Revised September 2011

Over the odds: Information-based rescaling in subjective probability and prospect theory

Roger J. Bowden[1]

Abstract

The psychological basis of subjective probability is often identified with the odds at which agents will gamble. The same framework can be used in cumulative prospect theory to derive 'weighting functions' for outcomes that differ from the natural distribution in assigning additional weight to exceptionally good or bad outcomes. Rescaling is itself a form of measure change analogous to that used in derivatives pricing. The measure change is based on locational entropy, which cumulates the tail behaviour of the log odds function. Implications for risk management are considered.

Key words: Weighting function; subjective probability; measure change; locational entropy; log odds ratio; Friedman-Savage utility.

JEL classifications: D81, G11, C02.

MSC Classifications: 60E05, 94A17.

---

[1]Kiwicap Research Ltd, Wellington, New Zealand, and Institute of Finance, Faculty of Mathematics and Economics, Ulm University, Germany. Emails: roger.bowden@kiwicap.co.nz; roger.bowden@uni-ulm.de . The author is grateful to Matthias Boehm, Paul Koerbitz and participants at the Ulm Finance Workshop, Heiligkreuztal, July 2011.

1. Introduction

The distinction between objective and subjective probability has a venerable history in economics and statistics, originating with influential authors as Ramsay (1926), de Finetti (1937) and Savage (1954). While the epistemological foundations might differ, the general supposition was that subjectivist probability was nevertheless internally consistent, following the axioms of Von Neumann-Morgenstern (1947); Pfanzagl (1967). In more recent times, Tversky and Kahneman (1974, 1979) suggested that agents do not necessarily think and act in terms of the classic rules of probability. Their principal postulates may be summarised as: (a) framing is important (the way choices are presented); (b) people overestimate probabilities of rare events; and (c) in thinking about compound gambles, people do not recombine probabilities correctly. In particular, cumulative prospect theory (Tversky & Kahneman (1992)) could be applied to investment activity, in terms of which investors would systematically overweight one or both of the tails, relative to the true or natural distribution. Prelec (1998) provides an axiomatic treatment of the choice theory involved. Subsequently, the approach was utilised to explain some well known anomalies in asset pricing, such as the equity premium puzzle (Benartzi and Thaler (1995)), size and value anomalies (Giorgi & Hens (2006)), or the momentum effect (Grinblatt & Han (2005), Menkhoff and Schmeling (2006)).

Cumulative prospect theory substitutes a 'weighting function' of wealth or return outcomes for the original or natural probability distribution. The result, as it applies to investment, is analogous to a subjective probability measure, although the resulting decision rules do necessarily adhere to those that might be derived from the Von Neumann-Morgenstern continuity and independence axioms. Likewise, the classical utility function is replaced by a 'value function', resulting in a decision criterion analogous to expected utility. Nevertheless, prospect theory shares with the more traditional subjectivist approach a basis that even where objective or frequentist probabilities might exist, economic agents map such probabilities into their own more subjective measure. It is the operational specification of this correspondence that is studied in the present paper. In this aspect, there is a margin of choice for the investigator as to the transformation or rescaling that maps between natural probability and weighting function outcomes. Tversky and Kahneman (1992b) used a power law such that the rescaling is greater at either tail of the natural distribution function (see section 2 below). In the last analysis, any such choice has to be tested against results (e.g. Boehm 2011), or based on a revelation experiment or survey.

However, the choice of transformation to put to the test might itself be guided by psychological heuristics or biases. Subjectivist authors, notably de Finetti (1937, 1970), drew on the widespread understanding and use of fractional odds in betting, which continues to this day. The odds at which a subject is willing to bet can be taken as a proxy for the probability that he or she has assigned to the outcome. Taking the log of the odds ratio enables the subject to rank events in an intuitively appealing way. However, at any given support point, two distributions might have the same tail probability but could differ in their tail length and therefore the prospect of large gains or losses. It seems reasonable to suppose that investors or gamblers are influenced by the odds over the entire tail area, rather than just the odds at any particular point. If so, then a behavioural rescaling function can be based on the average log odds over the tail area. This is equivalent to rescaling using locational entropy (Bowden 2010, 2011) which captures the changes in uncertainty as one moves along the horizontal axis. Hence there is a link to the way that investors might react to different degrees of information, as it impacts on the tail length of the original or natural density.

Comparison of alternative rescaling candidates is facilitated by adopting the formal framework of changes in measure based on Radon-Nikodym derivatives. This will be familiar from the theory of derivatives pricing in complete markets, where it transforms to a risk neutral measure. But the same framework can be applied to the quite different context of cumulative prospect theory, involving substitution of a weighting function for the natural measure. It provides a way to organise and evaluate alternative candidates for rescaling functions. The rescaling framework can also be used in the context of more traditional subjective probability, with much the same shades of meaning. In the development of the present paper, the term ‘subjective probability’ can refer to either framework. Likewise, the ‘utility function’ can refer to either the value function of prospect theory or the more traditional cardinal preference function of the expected utility framework.

The scheme of exposition is as follows. Section 2 develops the general rescaling framework for subjective probability, drawing on the analogy with conversion to risk neutral measure in derivatives pricing. The potential problem of observational equivalence, or utility equivalence, is pointed out; in particular, as providing a possible alternative basis for Friedman-Savage (1948) type utility functions. Possible extensions are briefly explored in the same framework. Section 3 utilises the rescaling framework to develop and motivate the information based approach to the choice of scaling functions. It is motivated in terms of the conditional expected log odds function as it applies to the natural distribution's tails. Section 4 considers the issue of risk management if it is known that a predisposition does exist to overweight tail probabilities.

2. Probability rescaling as a measure change

In what follows, a random variable X (e.g. investment returns or wealth states) maps from an underlying event space Ω into the real line, or some subset such as $\Re_+$; so $X(\omega) = x \in \Re$. There is a natural measure under which the random variable induces a real valued distribution function *F(x)*. Investors may, or may not, be aware of or able to assess the true *F*. But in its place, they act in terms of a subjective distribution or weighting function *W*. In what follows it is assumed that subjective probabilities (respectively weighting functions) can be regarded as transformed or rescaled versions of an original or natural measure. The task is to specify procedures to characterise the way in which this might be done, encompassing conjectured natural biases or behavioural regularities.

The process of changing the probability measure has become familiar from derivatives pricing in complete markets[2]. In that context, if F is the original measure and W the equivalent martingale (risk neutral) measure, the no-arbitrage value of an end of period payoff *a(x),* discounted where necessary, is given by

$$v = E_w[a(x)] = E_f[\xi_w(x)a(x)].$$

Here the function $\xi_w(x)$ reflects an adjustment of the natural measure for risk. It can be identified with dW/dF, the Radon Nikodym derivative of one measure with respect to the other, evaluated at point x. If the densities exist, then

$$w(x) = \xi_w(x)f(x),$$

with the requirements $\xi_w(x) \geq 0$; $E_f[\xi_w(x)] = 1$ motivating rescaling as a description of the process. The objective of the present section is to show that a similar idea can be used as a framework for subjective probability rescaling.

2.1 Subjective probability as rescaling

The same mathematical treatment can be applied to subjective probabilities or weighting functions. However, the reference in this context is to psychological predisposition, rather

[2] The original reference is Harrison and Kreps (1979). For general treatments of measure changes, see Billingsley (1995), and in Finance, Duffie (2001).

than to risk neutrality under complete markets. Investors mentally transform the natural probabilities into subjective probabilities (or weighting functions), or act as though they do. Thus it might be known that returns in general have a natural distribution *F*. But investors do not act as though this is their decision basis; instead they think and act in terms of a subjective distribution function *W* (or the 'cumulative weighting function' in prospect theory). In line with the prospect theory findings as to rare events, *W* might have longer tails than *F*, to a degree that could differ as between the upper and lower tails. Thus at the upper end, an investment in a dotcom IPO or an oil explorer might turn into a life changing outcome for the lucky investor. At the lower end, a proposed portfolio could decimate the investor's carefully accumulated retirement capital. In each case the true probability is small, but the investor reweights it to become larger, and acts accordingly.

The basic version of the measure transformation can be represented as $W = \psi(F)$. The function $\psi(F)$ is specified as non decreasing (the monotonicity assumption) and is normalised so that $W(0) = 0$, $W(1) = 1$. The derivative, if it exists, is written as $\xi_w^0(F) = \dfrac{d\psi}{dF}$ so that $dW = \xi_w^0(F)dF$. Correspondingly, let

(1) $\xi_w(x) = \xi_w^0(F(x)) = \Psi'(F(x))$.

Then $dW(x) = \xi_w(x)dF(x)$, and $w(x) = \xi_w(x)f(x)$ if the densities exist. The specifications imposed on ψ imply that the function ξ is nonnegative, and $E_f[\xi_w(x)] = 1$. Thus $\xi_w(x)$ can be interpreted as a Radon Nikodym measure for a change from the natural to the subjective distribution (weighting function) of the investment outcome *x*. The general import is that the function rescales to a greater degree at one, or possibly both, tails of the original, making compensating re-weightings elsewhere. Thus if investors act as though events of rare good fortune have an inflated probability, they will act as though $1-W(F) > 1-F$ for outcomes *x* such that $F(x)$ is near unity; equivalently, $\xi_w(x) = \xi_w^0(F(x)) >> 1$ in this zone.

Operationally, one might start with the suggestion that a given nonnegative function $s(F)$ could serve as a possible rescaling function. Normalising as $\xi_w^0(F) = s(F) / \int_0^1 s(F)dF$ and setting $\xi_w(x) = \xi_w^0(F(x))$ will assure that $E_f[\xi_w(x)] = 1$, providing the integral $\int_0^1 s(F)dF$ exists.

Remark 1

The usual assumption in prospect theory is that the reweighting represents distortions that apply to *F*, e.g. in the tail areas of special interest, and that the value of *F* is a sufficient statistic for this purpose. However, introspection might suggest that this is not entirely true. There might be a point where the utility function (or 'value function', in prospect theory terminology) adopts sharper curvature upward, as in a life changing lottery win or hedge fund outcome. Similarly on the downside, there may be a critical point that signals an adverse lifestyle transition. In either case, there could be a sympathetic response in terms of subjective probability rescaling. Two general ways of handling this are as follows.

(2a) $W(x) = W(F(x), x)$;

or more specifically,

(2b) $W(x) = W(F(x), u(x))$.

In expression (2a) marginal sensitivity to *x* might be important only around lower or upper critical values, say x_L or x_U, so that $\frac{\partial W}{\partial x}$ is changing rapidly in these zones. In such terms,

$$dW(x) = \frac{\partial W}{\partial F}dF + \frac{\partial W}{\partial x}dx = [\frac{\partial W}{\partial F} + \frac{1}{f(x)}\frac{\partial W}{\partial x}]dF(x),$$

assuming a density for *F* exists. So

(3a) $$\xi_W(x) = \frac{\partial W}{\partial F} + \frac{1}{f(x)}\frac{\partial W}{\partial x}.$$

Similarly for specification (2b),

$$dW(x) = [\frac{\partial W}{\partial F} + \frac{u'(x)}{f(x)}\frac{\partial W}{\partial x}]dF(x),$$

so that

(3b) $$\xi_W(x) = \frac{\partial W}{\partial F} + \frac{u'(x)}{f(x)}\frac{\partial W}{\partial x}.$$

The methodological point in either case is that the rescaling framework continues to apply even where there is a lack of separability as between subjective probability and utility.

Remark 2

In cumulative prospect theory the zone of integration may be split into positive and negative values of *x* to reflect distinct re-weightings of the left and right hand tails:

$$\psi(F) = \begin{matrix} \psi^-(F); & x \le 0 \\ \psi^+(\overline{F}); & x > 0, \overline{F} = 1 - F. \end{matrix}$$

In this case, some adaptations are necessary with differentials:

$$\xi_w(x) = \frac{d}{dF}\psi^-(F(x)); \quad x \le 0$$

$$\xi_w(x) = -\frac{d}{dF}\psi^+(1-F(x)); \quad x > 0\,.$$

Thus the investor weighting function chosen by Tversky-Kahneman (1992) can be cast as:

$$\text{(4)} \qquad \psi(F) = \begin{cases} k\dfrac{F^{\gamma^-}}{(F^{\gamma^-} + (1-F)^{\gamma^-})^{1/\gamma^-}}; F \le F(0) \\ 1 - \dfrac{\overline{F}^{\gamma^+}}{(\overline{F}^{\gamma^+} + (1-\overline{F})^{\gamma^+})^{1/\gamma^+}}; \overline{F} = 1-F,\ F > F(0) \end{cases}$$

where γ^-, γ^+ are positive constants. As it stands, the resulting cumulative weighting function is not continuous at zero, and may not even be monotonic. However, the constant *k* can be chosen[3] so that the two halves splice together with no break at *F(0)*. For example, if $\gamma^- = \gamma^+ = \gamma$, then $k = \dfrac{\Delta}{{F_0}^{\gamma}} - \mu_0^{\gamma}; \Delta_0 = (F_0^{\gamma} + (1-F_0)^{\gamma})^{1/\gamma},\ \mu_0 = \dfrac{1-F_0}{F_0}$.

2.2 Expected utility and observational equivalence

In prospect theory discussions, the utilities of classical Von Neumann-Morgenstern choice theory are recast as 'value functions', and the 'weighting function' replaces probabilities, to arrive at a decision criterion that corresponds to expected utility. As earlier pointed out, however, much the same rescaling process can apply to either. In what follows, the more familiar expected utility framework is utilised for exposition.

Expected utility under the natural measure is

$$\text{(5a)} \qquad E_f[u(x)] = \int_{-\infty}^{\infty} u(x)dF(x)\,.$$

Under the subjective measure it is

$$\text{(5b)} \qquad E_w[u(x)] = \int_{-\infty}^{\infty} u(x)dW(x)\,.$$

Note that expression (5b) could alternatively be written as

[3] An alternative is to adjust both the upper and lower branches, the upper branch being adjusted by $(1+k)(1-\Psi_+(F)) - k$. Set *k* so that $k\Psi_-(0) = \Psi_+(0) - k(1-\Psi_+(0))$.

$$E_W[u(x)] = \int_{-\infty}^{\infty} \tilde{u}(x) dF(x)\,; \tilde{u}(x) = \xi_W(x) u(x)\,,$$

in other words, as a rescaling of utility instead of probability. For any given application (choice of *F, ψ*) there may be observational equivalence between transformed probability and transformed utility, although the equivalent utility function will not remain invariant to different specifications of *F(x).*

Most subjective probability measures entail overweighting the tails of the natural distribution. This means that the original utility function, evaluated at subjective probabilities, is expected utility equivalent under the natural measure to a transformed utility function $\tilde{u}(x)$ that overemphasises extreme values of the domain, perhaps to the point of a radical change in shape or behaviour. To illustrate, suppose $\xi_W(x)$ is a strictly convex function $\xi''(x) \geq 0$ with $\xi_W(x) \leq 1, a < x < b; \xi_W(x) > 1, \textit{otherwise}$. Then apparent risk neutrality under subjective measure *W* corresponds to Friedman-Savage utility functions with natural probabilities. To see this, note that the function $\tilde{u}(x) = x\xi_W(x)$ has the same shape as the typical Friedman-Savage utility function (see figure 1). Thus

$$E_W[x] = \int_{-\infty}^{\infty} xw(x)dx = \int_{-\infty}^{\infty} x\xi_W(x) f(x)dx = \int_{-\infty}^{\infty} \tilde{u}(x) f(x)dx = E_f[\tilde{u}(x)]\,.$$

In such cases, one would not know whether a risk-neutral investor is acting according to a subjective probability that overweights the tails, or on the other hand, is maximising an Friedman-Savage type utility function under natural probabilities. Only if the same investor was presented with range of different situations – and therefore different F's – would one be able to detect the difference.

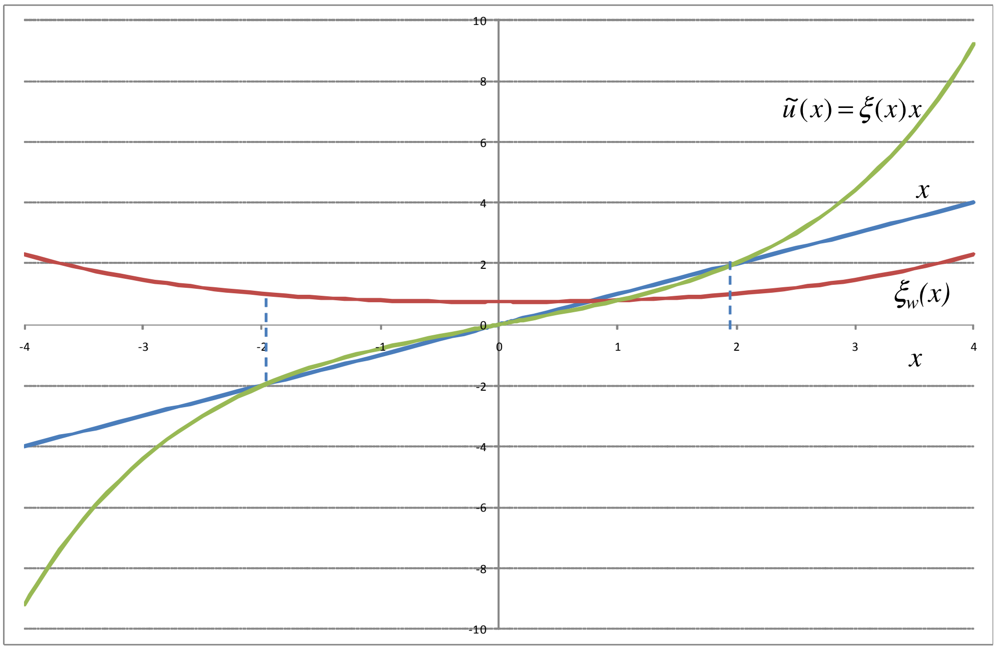

Figure 1: Equivalent utility function

3. Information-based rescaling

Subjective probability, in di Finetti's (1937) formulation, reflects the odds at which a person is just prepared to bet on the given event. Indeed, the longstanding use of fractional odds by bookmakers implies that people are comfortable in thinking and acting in such terms, as a matter of the psychology of gambling. In this respect, the use of the log odds facilitates judgments involving transitivity; thus if event A is viewed as twice as likely to happen than B, and B three times as C, then the subject should intuitively consider event A as 6 times more likely than C (rather than 5). More generally, a predisposition[4] to think in terms of logs is familiar from everyday life, which is why % changes in salary or wage determinations mean more than dollar changes. In the present context, this could mean that in their subjective probability rescaling, people react to the log of tail areas; 'noticeability' is based on log F and log (1-F). In addition, investors are assumed to react to tail length, proxied in terms of the average log odds before, or after, the chosen point. The present section formalises such behaviour in terms of the log odds function and its cumulative, the locational entropy function.

[4] This is sometimes referred to as 'Gibrat's Law'. However, in its original formulation (Gibrat (1931)), this applied to growth in firm size as a rationale for the lognormal distribution, with no specific reference to subjectivism.

3.1 Log odds and locational entropy

The log odds function will be taken as referring to the event that the random variable $X > x$ versus $X \leq x$:

$$\lambda(x) = \log\frac{1-F(x)}{F(x)}; \quad \Lambda(F) = \log\frac{1-F}{F}.$$

An associated concept is locational entropy (Bowden 2011), which is the differential entropy attached to a binary random variable that for any given marker point x takes value 1 if $X \leq x$ or 0 if $X > x$. The locational entropy function is defined as

(6a) $$h(x) = -[F(x)\ln F(x) + (1-F(x))\ln(1-F(x))].$$

It can alternatively be presented as a function of *F:*

(6b) $$H(F) = -[F\ln F + (1-F)\ln(1-F)],$$

so that $h(x) = H(F(x))$. Viewed as a function of x, locational entropy is positive definite, symmetric or asymmetric about the median x_m corresponding to the shape of the density $f(x)$, with $\lim_{x\to\pm\infty} h(x) = 0$. It has a maximum value of ln2 ~ 0.69 at the median, and expected value $E_f[\log(h(x)] = 1/2$. Locational entropy can be identified with the cumulative of the log odds function, such that:

(7) $$H(F) = \int_0^F \Lambda(u)du; \quad h(x) = \int_{-\infty}^{x} \lambda(s)f(s) = F(x)E[\lambda(s) \mid s \leq x].$$

The log odds function can be generalised so that the respective tails F, $1-F$ are weighted unequally. The version used has constant weights:

(8) $$\Lambda_\theta(F) = \log[\frac{(e(1-F))^{1-\theta}}{(eF)^\theta}], 0 < \theta < 1; \lambda_\theta(x) = \Lambda_\theta(F(x)).$$

The cumulative corresponding to (6a) gives biased locational entropy as

(9a) $$h_\theta(x) = -2[\theta F(x)\ln(F(x)) + (1-\theta)(1-F(x)\ln(1-F(x))]; 0 < \theta < 1,$$

(9b) $$H_\theta(F) = -2[\theta F\ln F + (1-\theta)(1-F)(\ln(1-F)]; 0 < \theta < 1.$$

Thus $H_\theta'(F) = \Lambda_\theta(F)$. The normalising factor *exp(1-2θ)* implicit in expressions (8),(9) ensures that $E[\lambda_\theta(x)] = 0$ while $h_\theta(x)$ remains positive with $\lim_{x\to\pm\infty} h_\theta(x) = 0$. The special case $\theta = ½$ reverts to the unbiased version (6a,b).

3.2 Information-based scaling functions

As with classical entropy, inverse constructs derived from locational entropy may be identified with information[5]. In the current context, this can refer to either the additive inverse ($\propto -h(F)$) as type A, or else the multiplicative inverse ($\propto 1/h(F)$)) as type B. The former corresponds to the usual definition of information, and the resulting density $w(x) = \xi(x) f(x)$ is related to a unit shifted versions of *f(x)* that appear in locational entropy[6]. The latter corresponds to the information measure of classical statistics, i.e. the inverse of the expected likelihood function.

The precise variant aside, locational information provides a potential measure for subjective over- or under-valuation of the tails, which according to natural measure are regions where information is highest. The rescaling to be applied at any given point incorporates the average log odds up to that point; or reversing sign, after that point. Thus in a lottery type situation, one could visualise the subject, in reweighting any given point *x*, as weighing up the probabilities of the entire range of outcomes X ≥ x. As *x* increases, and the prospective prizes become greater, the subject becomes more predisposed to do this. The density rescaling at any given point *x* is linked to the cumulative log odds of points thereafter, and hence to the information at that point.

3.2 Rescaling

The rescaling functions incorporate an affine normalisation of the inverse information functions, to ensure that $\Psi(0) = 0; \Psi(1) = 1$ and $E_f[\xi_w(x)] = 1$. The type A version gives rise to:

(10a) $\xi^0_{wA}(F;\beta,\theta) = 1 + \beta(0.5 - h_\theta(F))$, $\xi_{wA}(x;\beta,\theta) = \xi_{wA}(F(x);\beta,\theta)$;

(10b)
$$\Psi(F;\beta,\theta) = -0.5 * \beta(1-\theta) + (1 + 0.5 * \beta)F + \beta[\theta F^2(\ln F - 0.5) - (1-\theta)(1-F)^2(\ln(1-F) - 0.5)].$$

---

[5] Claude Shannon, the originator of entropy, identified it with information, but later scholars, notably Norbert Wiener, identified information with the negative of entropy (Gleick (2011, ch. 9)).

[6] The rescaled density can be written as $w(x) = 1 + \frac{\beta}{2}(1 - \tilde{f}(x))$. The embedded density $\tilde{f}(x) = F(x) f_L(x) + (1 - F(x) f_R(x)$, where $f_L(x) = -f(x)\ln F(x)$ and $f_R(x) = -f(x)\ln(1 - F(x))$ are respectively the unit left and right shifted versions of *f(x)* (Bowden (2011).

The parameter $\beta > 0$ can be chosen[7] to reflect the slope or sensitivity of the scaling function. Note that $\xi_w^0{}'(F;\beta,\theta) = -\beta\Lambda_\theta(F)$. As the log odds function is such that $\Lambda\theta'(F) < 0$, this means that the proposed rescaling function $\xi_w^0(F;\beta,\theta)$ is convex in F. Figures 2a,b illustrate, in both cases with type A informational inverses. As a general observation, information based rescaling weights the tail areas by more than does the Tversky-Kahneman formula (§1.2, expression (4)), drawing more mass away from the median area.

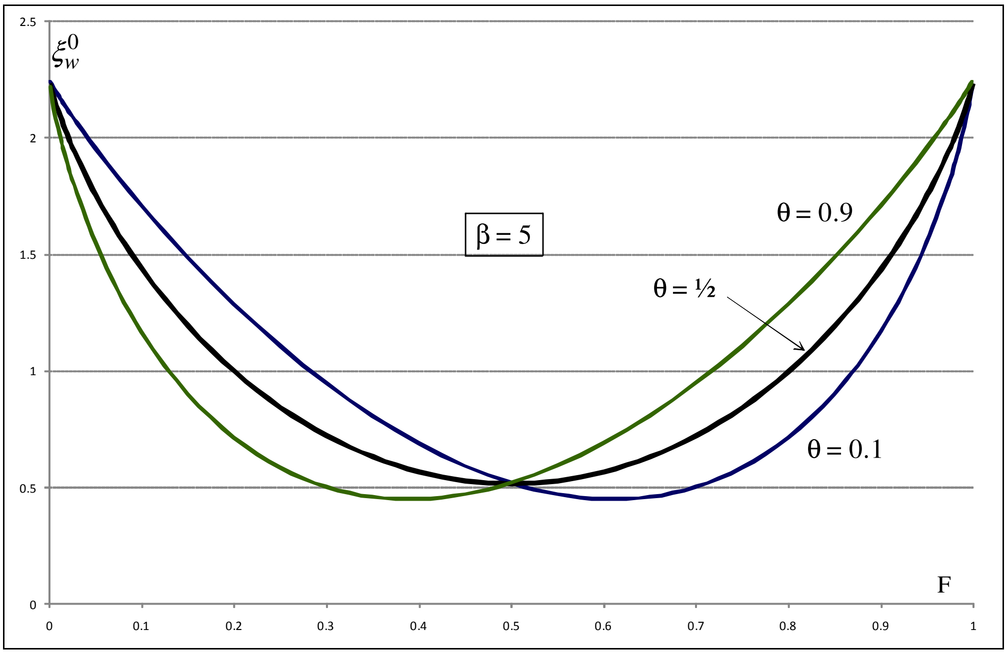

Figure 2a: Density scaling functions

[7] Note that there are limits on β to ensure that ξ > 0. For θ=½, the limit is $\beta < \dfrac{1}{\ln 2 - 0.5} = 5.177399$.

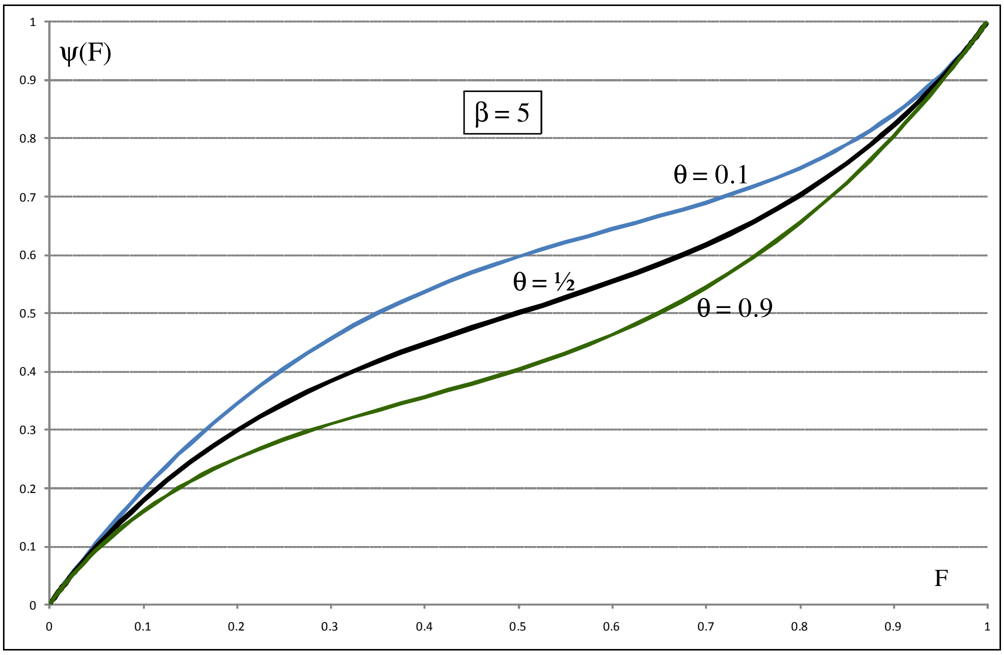

Figure 2b: Cumulative scaling functions.

Turning to type B informational inverses, the general version is

(11) $$\xi^0_{wB}(F;\beta,\theta) \propto \frac{1}{H_\theta(F)^\beta}; 0<\beta,\theta<1,$$

with the proportionality constant determined by the unit expected value. The extent to which both tails are inflated can be adjusted with the parameter β; a higher value induces a more extreme tail rescaling. Accentuating one of other of the tails can be achieved via the asymmetry parameter θ, as for type A. The general rescaling effect of (11) is to inflate the extreme tails by more than the additive version (10).

Note that as α→1, the integral $\int_0^1 \xi^0_{wB}(F;\alpha,\theta)dF$ becomes divergent, which means that $E_f[\xi_{wB}(x);\alpha,\theta]$ does not technically exist. Numerically, the divergence becomes a problem from about α = 0.85 upwards. However, if the objective is just to develop a scaling function to cover all except minute tail areas, an ad hoc truncation is to convert the closed interval [0,1] into the open interval (0,1), e.g. as $[0+\delta, 1-\delta]$ for small δ, and renormalising to start with $\psi(\delta)$ reset to zero.

Figure 3 compares natural and rescaled densities for type A and B informational rescaling. The natural density is taken as Cauchy, a long tailed distribution that has been proposed for use with financial returns data (Suàrez and Menèndez 2005). The rescalings inflate the tails, drawing mass away from the median area, the more so with type B.

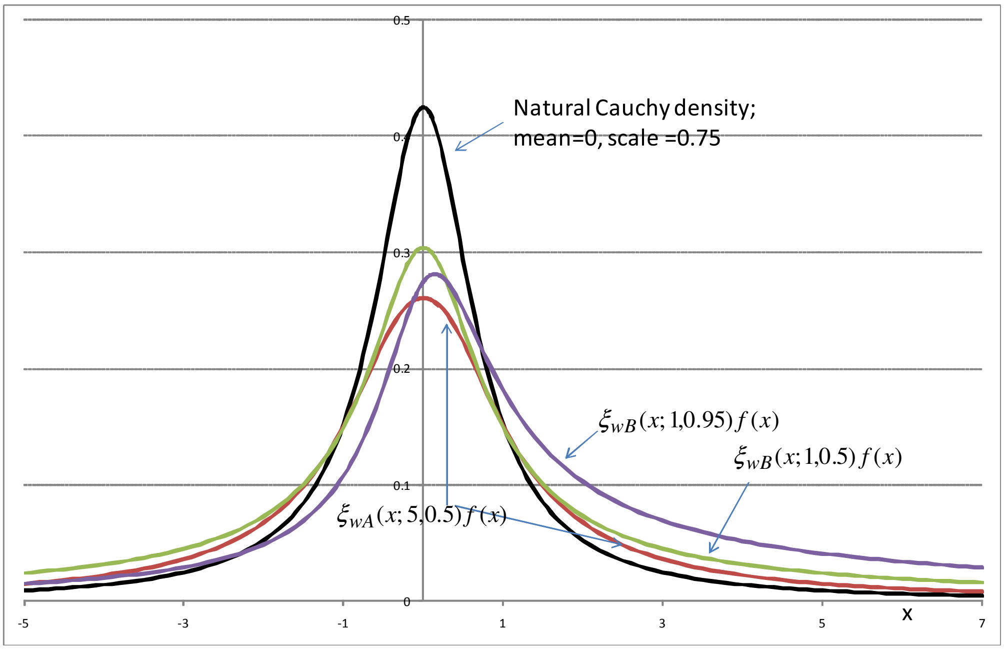

Figure 3: Cauchy distribution: Selection of natural and subjective densities.

4. Risk and subjective probability

Over-emphasising right hand tail probabilities as a basis for action raises an issue as to whether investors are exposing themselves to an undue amount of risk – rose tinted spectacles come with a cost. Alternatively, investors who overemphasise the left hand tail might be excessively reluctant to undertake potentially profitable opportunities. In this case, subjective probabilities would entail an opportunity cost. In making such judgements[8], actions that would otherwise have been taken under the natural distribution are an appropriate benchmark. Accordingly, a range of comparative risk management and investment criteria can be related to properties of the subjective rescaling function $\xi_w(x)$. The following lemma expresses common risk management metrics in such terms:

[8] One could imagine back office staff as risk managers, attempting to monitor the front office dealers, who might exhibit excessive bonus seeking behaviour and be predisposed to right hand tail inflation.

Lemma

Let $\Phi_f(X) = \int_{-\infty}^{X} F(x)dx$ *; similarly* $\Phi_W(X)$ *. Then*

(12a) $F(X)E_f[(\xi_W(x)-1) \mid x \le X] = W(X) - F(X)$

(12b) $\int_{-\infty}^{X} (x-X)(1-\xi_W(x))f(x)dx = \Phi_W(X) - \Phi_f(X)$

(12c) $\int_{-\infty}^{X} x(1-\xi_W(x))f(x)dx = F(X)E_f[x \mid x \le X] - W(x)E_W[x \mid x \le X]$.

[The proof is straightforward and will be omitted.]

Condition (12a) relates to first order stochastic dominance (FSD). The natural distribution F is FSD over the subjective distribution *W* if for all X, $E_f[(\xi_W(x)-1) \mid x \le X] \le 0$. An equivalent condition is that $\psi(F) \le F$, all F; the graph of Ψ against F remains below the 45° line.

Expression (12b) gives rise to the second order stochastic dominance condition as $\int_0^F (\psi(\tilde{F}) - \tilde{F})d\tilde{F} \le 0$ all F. In other words, the subjective distribution *W* is SSD over the natural *F* if the accumulated area between *Ψ(F)* and the 45° line remains negative as *F* increases.

Condition (12c) casts the comparison in terms of conditional value at risk (CVaR), or expected shortfall. The RHS of (10c) is an index of the relative downside exposure, namely the CVaR's adjusted for the regime probabilities under the alternative measures. If the left hand side is positive for all *X*, then the downside apparent cost of risk is less with the adjusted measure *W*.

To result in distribution *W* being second order stochastic dominant over *F*, a possible scaling function would be one which commences with values less than unity in the left hand tail, and rises thereafter. A logistic growth curve for ξ against *F* would be one such shape. But W SSD F is not ideal from the point of view of risk management. Investors who believe the subjective version *W* could well be risk averse in terms of their utility function and select an investment or portfolio accordingly. But the fact that *W* is

SSD over the true *F* means they are underestimating the true risk, for the left hand tail area under *F* is greater.

Scaling functions and their subjective distributions can therefore be hazardous unless they overweight both tails, not just the desirable right hand tail. Correspondingly, the graph of Ψ(F) against F should resemble the typical cubic power shape as in figure 2. All the rescaling functions considered in the present paper have this property. From the point of view of evolutionary economics, it has some advantages. Such biases help to keep investors out of trouble on the downside, yet benefit from risk-seeking upside gambles that might otherwise be unwarranted in terms of more objective probability measures[9].

References

Benartzi, S. And R.H. Thaler (1995), Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics*, 100, 73-92.

Billingsley, P. (1995), *Probability and Measure*, 3rd ed. New York: Wiley.

Boehm, M. (2011), Drivers of relative attractiveness for cumulative prospect theory investors. Working paper, Institute of Finance, Ulm University.

Bowden. R.J. (2011), Information, measure shifts and distribution metrics. *Statistics*, forthcoming, available online at DOI:10.1080/02331888.2010.508560 .

Bowden, R.J. (2010), Directional entropy and tail uncertainty, with applications to financial hazard. *Quantitative Finance*, 11, 437- 446.

De Finetti, B. (1937), La Prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7, 1-68.

De Finetti (1970), *Theory of Probability*, New York: Wiley. Trans. A. Machi &A.F.M. Smith vols 1,2 New York: Wiley, 1974-5.

Duffie, D. (2001), *Dynamic Asset Pricing*, 3rd ed. Princeton: University Press.

Friedman, M., and L.J. Savage (1948), Utility analysis of choices involving risk. *Journal of Political Economy* 56, 279–304.

Gibrat R. (1931), *Les Inégalités Ѐconomiques*, Paris: Recueil Sirey.

9 The Conquistadores in Chile and Peru are a possible instance. They found very little gold, relative to the optimistic prior beliefs that drove their expeditions – the headdresses and ornaments of the indigenous Aymara and Quechua cultures were mostly gold plate. Yet the Spanish colonialists benefited in the long run from other products.

Giorgi, E.D. and T. Hens (2006), Making prospect theory fit for Finance. *Financial Markets and Portfolio Management*, 20, 339-360.

Gleick, J. (2011), *The Information*. New York: Pantheon/Random House.

Grinblatt, M., and B. Han (2005), Prospect theory, mental accounting and momentum. *Journal of Financial Economics*. 78, 311-339.

Harrison, M., and D. Kreps (1979), Martingales and arbitrage in multiperiod security markets, *Journal of Economic Theory* 20, 381--408.

Menkhoff, L., and M. Schmeling (2006), A prospect theoretical interpretation of momentum returns. *Economics Letters*, 93, 360-366.

Prelec, D. (1998), The probability weighting function. *Econometrica*, 66,497-527.

F.P. Ramsey, F.P. (1926), Truth and probability, Ch.7 of R.B. Braithwaite (ed), *The Foundations of Mathematics and Other Logical Essays*, 1931, London: Routledge. Also in D.H. Mellor (ed) *Foundations: Essays in Probability, Logic, Mathematics and Economics*, 1978, London: Routledge.

Pfanzagl, J (1967), Subjective Probability Derived from the Morgenstern-von Neumann Utility Theory. In M. Shubik. *Essays in Mathematical Economics In Honor of Oskar Morgenstern.* Princeton: University Press.

Savage, L.J. (1954), *The Foundations of Statistics*, New York: Wiley.

Suàrez, S. and S.C. Menèndez (2005), Computational tools for the analysis of market risk, in Menèndez, S.C., A.S. Callè and L. Seco (eds) *Risk management in Finance*, Proc. First Risk Lab international conference, Bilbao: Fundación B.B.V.A.

Tversky, A., and D. Kahneman (1974), Judgment under uncertainty: Heuristics and biases. *Science, New Series,* 185, 1124-1131.

Tversky, A., and D. Kahneman (1979), Prospect theory: An analysis of decision under risk. *Econometrica,* 47, 263-292.

Tversky, A., and D. Kahneman (1992), Advances in prospect theory: Cumulative representation under uncertainty. *Journal of Risk and Uncertainty,* 5, 297-323.

Von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton: University Press.