk aversion. Substituting (3) and (4) into $V(\mu^*, \sigma^*)$ gives

$$V(\mu,\sigma) == A - \lambda e^{-\lambda \log\mu + \frac{\lambda}{2}(\lambda+1)\log(1+\sigma^2/\mu^2)}, \qquad (6)$$

with

$$\frac{\partial V}{\partial \mu} = (A - V)\frac{\lambda}{\mu}\frac{\left[\mu^2 + (\lambda+2)\sigma^2\right]}{\mu^2 + \sigma^2},$$

$$\frac{\partial V}{\partial \sigma} = -(A - V)\frac{\lambda(1+\lambda)\sigma}{\mu^2 + \sigma^2}$$

and

$$S = \frac{(1+\lambda)\mu\sigma}{\mu^2 + (\lambda+2)\sigma^2}. \qquad (7)$$

So S is a function of μ and since

$$\frac{\partial S}{\partial \mu} = S\frac{(\lambda+2)\sigma^2 - \mu^2}{\mu\left[\mu^2 + (\lambda+2)\sigma^2\right]}. \qquad (8)$$

This will be negative, showing decreasing absolute risk aversion, which is often considered plausible, if $\mu^2 > (\lambda+2)\sigma^2$. It could be positive for low mean income (or high λ and variance), but as will be seen in the next section, quasiconcavity requirements are then infringed. Sinn (1989, p.152) presents a utility function of the form (6), although his derivation followed a quite different path.

Assuming log normality is not essential, of course. Suppose x has a Pareto distribution

$$\theta\phi^\theta \frac{1}{x^{\theta+1}} \quad x \geq \phi.$$

The mean and variance of the distribution are

$$\mu = \frac{\theta\phi}{\theta - 1} \quad \text{and} \quad \sigma^2 = \frac{\theta\phi^2}{(\theta-1)^2(\theta-2)}$$

The log of a Pareto has the (location-scale) two parameter exponential density

$$f(y) = \frac{1}{\alpha}e^{\frac{-(y-\beta)}{\alpha}}, \quad \beta < y < +\infty,$$

where $\alpha = 1/\theta$ and $\beta = \log\phi$. The mean and variance of the exponential are

$$\mu^* = \alpha + \beta \quad \text{and} \quad \sigma^{*2} = \alpha^2.$$

Taking u(x) = log x gives $V(\mu^*,\sigma^*) = \mu^*$ as in the log normal case, but expressing μ^* in terms of μ and σ^2 gives instead of (3)

$$V(\mu,\sigma) = \mu^* = \log\left\{\frac{\sigma^2}{\mu}\left[1+\frac{\mu^2}{\sigma^2}-\sqrt{1+\frac{\mu^2}{\sigma^2}}\right]\right\}+\frac{\sigma^2}{\mu^2}\left(\sqrt{1+\frac{\mu^2}{\sigma^2}}-1\right).$$

Again, taking

$$u(x) = A-\lambda x^{-\lambda} = A-\lambda e^{-\lambda\log x} = A-\lambda e^{-\lambda y}$$

and noting that

$$\int_\beta^\infty \frac{\lambda}{\alpha}e^{-\lambda y-\frac{1}{\alpha}(y-\beta)}dy = \frac{\lambda}{\alpha\lambda+1}e^{-\beta\lambda},$$

we get

$$V(\mu,\sigma) = A-\frac{\lambda\left(1+\sqrt{1+\mu^2/\sigma^2}\right)}{\left(1+\sqrt{1+\mu^2/\sigma^2}\right)+\lambda}\left\{\frac{\sigma^2}{\mu}\left[1+\frac{\mu^2}{\sigma^2}-\sqrt{1+\frac{\mu^2}{\sigma^2}}\right]\right\}^{-\lambda}$$

instead of (6). Many other $V(\mu,\sigma)$ can be obtained corresponding to various u(x) and distributions of x, although, depending on the tractability of the integrals involved, mathematical expressions can sometimes be difficult.

4. THE QUASICONCAVITY CONDITIONS

We also want quasiconcavity of V with respect to μ and σ, so that the indifference curves are convex. The quasiconcavity condition is non-negativity of

$$\left[-\frac{\partial^2 V}{\partial\sigma^2}\left(\frac{\partial V}{\partial\mu}\right)^2+2\frac{\partial^2 V}{\partial\mu\partial\sigma}\left(\frac{\partial V}{\partial\mu}\right)\left(\frac{\partial V}{\partial\sigma}\right)-\frac{\partial^2 V}{\partial\mu^2}\left(\frac{\partial V}{\partial\sigma}\right)^2\right]\Big/\left(\frac{\partial V}{\partial\mu}\right)^3. \tag{9}$$

While this will hold for low values of σ for all utility functions and distributions, Appendix B shows it will not remain so over the whole of (σ,μ) space for the class of utilities u(x) =u(g(y)) defined in Section 2. The quasiconcavity region depends on both the properties of the transformation to location-scale and the behaviour of the indifference curve of $V(\mu^*,\sigma^*)$ in (σ^*,μ^*) space. Seeking a single formula covering all transformations and utility functions leads to extremely unwieldy algebraic terms. For the case of x log normal x and any u(x) = u(g(y)), concave in y as well as x, non negativity of (9) requires that

$$\frac{\mu(\mu^2+\sigma^2)}{\left[\mu^2+\sigma^2\left(\frac{S^*}{\sigma^*}+2\right)\right]^3}\left\{\left(\frac{S^*}{\sigma^*}+1\right)\left[\mu^2-\sigma^2\left(\frac{S^*}{\sigma^*}+2\right)\right]+\frac{\sigma^2}{\sigma^{*2}}\left(\frac{dS*}{d\sigma*}-\frac{S^*}{\sigma*}\right)\right\}, \qquad (10)$$

with

$$\frac{dS^*}{d\sigma^*}=\frac{\partial S^*}{\partial\sigma^*}+S^*\frac{\partial S^*}{\partial\mu^*},$$

be non-negative. This shows that quasiconcavity certainly holds in the interval

$$\sigma<\mu\left(\frac{S^*}{\sigma^*}+2\right)^{-\frac{1}{2}}, \qquad (11)$$

when the elasticity of the indifference curve in (σ^*,μ^*) space

$$\frac{d\log S^*}{d\log\sigma^*}$$

is equal to unity. When the elasticity is greater than unity the interval expands, since σ can become somewhat larger, and when the elasticity is less than unity the interval contracts correspondingly. But it is difficult to make more clear-cut statements without considering specific utility functions. Similar conditions to (10) can be obtained for other distributions transformable to location-scale form, but in the case of any particular utility function, it is usually much easier to proceed by obtaining S as in the previous Section and then examining the positivity of

$$\frac{dS}{d\sigma}=\frac{\partial S}{\partial\sigma}+\frac{\partial S}{\partial\mu}\frac{d\mu}{d\sigma}=\frac{\partial S}{\partial\sigma}+S\frac{\partial S}{\partial\mu}.$$

This is what Feldstein (1969) did for the case of u(x) = log x with x log normal. Simple differentiation of (5) shows

$$\frac{dS}{d\sigma}=\frac{\mu(\mu^2+\sigma^2)(\mu^2-2\sigma^2)}{(\mu^2+2\sigma^2)^3}$$

so that convexity of the indifference curve requires $\sigma<\mu/\sqrt{2}$. This has already been obtained more tediously in Appendix B and would also follow from (11). As mentioned in the previous section, Elton and Gruber (1974) also examined this case, but unlike Feldstein, did not think the lack of convexity significant[3]. They (p. 487) equated the slope of the frontier $\mu = \mu(\sigma)$ to (5) saying equality could hold for up to two portfolios. If just one exists they see no problem, while if both exist, they choose that with the highest $V(\mu,\sigma)$. If neither exist, they say an end point (if any) of the frontier is optimal. This cannot be quite right as is easily seen by considering a straight line frontier. This could be tangential from below to the convex section ($\sigma < \mu/\sqrt{2}$) of an indifference curve, so a solution exists[4], but eventually cross it somewhere in its concave region ($\sigma > \mu/\sqrt{2}$) and reach higher level indifference curves[5]. The true rationale for assuming optimality obtained by equating the slope of the frontier to S must depend on one of two circumstances. The first is that the frontier could itself be more concave than is the indifference curve in its concave region. Then the tangent to the convex section gives the only contact point and the frontier cannot recross the indifference curve. The second is that the end point of the frontier is reached before it can recross the indifference curve.

Returning to another example of Section 3, the utility function

$$u(x) = A - \lambda x^{-\lambda}$$

had $V(\mu,\sigma)$ given by (6) and S by (7). Differentiation of the latter gives

$$\frac{dS}{d\sigma} = \frac{(\lambda+1)\mu(\mu^2+\sigma^2)[\mu^2-(\lambda+2)\sigma^2]}{[\mu^2+(\lambda+2)\sigma^2]^3}.$$

So again there is an interval ($\sigma < \mu/\sqrt{\lambda+2}$) where the indifference curve is convex and one where it is concave ($\sigma > \mu/\sqrt{\lambda+2}$). Depending on whether λ is negative or positive, the convex interval can be larger or smaller than in the previous case. It is worth noting from (8) that in this convex region the desirable property of decreasing absolute risk aversion holds. Once again, the equivalence of maximising expected utility to the $\boldsymbol{\mu/\sigma}$ procedure of equating the slopes of frontier and indifference curve will depend on the frontier being more concave than the indifference curve in its concave region. In many applied situations that may be plausible enough. For example, in production

[3] Possibly because Feldstein choose to illustrate by the case of of a single risky asset and a riskless one, with the only decision variable being the proportion of wealth assigned to the latter. But then $\boldsymbol{\mu/\sigma}$ equivalence actually holds, because it is a case of the 'location-scale' condition. Several authors criticised Feldstein on this basis including (implicitly) Bierwag (1974), Mayshar (1978) and Meyer (1987). But Elton and Gruber's case was not of this type.

[4] The other solution, where the line is tangential from above to a lower indifference curve could also exist, although it would not be of interest.

[5] The slopes of these indifference curves go to 0 as $\sigma \to \infty$.

economic problems the frontier will often be a marginal cost curve and under the assumption of increasing marginal cost, the slope will not only decrease with σ but could even become negative before the concave region of the indifference curve is reached.

To summarise: if y = h(x) is concave and y follows a location-scale distribution, then for every concave utility function u(x) that is also concave as a function of y, $\boldsymbol{\mu}/\boldsymbol{\sigma}$ analysis is compatible with maximising expected utility in the sense that the maximum will occur on the frontier. But although the indifference curve is initially convex as σ increases, it can become concave at a point determined by the distribution and utility function. Then the optimality of the tangential point of the frontier and the convex phase of the indifference curve depends on the concavity or terminal of the frontier.

5. SLOPES OF INDIFFERENCE CURVES AND ABSOLUTE RISK AVERSION

The slope S of the (μ,σ) indifference curve is the positive quantity

$$S = -\frac{\partial V}{\partial \sigma} / \frac{\partial V}{\partial \mu} .$$

For location-scale distributions, Meyer (1987) discussed the correspondence between the sign of

$$\frac{\partial S^*}{\partial \mu^*} \quad \text{and} \quad \frac{\partial R(y)}{\partial y}$$

where R(y) is the Arrow-Pratt measure of absolute risk aversion $-u''(y)/u'(y)$. Since $R'(y)$ is $-u'''(y)/u'(y) + R^2(y)$, it cannot be negative if $u'''(y)$ is. It is then easy to show that S^* increases with μ^* if R(y) increases with y, is constant if R(y) is, but may decrease with μ^* if R(y) decreases with y[6]. This raises the matter of possible correspondences that might exist in our non location-scale situation. Unsurprisingly, matters are more complicated, but some parallels emerge. As is easily seen

$$R(x) = -\frac{\partial^2 u(x)}{\partial x^2} / \frac{\partial u(x)}{\partial x} = R(y)y'(x) - y''(x)/y'(x)$$

and

[6] Various authors, most recently Eichner and Wagener (2005), have considered the implications of the signs of higher order derivatives of $u(y)$ for further correspondences in the location-scale situation.

$$\frac{\partial R(x)}{\partial x} = R'(x) = R'(y)(y'(x))^2 + R(y)y''(x) - y'''(x)/y'(x) + (y''(x)/y'(x))^2. \quad (12)$$

Even if $R'(y)$ is positive, $R'(x)$ need not be because the second term is negative as the third could also be. If we take y = log x, (12) simplifies to

$$R'(x) = \frac{R'(y)}{x^2} - \frac{1}{x^2}[R(y)+1]. \quad (13)$$

Even if $R(y)$ shows increasing absolute aversion, $R'(x)$ could still be negative. For example, the familiar quadratic utility function

$$u(y) = y - by^2, \quad y < 1/2b,$$

of course displays increasing absolute risk aversion in y, but

$$R(x) = -\frac{\partial^2 u}{\partial x^2} / \frac{\partial u}{\partial x} = \frac{1}{x}\left(1 + \frac{2b}{1 - 2b\log x}\right)$$

and as this is a product of a term decreasing in x with one increasing in x, it need not be increasing overall.

For x log normal (B5) applies

$$\frac{\partial S}{\partial \mu} = \frac{1}{\left[\mu^2 + \sigma^2\left(\frac{S^*}{\sigma^*} + 2\right)\right]^2} \left\{ \begin{array}{c} \left(\frac{S^*}{\sigma^*} + 1\right)\left[\sigma^2\left(\frac{S^*}{\sigma^*} + 2\right) - \mu^2\right] + \frac{\sigma(\mu^2 + 2\sigma^2)}{\sigma^*}\frac{\partial S^*}{\partial \mu^*} \\ -\frac{\sigma^2}{\sigma^{*2}}\left(\frac{\partial S*}{\partial \sigma *} - \frac{S^*}{\sigma *}\right) \end{array} \right\}. \quad (14)$$

In (13) constant absolute risk aversion in y implies $R'(y) = 0$ and decreasing risk aversion in in x. In (14), if S^* has the constant relative risk aversion form $\lambda\sigma^*$, then the second and third terms in chain brackets are zero and decreasing absolute risk aversion holds in (σ, μ) space provided (11) holds. But that was the condition for validity in the sense of a convex indifference curve. So with this S^* there is only decreasing absolute risk aversion and the signs of (13) and (14) agree.

If $R'(y)$ is negative in (13), $R'(x)$ certainly is. $R'(y)$ negative requires $u'''(y)$ positive, although it is not guaranteed by it. $R'(y)$ is certainly positive if $u'''(y)$ is negative , but as already mentioned, this need not imply $R'(x)$ positive. Since

$$\frac{\partial S^*}{\partial \mu *} = -\left(\frac{\partial^2 V}{\partial \mu^* \partial \sigma^*} + S^* \frac{\partial^2 V}{\partial \mu^{*2}} \right) / \frac{\partial V}{\partial \mu *}$$

and

$$\frac{\partial S^*}{\partial \sigma *} = -\left(\frac{\partial^2 V}{\partial \sigma^{*2}} + S^* \frac{\partial^2 V}{\partial \mu^* \partial \sigma^*} \right) / \frac{\partial V}{\partial \mu *}$$

and positive $u'''(y)$ means the second derivative with respect to μ^* and σ^* is positive, the derivative of S^* with respect to μ^* can then be negative and the second term in (14) is negative. If $u'''(y)$ is negative, both derivatives of S^* with respect to μ^* and σ^* are positive and the second term is positive. But (14) can still be negative through the first term if σ is sufficiently smaller than

$$\mu\left(\frac{S^*}{\sigma^*} + 2\right)^{-\frac{1}{2}}.$$

So the correspondences between the properties of the Arrow-Pratt measure for risk aversion of u(x) and the risk aversion properties of its expectation, $V(\mu,\sigma)$, when that exists for a location-scale distribution, do not carry over comprehensively to our class of utility functions and distributions transformable to location-scale form, although there are at least some parallels.

6. SUMMARY AND CONCLUDING REMARKS

We have shown that if x follows a non location-scale two parameter distribution, but y = h(x) does follow a location-scale distribution, where h(x) is concave, conclusions follow about compatibility of the expected utility and $\boldsymbol{\mu}/\boldsymbol{\sigma}$ approaches for all utility for which $u(x) = u(h^{-1}(y))$ is concave in y as well as x. As was shown in Section 3, $E(u(x)) = V(\mu,\sigma)$ is increasing in μ and decreasing in σ and then maximising expected utility is compatible with $\boldsymbol{\mu}/\boldsymbol{\sigma}$ analysis in the sense of the maximum occurring on the efficient frontier. Full agreement of the expected utility approach with $\boldsymbol{\mu}/\boldsymbol{\sigma}$ analysis would require that the optimum occur at the tangential point of the indifference curve and the efficient frontier. However, as was described in Section 4, the indifference curve, although initially convex, can eventually become concave. While there is always a tangential point of the convex portion of the indifference curve and the frontier, the optimality of this point requires either that the degree of concavity of the efficient frontier exceeds that of the indifference curve, or if it does not, that the frontier terminates before it can recross that curve.

Requiring that $u(x) = u(h^{-1}(y))$ is concave in y as well as x does restrict the class of utility functions somewhat, as was discussed in Section 2. But, as mentioned in the Introduction, strong restrictions are implied by the very existence of $E(u(x)) = V(\mu, \sigma)$ in the location-scale case. The permissible class of utility functions is important, containing familiar utility functions such as log x that do not possess expectations under normality of x, but do under lognormality or Pareto distributions, which are also more appropriate for modelling income or wealth than location-scale distributions. However, we do not claim that there cannot be any utility functions, other than those of this class , for which expected utility and μ / σ analysis are compatible.

It is worth mentioning that there may even be a generalisation of Meyer's (1987) 'location scale condition'. An originally multiparameter distribution, employed in circumstances where all parameters except two are held constant, effectively becomes a two parameter distribution. This distribution will not belong to the location-scale family if the variable parameters do not correspond to location and scale. But perhaps the distribution could be of this paper's type where a concave transformation y = h(x) produces a location-scale distribution for y. Whether there really are applied problems of this nature will not be pursued here.

Finally, the investigation in section 5 did not reproduce the relatively clear-cut correspondences between the properties of the Arrow-Pratt measure for u(x) and the properties of $V(\mu, \sigma)$ that hold for location-scale distributions. While there are some parallel properties, the situations are more complex. However, we feel the matters of key importance were those covered in Sections 3 and 4 relating to the monotonicity and quasiconcavity conditions.

APPENDIX A

EQUIVALENCE OF E(U) TO μ / σ ANALYSIS FOR LOCATION-SCALE DISTRIBUTIONS

Let the distribution be

$$\frac{1}{\alpha} f\left(\frac{x-\beta}{\alpha}\right) , \qquad \beta - c_l \alpha \le x \le \beta + c_u \alpha ,$$

with β and α positive and where c_l or c_u are constants independent of β and α , one or both of which can be infinite. We will denote the mean and standard deviation of the distribution by $\mu*$ and $\sigma*$. So a normal distribution has both c_l and c_u infinite, $\beta = \mu*$ and $\alpha = \sigma*$. A Gumbel distribution has c_l and c_u infinite, but $\beta = \mu * -\gamma\sigma * \sqrt{6}/\pi,$ where γ $(\approx .5772)$ is Euler's constant, and $\alpha = \sigma * \sqrt{6}/\pi.$ The two parameter exponential has c_l zero and c_u infinite, $\beta = \mu * -\alpha$ and $\alpha = \sigma.*$ The uniform distribution has both c_l and c_u equal to $\sqrt{3}$, $\beta = \mu *$ and $\alpha = \sigma *$. Generally,

$$\mu = \frac{1}{\alpha}\int_{\beta - c_l\alpha}^{\beta + c_u\alpha} xf\left(\frac{x-\beta}{\alpha}\right)dx \ .$$

Putting

$$w = \frac{x-\beta}{\alpha}$$

it becomes

$$\mu^* = \beta + \alpha\int_{-c_l}^{c_u} wf(w)dw = \beta + \alpha k_1 \ ,$$

where k_1 is a constant. Similarly $\sigma^* = \alpha\, k_2$, where k_2 is another constant. Notice that since

$E(w) = (\mu^* - \beta)/\alpha = k_1$

$$\int_{-c_l}^{c_u}(w-k_1)f(w)dw = 0. \qquad \text{(A1)}$$

The expectation of u(x) is

$$V = \frac{1}{\alpha}\int_{\beta - c_l\alpha}^{\beta + c_u\alpha} u(x)f\left(\frac{x-\beta}{\alpha}\right)dx$$

and is presumed to exist. This places constraints on choice of u(x), for example, ruling out log x[7].

Substituting w for x

$$V = \int_{-c_l}^{c_u} u(\beta + \alpha\, w)f(w)dw = \int_{-c_l}^{c_u} u(\mu * + \sigma * (\frac{w-k_1}{k_2}))f(w)dw\,.$$

Then, for convenience taking the limits of integration as understood,

$$\frac{\partial V}{\partial \mu *} = \int u'(...)f(w)dw,$$

which is positive since u(x) is increasing in x.

[7] Clearly a device like u(x) = 0 for $x \le 1$ and u(x) = log x, for x > 1, gives a function with finite expectation, but then the requirements for the proofs that follow do not hold.

$$\frac{\partial V}{\partial \sigma *} = \frac{1}{k_2} \int u'(...)(w - k_1) f(w) dw$$

and since $u'(x)$ is decreasing in x, negative values of $w - k_1$ are being multiplied by larger values than are positive values. So, in view of (A1), the derivative of V with respect to $\sigma *$ is negative. Then the slope of an indifference curve

$$S* = \frac{d\mu *}{d\sigma *} = -\frac{\partial V}{\partial \sigma *} / \frac{\partial V}{\partial \mu *}$$

is positive.

$$\frac{\partial^2 V}{\partial \mu *^2} = \int u''(...) f(w) dw$$

is negative since $u''(x)$ is negative and

$$\frac{\partial^2 V}{\partial \sigma *^2} = \frac{1}{k_2^2} \int u''(...)(w - k_1)^2 f(w) dw$$

is negative for the same reason. The sign of

$$\frac{\partial^2 V}{\partial \mu * \partial \sigma *} = \frac{1}{k_2} \int u''(...)(w - k_1) f(w) dw$$

is unclear. It is positive if $u'''(x)$ is positive and negative if $u'''(x)$ is negative, but in either situation the Cauchy-Schwarz inequality implies

$$\frac{\partial^2 V}{\partial \mu *^2} \frac{\partial^2 V}{\partial \sigma *^2} - \left(\frac{\partial^2 V}{\partial \mu * \partial \sigma *} \right)^2$$

is non-negative. So V is a concave function of μ^* and σ^*, indifference curves are convex, and the equivalence of $\boldsymbol{\mu} / \boldsymbol{\sigma}$ analysis to the expected utility approach is evident. While concavity is not essential for convexity of an indifference curve and that quasiconcavity will suffice, concavity implies quasiconcavity. Also, the main focus of this paper is to utilise these results to extend the equivalence to distributions that are not location-scale, but are transformable to that family.

APPENDIX B

EQUIVALENCE FOR TRANSFORMABLE DISTRIBUTIONS AND RESTRICTED U(X)

Proving Monotonicity

To show $V(\mu, \sigma)$ satisfies monotonicity.we have to prove that the terms of (1)

$$\frac{\partial\mu^*}{\partial\mu} \quad \text{and} \quad \frac{\partial\sigma^*}{\partial\sigma}$$

are positive and that the terms of (2)

$$\frac{\partial\mu^*}{\partial\sigma} \quad \text{and} \quad \frac{\partial\sigma^*}{\partial\mu}$$

are negative. The distribution of x is not location-scale, but that of y=h(x) is. Now h(x) is assumed increasing and concave. The inverse transformation x=g(y) must be such that x=g(h(x)). Then

$$1 = \frac{\partial g}{\partial y}\frac{\partial h}{\partial x}.$$

So g(y) is increasing in y. Also

$$0 = \frac{\partial^2 g}{\partial y^2}\left(\frac{\partial h}{\partial x}\right)^2 + \frac{\partial g}{\partial y}\frac{\partial^2 h}{\partial x^2}$$

and since h(x) is concave, its second derivative is negative and so the second derivative of g(y) is positive. That is, g(y) is increasing and convex. Clearly

$$\mu = \frac{1}{\alpha}\int_{\beta - c_l\alpha}^{\beta + c_u\alpha} g(y) f\left(\frac{y-\beta}{\alpha}\right) dy$$

or, with the same substitution as in Appendix A,

$$\mu = \int_{-c_l}^{c_u} g(\mu^* + \sigma^*(\frac{w-k_1}{k_2}))f(w)dw.$$

Taking the limits of integration as understood and, for convenience, omitting arguments of g

$$\frac{\partial\mu}{\partial\mu^*} = \int g' f(w)dw,$$

which is positive since g(y) is increasing in y.

$$\frac{\partial\mu}{\partial\sigma^*} = \frac{1}{k_2}\int g'(w-k_1)f(w)dw$$

and since $g'(y)$ is increasing in y, negative values of $w-k_1$ are being multiplied by smaller values than are positive values. So, in view of (A1), the derivative with respect to σ^* is positive. Also

$$\sigma^2 = \frac{1}{\alpha}\int_{\beta - c_l\alpha}^{\beta + c_u\alpha} \{g(y) - \mu\}^2 f\left(\frac{y-\beta}{\alpha}\right)dy$$

or

$$\sigma^2 = \int_{-c_l}^{c_u} \left\{ g(\mu^* + \sigma^* (\frac{w-k_1}{k_2})) - \mu \right\}^2 f(w)dw.$$

Then

$$\frac{\partial \sigma^2}{\partial \mu^*} = 2\sigma \frac{\partial \sigma}{\partial \mu^*} = 2\int (g - \mu) g' f(w)dw.$$

Noting that, by definition,

$$0 = \int (g - \mu) f(w)dw \qquad \text{(B1)}$$

and remembering g' is increasing, negative values of $g - \mu$ are being multiplied by smaller values than are positive values and so (B1) implies the derivative of σ with respect to μ^* is positive. Again,

$$\frac{\partial \sigma^2}{\partial \sigma^*} = 2\sigma \frac{\partial \sigma}{\partial \sigma^*} = \frac{2}{k_2}\int (g - \mu) g'(w - k_1) f(w)dw.$$

When $w < k_1$ and $g < \mu$ the sign inside the integral is positive. When $w > k_1$ and $g < \mu$ it is positive and when $w > k_1$ and $g > \mu$ the sign inside the integral is also positive[8]. Then the fact that g' is increasing ensures the derivative of σ with respect to σ^* is positive. So the terms in the matrix

$$M = \begin{bmatrix} \frac{\partial \mu}{\partial \mu^*} & \frac{\partial \mu}{\partial \sigma^*} \\ \frac{\partial \sigma}{\partial \mu^*} & \frac{\partial \sigma}{\partial \sigma^*} \end{bmatrix}$$

are all positive. Now

$$\begin{bmatrix} d\mu \\ d\sigma \end{bmatrix} = \begin{bmatrix} \frac{\partial \mu}{\partial \mu^*} & \frac{\partial \mu}{\partial \sigma^*} \\ \frac{\partial \sigma}{\partial \mu^*} & \frac{\partial \sigma}{\partial \sigma^*} \end{bmatrix} \begin{bmatrix} d\mu^* \\ d\sigma^* \end{bmatrix}$$

and of course

---

[8] Since μ is the mean of a positive increasing convex function of y, it will be much larger than k_1, which is zero for a normal or uniform, .5572 for a Gumbel and 1 for a two parameter exponential.

$$\begin{bmatrix} d\mu^* \\ \\ d\sigma^* \end{bmatrix} = \begin{bmatrix} \frac{\partial \mu^*}{\partial \mu} & \frac{\partial \mu^*}{\partial \sigma} \\ \frac{\partial \sigma^*}{\partial \mu} & \frac{\partial \sigma^*}{\partial \sigma} \end{bmatrix} \begin{bmatrix} d\mu \\ \\ d\sigma \end{bmatrix}.$$

So

$$\begin{bmatrix} \frac{\partial \mu^*}{\partial \mu} & \frac{\partial \mu^*}{\partial \sigma} \\ \frac{\partial \sigma^*}{\partial \mu} & \frac{\partial \sigma^*}{\partial \sigma} \end{bmatrix} = M^{-1}$$

and evidently the diagonal elements are positive and the off diagonal elements negative. So

$$\frac{\partial \mu^*}{\partial \mu} \quad \text{and} \quad \frac{\partial \sigma^*}{\partial \sigma}$$

are positive and

$$\frac{\partial \mu^*}{\partial \sigma} \quad \text{and} \quad \frac{\partial \sigma^*}{\partial \mu}$$

are negative and $V(\mu,\sigma)$ satisfies monotonicity.

Investigating Quasiconcavity

We need $V(\mu,\sigma)$ quasiconcave in μ and σ to ensure the indifference curve is convex. The condition for this is that

$$\left[-\frac{\partial^2 V}{\partial \sigma^2}\left(\frac{\partial V}{\partial \mu}\right)^2 + 2\frac{\partial^2 V}{\partial \mu \partial \sigma}\left(\frac{\partial V}{\partial \mu}\right)\left(\frac{\partial V}{\partial \sigma}\right) - \frac{\partial^2 V}{\partial \mu^2}\left(\frac{\partial V}{\partial \sigma}\right)^2 \right] \Big/ \left(\frac{\partial V}{\partial \mu}\right)^3 \qquad \text{(B2)}$$

be positive. Now

$$\frac{\partial^2 V}{\partial \mu^2} = \frac{\partial}{\partial \mu^*}\left(\frac{\partial V}{\partial \mu}\right)\frac{\partial \mu^*}{\partial \mu} + \frac{\partial}{\partial \sigma^*}\left(\frac{\partial V}{\partial \mu}\right)\frac{\partial \sigma^*}{\partial \mu}$$

or, working out terms

$$\frac{\partial^2 V}{\partial \mu^{*2}}\left(\frac{\partial \mu^*}{\partial \mu}\right)^2 + 2\frac{\partial^2 V}{\partial \mu^* \partial \sigma^*}\frac{\partial \mu^*}{\partial \mu}\frac{\partial \sigma^*}{\partial \mu} + \frac{\partial^2 V}{\partial \sigma^{*2}}\left(\frac{\partial \sigma^*}{\partial \mu}\right)^2 + \frac{\partial V}{\partial \mu^*}\frac{\partial^2 \mu^*}{\partial \mu^2} + \frac{\partial V}{\partial \sigma^*}\frac{\partial^2 \sigma^*}{\partial \mu^2}.$$

Similarly

$$\frac{\partial^2 V}{\partial \sigma^2}$$

is

$$\frac{\partial^2 V}{\partial \mu^{*2}}\left(\frac{\partial \mu^*}{\partial \sigma}\right)^2 + 2\frac{\partial^2 V}{\partial \mu^* \partial \sigma^*}\frac{\partial \mu^*}{\partial \sigma}\frac{\partial \sigma^*}{\partial \sigma} + \frac{\partial^2 V}{\partial \sigma^{*2}}\left(\frac{\partial \sigma^*}{\partial \sigma}\right)^2 + \frac{\partial V}{\partial \mu^*}\frac{\partial^2 \mu^*}{\partial \sigma^2} + \frac{\partial V}{\partial \sigma^*}\frac{\partial^2 \sigma^*}{\partial \sigma^2}$$

and

$$\frac{\partial^2 V}{\partial \mu \partial \sigma}$$

is

$$\frac{\partial^2 V}{\partial \mu^{*2}}\left(\frac{\partial \mu^*}{\partial \mu}\right)\left(\frac{\partial \mu^*}{\partial \sigma}\right) + 2\frac{\partial^2 V}{\partial \mu^* \partial \sigma^*}\left(\frac{\partial \mu^*}{\partial \sigma}\frac{\partial \sigma^*}{\partial \mu} + \frac{\partial \mu^*}{\partial \mu}\frac{\partial \sigma^*}{\partial \sigma}\right) + \frac{\partial^2 V}{\partial \sigma^{*2}}\left(\frac{\partial \sigma^*}{\partial \mu}\right)\left(\frac{\partial \sigma^*}{\partial \sigma}\right)$$
$$+ \frac{\partial V}{\partial \mu^*}\frac{\partial^2 \mu^*}{\partial \mu \partial \sigma} + \frac{\partial V}{\partial \sigma^*}\frac{\partial^2 \sigma^*}{\partial \mu \partial \sigma}.$$

Substituting these into the numerator of (B2) many terms cancel and the remaining terms give

$$\left[-\frac{\partial^2 V}{\partial \sigma^{*2}}\left(\frac{\partial V}{\partial \mu^*}\right)^2 + 2\frac{\partial^2 V}{\partial \mu^* \partial \sigma^*}\left(\frac{\partial V}{\partial \mu^*}\right)\left(\frac{\partial V}{\partial \sigma^*}\right) - \frac{\partial^2 V}{\partial \mu^{*2}}\left(\frac{\partial V}{\partial \sigma^*}\right)^2\right]\left(\frac{\partial \mu^*}{\partial \mu}\frac{\partial \sigma^*}{\partial \sigma} - \frac{\partial \mu^*}{\partial \sigma}\frac{\partial \sigma^*}{\partial \mu}\right)^2$$

$$+\left[\frac{\partial V}{\partial \mu^*}\frac{\partial^2 \mu^*}{\partial \mu^2} + \frac{\partial V}{\partial \sigma^*}\frac{\partial^2 \sigma^*}{\partial \mu^2}\right]\left(\frac{\partial V}{\partial \mu^*}\frac{\partial \mu^*}{\partial \sigma} + \frac{\partial V}{\partial \sigma^*}\frac{\partial \sigma^*}{\partial \sigma}\right)^2$$

$$+\left[\frac{\partial V}{\partial \mu^*}\frac{\partial^2 \mu^*}{\partial \sigma^2} + \frac{\partial V}{\partial \sigma^*}\frac{\partial^2 \sigma^*}{\partial \sigma^2}\right]\left(\frac{\partial V}{\partial \mu^*}\frac{\partial \mu^*}{\partial \mu} + \frac{\partial V}{\partial \sigma^*}\frac{\partial \sigma^*}{\partial \mu}\right)^2$$

$$+2\left[\frac{\partial V}{\partial \mu^*}\frac{\partial^2 \mu^*}{\partial \mu \partial \sigma} + \frac{\partial V}{\partial \sigma^*}\frac{\partial^2 \sigma^*}{\partial \mu \partial \sigma}\right]\left(\frac{\partial V}{\partial \mu^*}\frac{\partial \mu^*}{\partial \mu} + \frac{\partial V}{\partial \sigma^*}\frac{\partial \sigma^*}{\partial \mu}\right)\left(\frac{\partial V}{\partial \mu^*}\frac{\partial \mu^*}{\partial \sigma} + \frac{\partial V}{\partial \sigma^*}\frac{\partial \sigma^*}{\partial \sigma}\right).$$

The first term of this expression is the product of a squared term and the condition that $V(\mu^*, \sigma^*)$ is quasiconcave. But that condition must be true because Appendix A proved $V(\mu^*, \sigma^*)$ concave in μ^* and σ^*. So the first term is non-negative. However, it is unclear that the sum of all terms is non-negative for all (μ , σ) space. In the extreme situation of u(x) = h(x) = y, $V(\mu^*, \sigma^*) = \mu^*$ and all second derivatives of V with respect to μ^* and σ^* are zero and the first term of the expression is zero. As might be expected the remaining terms reduce to

$$\left[-\frac{\partial^2 \mu^*}{\partial \sigma^2}\left(\frac{\partial \mu^*}{\partial \mu}\right)^2 + 2\frac{\partial^2 \mu^*}{\partial \mu \partial \sigma}\left(\frac{\partial \mu^*}{\partial \mu}\right)\left(\frac{\partial \mu^*}{\partial \sigma}\right) - \frac{\partial^2 \mu^*}{\partial \mu^2}\left(\frac{\partial \mu^*}{\partial \sigma}\right)^2\right]. \qquad \text{(B3)}$$

So we require μ^*, the expectation of the transformation, to be quasiconcave in μ and

σ. Now if we take x log normal and y = h(x) = log x, so that y is normal, (B3) can be shown by rather tedious evaluation of terms to be

$$\frac{(\mu^2 - 2\sigma^2)}{\mu^2(\mu^2 + \sigma^2)^2}$$

and this can be negative unless $\mu > \sqrt{2}\sigma$. Remaining with x lognormal, but any u(x) that is a concave function of y = log x, very laborious manipulation enables (B2) to be written

$$\frac{\mu(\mu^2 + \sigma^2)}{\left[\mu^2 + \sigma^2\left(\frac{S^*}{\sigma^*} + 2\right)\right]^3} \left\{ \left(\frac{S^*}{\sigma^*} + 1\right)\left[\mu^2 - \sigma^2\left(\frac{S^*}{\sigma^*} + 2\right)\right] + \frac{\sigma^2}{\sigma^{*2}}\left(\frac{dS^*}{d\sigma^*} - \frac{S^*}{\sigma^*}\right) \right\}, \qquad \text{(B4)}$$

where

$$\frac{dS^*}{d\sigma^*} = \frac{\partial S^*}{\partial \sigma^*} + S^* \frac{\partial S^*}{\partial \mu^*}.$$

Another formula obtained during the derivation of (B4) is

$$\frac{\partial S}{\partial \mu} = \frac{1}{\left[\mu^2 + \sigma^2\left(\frac{S^*}{\sigma^*} + 2\right)\right]^2} \left\{ \begin{array}{c} \left(\frac{S^*}{\sigma^*} + 1\right)\left[\sigma^2\left(\frac{S^*}{\sigma^*} + 2\right) - \mu^2\right] + \frac{\sigma(\mu^2 + 2\sigma^2)}{\sigma^*}\frac{\partial S^*}{\partial \mu^*} \\ -\frac{\sigma^2}{\sigma^{*2}}\left(\frac{\partial S^*}{\partial \sigma^*} - \frac{S^*}{\sigma^*}\right) \end{array} \right\}, \qquad \text{(B5)}$$

which is employed in Section 5.

Returning to (B4), the second term within chain brackets in (B4) could be positive, negative or zero depending on whether the elasticity of the slope along the indifference curve associated with $V(\mu^*, \sigma^*)$ is greater than, equal to, or less than, unity. The first term could be negative too, because $V(\mu^*, \sigma^*) = \mu^*$ implies $S^* = 0$ and all terms within the chain brackets vanish except $\mu^2 - 2\sigma^2$, implying, as before, $\mu > \sqrt{2}\sigma$ for positivity. So the region of quasiconcavity in (σ, μ) space depends on both the quasiconcavity of the transformation expectation and the elasticity of the indifference curve in (σ^*, μ^*).

Corresponding conditions to (B4) can be obtained for other distributions transformable to location-scale form, but attempting to obtain a single formula applicable to all the relevant distributions and utility functions seems to result in almost intractable algebraic expressions. In any event, the quasiconcavity region for particular cases is more easily obtained by direct examination of the convexity of the indifference curve in (σ, μ) space via the positivity of

$$\frac{dS}{d\sigma} = \frac{\partial S}{\partial \sigma} + S\frac{\partial S}{\partial \mu}$$

without explicit consideration of $V(\mu^*, \sigma^*)$ or its indifference curve. The examples in section 4 exemplify this.

It may be worth remembering that for sufficiently small σ, the expectation of any utility function can be written

$$E[u(x)] = V(\mu, \sigma) = u(\mu) + \frac{1}{2}\sigma^2 u''(\mu),$$

so that

$$S = -\sigma\frac{u''(\mu)}{u'(\mu)} = \sigma R_A(\mu),$$

where R_A is the Arrow-Pratt coefficient of absolute risk aversion. Then

$$\frac{dS}{d\sigma} = \frac{\partial S}{\partial \sigma} + S\frac{\partial S}{\partial \mu} = R_A(\mu)\left[1 + \sigma^2\frac{\partial R_A(\mu)}{\partial \mu}\right]$$

and this is positive, even if decreasing absolute risk aversion holds, if σ is small enough. So every indifference curve commences with a convex region, the extent of which depends on the properties of the distribution and the utility function.

REFERENCES

Aitchison, J. and J. A. C. Brown (1957) *The Log Normal Distribution*, Cambridge: Cambridge University Press.

Arnold, B. C. (1983) *Pareto Distributions*, Fairland, Maryland: International Co-operative Publishing House.

Bandourian, R., J. B. McDonald and R. S. Jurley (2002) "A comparison of parametric models of income distribution across countries and over time", *Estadistica,* 55, 135-152.

Bierwag, G. O. (1974), "The Rationale of Mean Standard Deviation Analysis: Comment", *American Economic Review*, 64, 431-433.

Eichner, T. and A. Wagener (2005) "Measures of risk attitudes: correspondence between mean-variance and expected utility approach", *Decisions in Economics and Finance*, 28, 53-56.

Elton, E. J. and M. J. Gruber (1974) "On the maximisation of the geometric mean with log normal return distribution", *Management Scien*ce, 21, 483-487.

Elton, E. J. and M. J. Gruber (1995) *Market Portfolio Theory and Investment Analysis* 5th Ed., New York: Wiley.

Feldstein, M. S.: (1969) "Mean-Variance Analysis in the Theory of the Firm under Uncertainty", *Review of Economic Studies*, 36, 5-12.

Mayshar, J. (1978), "A Note on Feldstein's Criticisms of Mean-Variance Analysis", *Review of Economic Studies*, 45, 197-199.

Meyer, J. (1987): "Two-Moment Decision Models and Expected Utility Maximisation", *American*

Economic Review, 77, 421-430.

Sinn, H.-W. (1989) *Economic Decisions under Uncertainty* 2nd Ed., New York: Springer-Verlag (Ist Ed. 1983, Amsterdam: North-Holland).