Measuring Risk Relatively

Dr John Fountain* **June 2002**

Abstract

Kullback Leibler (KL) discrepancy measures risk as differences in knowledge between trading partners. Duality theory and scoring rule methods show that certainty equivalents, Hicksian surplus and Pratt Arrow risk premia for CARA agents (exactly) and proper utilities (approximately) depend linearly on expected wealth and KL discrepancy. Three applications show (1) relative entropy, not Shannon entropy, is an appropriate measure of risk and value of information, (2) incomplete markets based on aggregates may be first best efficient, (3) CARA agents share risk efficiently by acting as if they maximize a representative aegnt's CARA utility function defined over aggregate wealth. (JEL C44, D81)

Introduction and Overview

This paper uses duality theory and the statistical concept of Kullback Leibler (KL) discrepancy to develop a new measure of risk based on differences in knowlege. We define knowledge by the ability to predict consequences accurately and assess predictive accuracy operationally using scoring rules for probabilty forecasters (Leonard J Savage(1971), Frank Lad(1996,Chapter 6), Jose Bernardo and Adrian Smith(1994, Chapter 2)). The Log scoring rule turns out to be a natural way of assessing differences in knowledge, precisely for constant absolute risk averse (CARA) agents and approximately for all "proper" utility of wealth functions (mixtures of CARA utilities). By "natural" we mean natural to economists, not necessarily to statisticians, as our argument relies on ideas of constrained optimal choice in markets and the indirect assessment of wealth risks through supporting budget sets, rather than directly in terms of quantities.

The paper has three substantive sections and a short conclusion. Section I analyzes a CARA agent's choice of wealth risks in competitive contingent claims markets. The ordinary and compensated demand functions, the indirect utility function and the expenditure function emerging from this analysis all have simple functional forms driven by two sets of two summary statistics: expected values of wealth and the Kullback-Leibler (KL) discrepancy (relative entropy) between the agent's beliefs and the beliefs expressed in market prices. KL discrepancies are expectations of several key concepts from Bayesian statistics: Log scores, weights of the evidence or Log Bayes factors in favour (±) of the relevant alternative predictive hypotheses.

Section II explains that smart money follows the weight of evidence. That is, variations across states in weights of evidence determine optimal variations in contingent wealth holdings for CARA agents. The overall benefits and risks of trade are measured by the two forms of KL discrepancy. After adjusting for endowment effects and levels of risk tolerance, one form measures Hicksian multiproduct consumer surplus in contingent claims markets while the other form measures the Pratt Arrow risk premium faced by a CARA agent. Two logically equivalent formulae for certainty equivalents are derived, one combining surplus with market value of wealth, the other combining Pratt-Arrow risk premia with subjective expectations of wealth. The section concludes by explaining the difference between *indirect* risk premia based on optimisation in markets and *direct* risk premia based on Jensen's inequality, by showing why wealth variance measures do not matter (directly) to CARA agents, and by indicating how exact results for CARA agents can be extended to all "proper" utility functions based on the work of John Pratt and Richard Zeckhauser(1987) and Patrick Brockett and L Golden (1987).

Section III contains three brief applications. The first provides sufficient conditions for the economic relevance of Shannon entropy as a measure of economic surplus and of risk. Low entropy in a price distribution is fully captured as Hicksian surplus for maximally uninformed (minimal entropy) CARA agents, but at a cost in terms of high relative risk. If the market is maximally uninformed a self referentially informed agent (low entropy in her beliefs) can achieve high surplus, but at a cost in terms of a risk premium directly proportional to the entropy difference between the agent's and the market's beliefs. The second application shows how KL discrepancy can be used in the analysis of costs, benefits, and risks, of trading in *incomplete* contingent claims markets. Market incompleteness may not lead to any loss of surplus value nor change any risk premia when trading agents share certain conditional inference strategies relating macro to micro states. Trading wealth contingent on sufficient statistics substitutes perfectly for trading in wealth contingent on detailed underlying micro states in such cases. The third application uses the Gorman polar form of the indirect certainty equivalent function for CARA agents to analyze efficient risk sharing arrangements. Efficient allocations depend not only on relative risk tolerances and endowments of aggregate state contingent wealth but also on the weight of the evidence for and against the competing hypotheses of participating traders. Pareto-efficiency in risk sharing follows the weight of the evidence. An exact, analytically tractable general equilibrium pricing formula for an exchange economy of heterogeneous CARA agents is derived showing that equilibrium relative prices in Log odds form are a linear function of a geometric mean of the beliefs of the participating agents expressed as odds and any covariance between risk attitudes and beliefs. Equilibrium prices also turn out to be supporting prices for the optimal choices of a representative CARA agent making expected utility maximizing choices about aggregate wealth, where the representative agent has risk tolerance the sum of the risk tolerances of the participating agents and beliefs in odds form a geometric mean of the odds assessed by the participating agents.

Section I *Optimal choice and Kullback Leibler discrepancy*

Consider a CARA agent with negative exponential utility of wealth function u(z)=$-e^{-\frac{1}{\tau} z}$, Pratt-Arrow risk tolerance index $\tau = -u' / u''$, and beliefs described by a coherent pmf (probability mass function) ***q(s)***={q(1)...q(n)} for a discrete uncertain state variable S with a finite realm of possible values $s\epsilon\{1,...n\}$. Following Frank Lad(1996,42-45) we define the *constituent events* for S as the logical truth value of propositions of the form "S= s", denoted by the assertion in brackets (S= s). Each (S= s), is an uncertain 0/1 valued quantity that can be publicly verified and therefore used to define tradeable units of contingent wealth paying $1 if the proposition "S= s" is true and $0 otherwise. We assume that the agent can buy and sell unit contingent wealth claims (S= s) in markets as a price taker subject to a budget constraint ***p•ω=p•z***, where ***z(s)***={z(1)...z(n)} is a non-negative vector of contingent wealth levels, ***ω(s)***={ω(1)....ω(n)} is the agent's non-negative endowment of contingent wealth, and ***p(s)***={p(1)...p(n)} is a set of non-negative prices for unit contingent claims (S= s). In our notation, n-tuples like ***q(s), ω(s)*** etc are indicated by bold italics, we typically suppress arguments when they can be readily inferred from the context, eg writing ***q*** for ***q(s),*** and ***x•y*** indicates the conventional inner product $\sum_S$ x(s)y(s) .

Prices ***p*** must satisfy the constraint $\sum_S p(s)=1$, otherwise arbitrage can create Dutch books against the market[1] . Since "the market" stands willing to buy or sell unit contingent claims at prices that are non-negative and sum to 1, prices ***p*** can be interpreted as the market's *beliefs* about the corresponding underlying states S. The market prices ***p*** faced by the agent are assumed to be determined in a bargaining process that we don't explicitly analyze in this paper (but Section III.C develops general competitive equilibrium prices for a pure exchange economy of CARA agents to analyze efficient risk sharing arrangements).

A CARA agent's expected utility maximisation problem subject to a budget constraint ***p•ω=p•z*** and dual expenditure minimisation problem subject to an expected utility (EU=eu) or certainty equivalent (CE=ce) constraint yield the following value functions and solutions (see Appendix A for details):

1.1• $EU(\boldsymbol{p},\omega,\boldsymbol{q},\tau)=-e^{-\frac{1}{\tau}(\boldsymbol{p}\bullet\omega+\tau D(\boldsymbol{p}\|\boldsymbol{q}))}$ indirect utility function

1.2• $\boldsymbol{z}^o(s,\boldsymbol{p},\omega,\boldsymbol{q},\tau)=\boldsymbol{p}\bullet\omega+\tau\{D(\boldsymbol{p}\|\boldsymbol{q})+\text{Ln}\,\frac{q(s)}{p(s)}\}$ ordinary demand function

1.3• $m(\boldsymbol{p},eu,\boldsymbol{q},\tau)=CE(eu)-\tau D(\boldsymbol{p}\|\boldsymbol{q})$ expenditure function

1.4• $\boldsymbol{z}^h(s,\boldsymbol{p},eu\boldsymbol{q},\tau)=CE(eu)+\tau\text{Ln}\,\frac{q(s)}{p(s)}$ compensated demand function

The expression D(***p***||***q***) in equations 1.1→1.3 is the Kullback Leibler (KL) discrepancy between two pmfs ***p*** and ***q***, defined as:

1.5• $D(\boldsymbol{p}\|\boldsymbol{q})=\sum_{s=1}^{n} p(s)\,\text{Ln}\,\frac{p(s)}{q(s)}$

KL discrepancy is an important concept in statistical information theory (Cover and Thomas(1992), measuring relative entropy, and in Bayesian statistical inference, measuring (an expectation of) comparative predictive performance of probability forecasters as assessed by the Log scoring rule, (Savage(1971), Lad(1996), Bernardo and Smith(1994), von Winterfeldt and Edwards (1986)). Although KL discrepancy is not a proper *distance* metric on pmfs, since neither the axiom of symmetry nor the triangle inequality is satisfied, it has many useful mathematical properties (Cover and Thomas(1991), Theorems 2.6.3 and 2.7.2): for any two pmfs ***a***, ***b***

KL1 D(***b***||***a***)≥0 with equality iff *b(s)=a(s)* for all states *s (non-negativity)*

KL2 D(***b***||***a***) is *convex* in the pair ***(b,a)***

The optimal choices and values 1.1-1.4 are directly connected the Log scoring rule. Scoring rules generally can be viewed either as incentive structures in a stylized belief elicitation game or as evaluative criteria operationally defining "good" forecasters , or both. A wishes to know what B believes about the state variable S or how well B cam predict S. A presents B with a payoff (score) function that depends on what happens, S=*s*, and what B reports will happen, where the report ***r*** is in the form of a pmf on S. A *proper* scoring rule maximizes A's expected score payoff when A reports her personal beliefs about S.

The Log scoring rule is one such proper scoring rule, paying Ln *r(s)* , the logarithm of the probability assessed for the event (S=*s*) that actually occurs .[2] Since logs of fractions are negative, payments are penalties according to the Log scoring rule, increasing in total and at the margin the farther the probability report *r*(s) for (S=*s*) is from 1, perfect, one-off prediction. Since perfect, clairvoyant prediction has a Log score of Ln(1)=0, predictive ability is worse the lower the realised Log score. We refer to the idea of measuring predictive ability by the Log scoring rule as L-predictive ability throughout the paper to distinguish it from other incentive structures or evaluative criteria for eliciting and valuing forecasting pmfs (eg quadratic, pseudo-spherical, linear, etc. type scoring rules as outlined in Lad (1996, Ch 6)).

The *relative* L-predictive abilities of different forecasters, one using pmf ***q***, the other using pmf ***p***, can be compared by examining the Log-scores difference: Ln *q(s)* - Ln *p(s)*= Ln $\frac{q(s)}{p(s)}$ for event (S=*s)* . The n-tuple of these differences, denoted $\boldsymbol{Ln\,\frac{q}{p}}\equiv(\mathrm{Ln}\,\frac{q(1)}{p(1)},,,\mathrm{Ln}\,\frac{q(n)}{p(n)})$, describes the variation in relative L-predictive ability between pmfs ***q*** and ***p*** across all constituent events for S. KL discrepancy, also know as *relative entropy*, is an *expectation* of this variation in relative L-predictive ability: $D(\boldsymbol{q}||\boldsymbol{p})=\boldsymbol{q\bullet Ln\,\frac{q}{p}}$ for the agent's expectation of its relative L-predictive ability and $D(\boldsymbol{p}||\boldsymbol{q})=\boldsymbol{p\bullet\, Ln\,\frac{p}{q}}$ for the market's expectation of its own relative L-predictive ability.

Relative entropy is a generalisation of the concept of Shannon entropy, $H(\boldsymbol{a(s)})=-\Sigma_S\, a(s)\mathrm{Ln}(a(s))$ for a pmf ***a(s)***. Shannon entropy is a *self referential* measure of L-predictive ability (an own expected Log score penalty for pmf ***a***). Extremes of Shannon entropy occur at the uniform distribution $\boldsymbol{n(s)}=\frac{1}{n}$ on S, with maximum entropy value Ln(*n*), and when pmf ***a(s)*** asserts probability 1 for one particular state and zero for all others, a minimum entropy value of 0. Since $D(\boldsymbol{a(s)}||\boldsymbol{n(s)})=\mathrm{Ln}(n)-H(\boldsymbol{a(s)})$, *improvements* in self referential L-predictive ability described as *reductions* in Shannon entropy H can be thought of as improvements in expected *relative* L-predictive ability compared to a uniform distribution or "random chance". Hence Shannon entropy measures relative knowledge with uniform chance as a specific comparator, whereas relative entropy , KL discrepancy, admits a variable range of comparator pmfs.

There is a useful Bayesian interpretation of differential Log scores and their expectations. Suppose ***p*** and ***q*** express predictive hypotheses H_p and H_q about the likelihood of specific states S. Extending the discourse about uncertainties to incorporate uncertainty about *hypotheses* H as well as about *states* S, let someone with a coherent joint pmf on hypotheses and states assess prior odds Prob(H_q)/Prob(H_p) on the two relevant hypotheses. By Bayes theorem, posterior odds on hypotheses conditional on evidence (S=*s)* are:

$$\frac{\text{Prob}\,(H_q|s)}{\text{Prob}(H_p|s)} = \left(\frac{q(s)}{p(s)}\right)\frac{\text{Prob}\,(H_q)}{\text{Prob}(H_p)}$$, where $\left(\frac{q(s)}{p(s)}\right)$ is the Bayes factor for event S=*s*.

Log Bayes factors, Ln($\frac{q(s)}{p(s)}$), are known as the *weight of the evidence* S=*s* in favor of H_q over H_p (Good(1998),Kass and Raftery (1995)). KL discrepancies D($\boldsymbol{q}||\boldsymbol{p}$)=$\boldsymbol{q \cdot Ln \frac{q}{p}}$ and D($\boldsymbol{p}||\boldsymbol{q}$)=$\boldsymbol{p \cdot Ln \frac{p}{q}}$ are the agent's and the market's respective expectations for the weight of evidence in their favour, and therefore for the amount of *change* from prior to posterior odds on hypotheses. By KL1, each agent expects a positive weight of evidence in favour of their own hypothesis and a negative weight of evidence in favour of their trading partner's hypothesis. The farther apart ***p*** and ***q*** are in KL discrepancy the greater the expected Log Bayes factors or weights of evidence, and hence the greater the expectation for change in the amount of possible "updating" from prior to posterior odds on *hypotheses*.

Units of Log score differences, KL discrepancies and weights of evidence are "nats" using natural logs or "bits" using logs to the base 2. How many nats does it take to make up an "important" score difference or weight of evidence? Table 1, taken from a recent review article by Kass and Raftery (1995), suggests a range of qualitative interpretations relevant to scientific investigation:

Table 1:Log Score differences and Weight of Evidence

$\frac{q(s)}{p(s)}$	$Ln\frac{q(s)}{p(s)}$	*Evidence for* ***q*** *over* ***p***
1 → 3	0 → 1.1	Barely mentionable
3 → 10	1.1 → 2.3	Positive
10 → 100	2.3 → 4.61	Strong
>100	over 4.61	Very Strong

The quantities $\boldsymbol{Ln\frac{q}{p}}$, D(***q***||***p***),and D(***p***||***q***) indicate the amount of variability (actual and expected) in *relative* L-predictive ability or Log Bayes factors or the weight of evidence for or against one's predictive hypothesis *relative to* some other predictive hypothesis. Hence they can be taken as a measure of *relative knowledge risk*. When ***q***=***p*** there is no relative variability, and hence no relative knowledge risk, however much "absolute" variability or imperfect forecasting there is in comparison to a clairvoyant forecaster. At an intuitive level, larger (actual or expected) variability in relative weights of evidence for the competing hypotheses at stake suggests greater relative risk. Equations 1.1-1.4 indicate that these relative *knowledge* risks drive optimally chosen *wealth* risks for CARA agents in contingent claims markets.

Section II *Smart money follows the weight of evidence*

Market opportunities for trading contingent claims make it possible to *simulate* relative knowledge risks $\boldsymbol{Ln\frac{q}{p}}$. From a position of intially certain wealth consider the transaction $\boldsymbol{Ln\frac{q}{p}}$, purchasing $\text{Ln}\frac{q(s)}{p(s)}$ units of (S=*s*) when the agent expects S=*s* to be more likely than the market and selling $|\text{Ln}\frac{q(s)}{p(s)}|$ units of (S=*s*) when she expects S=*s* to be less likely than the market. From equation 1.2, the variability across constituent events (S=*s)* in optimal wealth for CARA agents is proportional to this simulated knowledge risk transaction, where the agent's risk tolerance is the factor of proportionality. *Smart money follows the weight of evidence*.

This observation is true not only for actual optimal choices in each state, but also in expectation. The agent's subjective value of smart money follows the expected weight of evidence. To see this, consider the direct *certainty equivalent* CE referred to in 1.3-1.4 for a contingent commodity $\boldsymbol{z}$ with expected utility $EU(\tau,\boldsymbol{q},\boldsymbol{z})=eu$ is :

1.6• $CE(\boldsymbol{z})=CE(eu) = \tau[-Ln(-EU(\tau,\boldsymbol{q},\boldsymbol{z}))]$,

Viewing contingent wealth $\boldsymbol{z}$ as an optimal choice $\boldsymbol{z}^{o}(\boldsymbol{p},\boldsymbol{\omega},\boldsymbol{q},\tau)$ for some supporting budget set $((\boldsymbol{p},\omega))$,(see Appendix A for details), the *indirect* certainty equivalent for the trading opportunity $(\boldsymbol{p},\omega)$ is:

1.7• $CE(\boldsymbol{p},\boldsymbol{\omega},\boldsymbol{q},\tau)=\boldsymbol{p}\bullet\omega + \tau D(\boldsymbol{p}||\boldsymbol{q})$

$CE(\boldsymbol{p},\boldsymbol{\omega},\boldsymbol{q},\tau)$ is an indirect utility function, 1.1 in logarithmic form, depending only on the market value of the agent's wealth $\boldsymbol{p}\bullet\boldsymbol{\omega}$ and KL discrepancy $D(\boldsymbol{p}||\boldsymbol{q})$. It has the Gorman polar form $F(\boldsymbol{p},m)=A(\boldsymbol{p})m+B(\boldsymbol{p})$ in income m and prices $\boldsymbol{p}$ (Cornes(1992,pp.53, 82, 194) with useful aggregation properties that we revisit in section III

Expected wealth is also directly related to expected knowledge risks. From 1.2, the *agent's subjectively expected wealth* at her optimal choice, $\boldsymbol{q}\bullet\boldsymbol{z}^{o}$ is:

1.8• $\boldsymbol{q}\bullet\boldsymbol{z}^{o} =\boldsymbol{p}\bullet\omega + \tau[D(\boldsymbol{q}||\boldsymbol{p})+D(\boldsymbol{p}||\boldsymbol{q})]$

Since $D(\boldsymbol{q}||\boldsymbol{p})+D(\boldsymbol{p}||\boldsymbol{q})= \boldsymbol{q}\bullet Ln\frac{\boldsymbol{q}}{\boldsymbol{p}} - \boldsymbol{p}\bullet Ln\frac{\boldsymbol{q}}{\boldsymbol{p}}$ subjectively expected wealth at an optimal choice is the market value of the agent's wealth $\boldsymbol{p}\bullet\omega$ plus the risk tolerance adjusted *difference* between the agent's (positive) and the market's (negative) expectations of the agent's relative L-predictive ability in forecasting constituent state variable events. *Differences in expectations about relative* *knowledge risks translate into differences in expectations about wealth*, the more so the more risk tolerant the agent is.

1.7 expresses the agent's indirect certainty equivalent in terms of *market expectations* $\boldsymbol{p}$, for the agent's endowed wealth ω, and for her relative L-predictive ability $\boldsymbol{Ln}\frac{\boldsymbol{q}}{\boldsymbol{p}}$ compared to the market. Taking 1.7 and 1.8 together, the indirect certainty equivalent can also be expressed in terms of *subjective expectations* $\boldsymbol{q}$ of wealth at an optimal choice less expectations about variations in relative L-predictive ability $\boldsymbol{Ln}\frac{\boldsymbol{q}}{\boldsymbol{p}}$ compared to the market:

1.9• $\mathrm{CE}(\boldsymbol{p}, \omega, \boldsymbol{q}, \tau) = \boldsymbol{q} \bullet \boldsymbol{z}^{o}(\boldsymbol{p}, \omega, \boldsymbol{q}, \tau)) - \tau \mathrm{D}(\boldsymbol{q} \| \boldsymbol{p})$

From 1.9, an *exact* measure of the *indirect* Pratt-Arrow risk premium $\mathrm{RP}(\boldsymbol{z}^{o})$ for a trading opportunity$(\boldsymbol{p}, \omega)$ to a CARA agent is:

1.10• $\mathrm{RP}(\boldsymbol{z}^{o}(\boldsymbol{p}, \omega, \boldsymbol{q}, \tau))) = \tau[\mathrm{D}(\boldsymbol{q} \| \boldsymbol{p})]$

Note that the indirect risk premium (1.10) is based on one form of KL discrepancy $\mathrm{D}(\boldsymbol{q}\|\boldsymbol{p})$ while the indirect certainty equivalent is based on the other $\mathrm{D}(\boldsymbol{p}\|\boldsymbol{q})$ (1.7).

1.10's indirect risk premium is determined by two considerations, one purely preference based, the agent's risk tolerance τ, the other partially constraint based, the KL discrepancy $\mathrm{D}(\boldsymbol{q}\|\boldsymbol{p})$. $\mathrm{D}(\boldsymbol{q}\|\boldsymbol{p})$ relative *knowledge* risks are a natural, standardized measure of the amount of *wealth* risk in a trading opportunity ($\boldsymbol{p}, \omega$) faced by a CARA agent with beliefs $\boldsymbol{q}$ and risk tolerance $\tau=1$. Any other CARA agent can compute her own wealth risks in the sense of Pratt-Arrow risk premia by scaling the *standardized* risk premium measure $\mathrm{D}(\boldsymbol{q}\|\boldsymbol{p})$ by her risk tolerance τ.

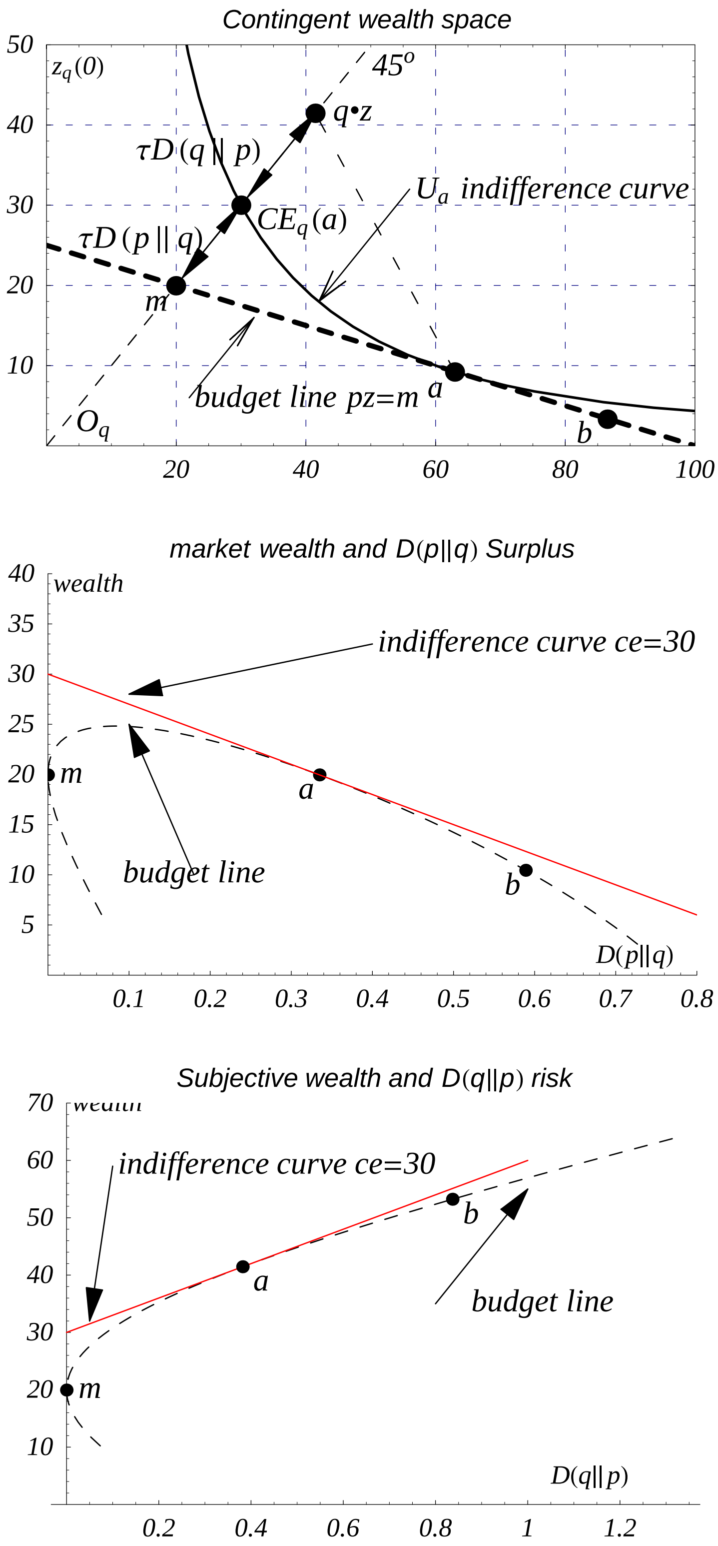
1.7 provides a valuation formula. The personal value of a risky trading opportunity ($\boldsymbol{p}, \omega$) is the sum of the market value of the agent's wealth endowment $\boldsymbol{p} \bullet \omega$ and the risk tolerance adjusted KL discrepancy $\mathrm{D}(\boldsymbol{p}\|\boldsymbol{q})$. We show in equation 1.11 below that $\mathrm{D}(\boldsymbol{p}\|\boldsymbol{q})$ is a *standardised* measure of economic surplus or gains from trade for CARA agents. Hence prices and endowments yielding higher market values for endowments $\boldsymbol{p}\bullet\omega$ and or standardised surplus D($\boldsymbol{p}$||$\boldsymbol{q}$) are preferred. When tradeoffs are necessary, incremental (±) standardised surplus D($\boldsymbol{p}$||$\boldsymbol{q}$) is valued in terms of market value of wealth at a constant marginal rate τ per unit. Notice that, absent endowment effects, prices $\boldsymbol{p}$ farther away from beliefs $\boldsymbol{q}$ in KL discrepancy D($\boldsymbol{p}$||$\boldsymbol{q}$) are preferred by all CARA agents, the more so the larger the agent's risk tolerance. Another way to say this is that CARA agents with given beliefs prefer *mean preserving price spreads* of market prices, as long as "mean" is understood as the *market* expectation of the agent's wealth $\boldsymbol{p}\bullet\omega$ and price "spread" is measured relative to the agent's beliefs $\boldsymbol{q}$ by KL discrepancy D($\boldsymbol{p}$||$\boldsymbol{q}$).

1.9 provides a complementary valuation formula in terms of the agent's expectations $\boldsymbol{q}$. The personal value of a risky trading opportunity is subjectively expected wealth from the associated optimal choice $\boldsymbol{q}\bullet\boldsymbol{z}^o(\boldsymbol{p}, \omega, \boldsymbol{q}, \tau))$ less an adjustment, the indirect risk premium, measured by τD($\boldsymbol{q}$||$\boldsymbol{p}$). CARA agents prefer trading opportunities yielding high subjectively expected wealth and smaller indirect risk premia. When tradeoffs are necessary, incremental (±) standardised risk in the form of D($\boldsymbol{q}$||$\boldsymbol{p}$) is valued in terms of subjectively expected wealth at a constant marginal rate τ per unit. Note that at given prices $\boldsymbol{p}$ D($\boldsymbol{q}$||$\boldsymbol{p}$)-mean preserving spreads of beliefs $\boldsymbol{q}$ are dis-preferred by CARA agents.[3]

The valuation formulas 1.7 and 1.9 are simply taking advantage of the isomorphism defined in duality theory between prices and quantities. Using this isomorphism, any two contingent wealth bundles ***x(s)*** and ***y(s)*** supported by budget sets ($\boldsymbol{p}_x, \omega_x$) and ($\boldsymbol{p}_y, \omega_y$) as optimal choices, $\boldsymbol{x}$= $z^o(\boldsymbol{p}_x, \omega_x, \tau, \boldsymbol{q})$ and $\boldsymbol{y}$=$z^o(\boldsymbol{p}_y, \omega_y, \tau, \boldsymbol{q})$, will will be ranked in preference as: EU(τ,$\boldsymbol{q}$,$\boldsymbol{x}$)≥EU(τ,$\boldsymbol{q}$,$\boldsymbol{y}$) iff $[\boldsymbol{p}_x\bullet\omega_x - \boldsymbol{p}_y\bullet\omega_y] + \tau[D(\boldsymbol{p}_x||\boldsymbol{q}) - D(\boldsymbol{p}_y||\boldsymbol{q})] \geq 0$ iff $[\boldsymbol{q}\bullet\boldsymbol{x} - \boldsymbol{q}\bullet\boldsymbol{y}] - \tau[D(\boldsymbol{q}||\boldsymbol{p}_x) - D(\boldsymbol{q}||\boldsymbol{p}_y)] \geq 0$. [Note that endowments are only determined up to an equivalence in terms of market value of endowed wealth.]

Figure 1 shows how these three equivalent ways of representing preferences and optimal choices are related for a simple 2-state example. The top graph shows the conventional representation in contingent commodity space and the two lower graphs the representations in terms of market wealth and D($\boldsymbol{p}||\boldsymbol{q}$)-surplus as well as in terms of subjectively expected wealth and D($\boldsymbol{q}||\boldsymbol{p}$)-risk respectively. The dashed line represents the budget line in all three graphs, and the points *m*,*a*,*b* on the budget lines represent the same contingent wealth risks in all graphs. The budget line representations in the lower graphs are obtained by calculating the supporting prices $\boldsymbol{p_z}$ for the agent for each point $\boldsymbol{z}$ on the budget line $\{\boldsymbol{z}:\boldsymbol{p}\bullet\boldsymbol{z}=\boldsymbol{p}\bullet\boldsymbol{\omega}\}$ as an optimal choice, (see the appendix for details), than calculating the corresponding levels of market wealth and standardized surplus, $\boldsymbol{p_z}\bullet\boldsymbol{z}$ and D($\boldsymbol{p_z}||\boldsymbol{q}$), and expected wealth and standardised risk, $\boldsymbol{q}\bullet\boldsymbol{z}$, and D($\boldsymbol{q}||\boldsymbol{p_z}$).

Figure 1



The expenditure function (1.3) shows that the more risk tolerant an agent is or the greater is the KL discrepancy D(***p***||***q***) between the agent's beliefs ***q*** and market prices ***p***, the cheaper it is for her to achieve any given level of expected utility. Hence D($\boldsymbol{p}||\boldsymbol{q}$), the market's expectation, and valuation, of its differential L-predictive ability or the weights of evidence in its favour , $\boldsymbol{Ln}\frac{\boldsymbol{p}}{\boldsymbol{q}}$, determines a standardized *cost savings* available to a CARA agent with unit risk tolerance and beliefs $\boldsymbol{q}$ trading at prices $\boldsymbol{p}$[4] .

These cost savings are experienced by the agent as economic surplus or gains from trade. Let indirect utility be eu_t =EU($\boldsymbol{t},\omega$) at prices $\boldsymbol{t}$ and eu_p =EU($\boldsymbol{p},\omega$) at prices $\boldsymbol{p}$. The Hicksian *equivalent* variation $\mathrm{EV}_{t\to p}$ in non contingent income for a price change from $\boldsymbol{t}$ to $\boldsymbol{p}$ is:

1.11•	$\mathrm{EV}_{t\to p}$	$=\mathrm{m}(\boldsymbol{t},eu_p)-\mathrm{m}(\boldsymbol{t},eu_t)$	by definition of EV
		$=\mathrm{CE}(\mathrm{eu}_p)-\mathrm{CE}(\mathrm{eu}_t)$	from 1.3
		$=(\boldsymbol{p}-\boldsymbol{t})\bullet\omega+\tau\{\mathrm{D}(\boldsymbol{p}\|\|\boldsymbol{q})-\mathrm{D}(\boldsymbol{t}\|\|\boldsymbol{q})\}$	from 1.7

The corresponding Hicksian *compensating* variation $\mathrm{CV}_{t\to p}$ is $\mathrm{m}(\boldsymbol{p},\mathrm{eu}_p)-\mathrm{m}(\boldsymbol{p},\mathrm{eu}_t)=\mathrm{CE}(\mathrm{eu}_p)-\mathrm{CE}(\mathrm{eu}_t)=(\boldsymbol{p}-\boldsymbol{t})\bullet\omega+\tau\{\mathrm{D}(\boldsymbol{p}||\boldsymbol{q})-\mathrm{D}(\boldsymbol{t}||\boldsymbol{q})\}$. As expected from the Gorman normal form of the certainty equivalent function, $\mathrm{EV}_{t\to p}=\mathrm{CV}_{t\to p}$.

Setting the price vector $\boldsymbol{t}$ in 1.11 to the agent's subjective probabilities $\boldsymbol{q}$, "fair odds" prices for this agent, the constant n-tuple of contingent wealth $\boldsymbol{q}\bullet\omega$ is supported as an optimal choice, D($\boldsymbol{q}||\boldsymbol{q}$)=0, and the EV/CV of 1.11 simplifies to:

1.12• $\mathrm{EV}_{q\to p}=\mathrm{CV}_{q\to p}=(\boldsymbol{p}-\boldsymbol{q})\bullet\omega+\tau\mathrm{D}(\boldsymbol{p}||\boldsymbol{q})$

or, reversing the direction of the price change, $\mathrm{EV}_{p\to q}=\mathrm{CV}_{p\to q}=(\boldsymbol{q}-\boldsymbol{p})\bullet\omega-\tau\mathrm{D}(\boldsymbol{p}||\boldsymbol{q})$

The endowment effect in 1.12, $(\boldsymbol{p}-\boldsymbol{q})\bullet\omega$, may be positive or negative in sign. It can be isolated conceptually by imagining an agent with extremely low risk tolerance, $\tau\to 0$, so that surplus from trade τD($\boldsymbol{p}||\boldsymbol{q}$) becomes vanishingly small. By 1.2, such an agent will optimally purchase nearly certain wealth equal to the market value of her endowment in all states, at any price. A change *to* fair odds pricing $\boldsymbol{q}$ from $\boldsymbol{p}$ in this limiting case has no effect on her risk taking *behaviour*. It simply brings her a windfall loss or gain depending on whether the market value of her intial endowment of risk $\boldsymbol{\omega}$ decreases, $\boldsymbol{q}\bullet\boldsymbol{\omega}<\boldsymbol{p}\bullet\boldsymbol{\omega}$, or increases, $\boldsymbol{q}\bullet\boldsymbol{\omega}>\boldsymbol{p}\bullet\boldsymbol{\omega}$.

If the initial endowment is a certainty, $\omega_s=\omega$ for all states, 1.12 reduces to $EV_{p\to q}=CV_{p\to q}=\tau D(\boldsymbol{p}||\boldsymbol{q})$. Quite apart from the statistical interpretations of section 2, KL discrepancy has a sound economic interpretation. It measures a *standardized* (τ=1, no wealth effects) Hicksian surplus or gains from trade in at prices $\boldsymbol{p}$ for a CARA agent with beliefs $\boldsymbol{q}$. The greater the KL discrepancy between market prices and beliefs, the greater the gains from trade. 1.12 indicates that these economic surpluses are lost completely in a change *to* fair odds pricing. Even a partial move towards fair odds pricing reduces these surpluses, in proportion to the change in KL discrepancy (1.11).

The notion that movements towards fair odds pricing in contingent claims markets create a *loss* that risk averse agents are willing to pay to avoid appears to run contrary to conventional intuition about the benefits of insurance. After all, a simple application of Jensen's inequality shows that the direct Pratt-Arrow risk premium, the difference between subjectively expected wealth and the certainty equivalent wealth, is positive for risk averse agents. Doesn't that indicate a willingness to pay to ***avoid*** risk?

Yes, but.... The direct Pratt-Arrow risk premium calculated according to Jensen's inequality provides a precise answer to a precise question about benefits of risk sharing for risk averse agents. But the particular valuation question being asked (how much wealth-for-sure compensates for a change in *quantities* from contingent wealth $\boldsymbol{z}$ to non contingent wealth $\boldsymbol{q}\bullet\boldsymbol{z}$?) is not the same question being asked when one is looking for Hicksian CV/EV valuations of *price* changes (how much wealth-for-sure compensates for a change in *prices* from $\boldsymbol{p}$ to $\boldsymbol{q}$?). The latter question assumes optimal behaviour subject to a budget constraint *before and after* a price change, while the former question does not.

In elementary price theory we teach that holding income fixed, changing relative prices may either help or hurt depending on the new substitution possibilities opened up compared to the old ones closed off. For CARA agents, holding market wealth fixed, D(***p***||***q***)-reducing price changes always hurt since the surplus from trade in wealth risks decreases. However, if the initial endowment ω happens to be supported as an optimal choice $\boldsymbol{z^o}$, fair odds pricing always has net benefical endowment effect that will more than outweigh the lost consumer surplus (expected wealth $\boldsymbol{q \bullet z^o} = \boldsymbol{q \bullet \omega}$ equals market wealth at new prices ***q*** which exceeds the minimum income at new prices to be as well off as initially, $\boldsymbol{p \bullet \omega} + \tau$D(***p***||***q***). When they occur, gains from fair odds pricing are the result of a tradeoff between endowment and substitution effects of price changes. Arguments based on Jensen's inequality and convex preferences alone mask this tradeoff.

Arguments for portfolio diversification provide an even more striking example of how important economic tradeoffs are obscured by simple applications of Jensen's inequality. Jensen's inequality implies that a diversified mixture of two indifferent wealth risks will be preferred to either one of them for an agent with convex preferences. But revealed preference implies that if either one of the indifferent risks cited in the argument happens to be an optimal choice in some budget set ($\boldsymbol{p,\omega}$) the mixture (diversified) option won't be *affordable* in that budget set, so the comparison is moot. Moreover, revealed preference also says that scaled down versions of the diversified option that are affordable in that budget set won't be *optimal* ! How much diversification is optimal? By 1.2 the optimal extent of diversification, like all other rational choices in a market, depends *both* on relative prices and on substitution possibilities in preference.

Suppose that the set of states can be described by two logically independent discrete quantities S and T with n_s and n_t distinct possible values, $s \epsilon \{s_1, ... s_{n_s}\}$ and $t \epsilon \{t_1, ... t_{n_t}\}$ respectively. There are $n_s n_t$ joint events of the form ((S,T)=(s,t)) with prices and beliefs ***p(s,t)*** and ***q(s,t)***. Trade in units of wealth conditional only on marginal S-events (S=s)=Σ_t((S,T)=(s,t)) or only on marginal T-events (T=t)=Σ_s((S,T)=(s,t)) may occur, at corresponding prices ***p(s)*** $=\Sigma_t\, p(s,t)$ and ***p(t)***$=\Sigma_s\, p(s,t)$ and with beliefs ***q(s)***$=\Sigma_t\, q(s,t)$ and ***q(t)***$=\Sigma_t\, q(s,t)$ respectively for these marginal events. KL discrepancies in contingent claims markets for joint events ((S,T)=(s,t)) can be neatly decomposed into sums of KL discrepancies for wealth contingent on marginal S-events, (S=s), and marginal T-events, (T=t), plus a residual:

1.13 $D(\boldsymbol{p(s,t)}\|\boldsymbol{q(s,t)})=D(\boldsymbol{p(s)}\|\boldsymbol{q(s)})+D(\boldsymbol{p(t)}\|\boldsymbol{q(t)})+\sum_s\sum_t p(s,t)\{\mathrm{Ln}(\frac{p(s,t)}{p(s)\,p(t)})-\mathrm{Ln}(\frac{q(s,t)}{q(s)\,q(t)})\}$

1.14 $D(\boldsymbol{q(s,t)}\|\boldsymbol{p(s,t)})= D(\boldsymbol{q(s)}\|\boldsymbol{p(s)})+D(\boldsymbol{q(t)}\|\boldsymbol{p(t)})+\sum_s\sum_t q(s,t)\{\mathrm{Ln}(\frac{q(s,t)}{q(s)\,q(t)})-\mathrm{Ln}(\frac{p(s,t)}{p(s)\,p(t)})\}$

The valuation formulae 1.13 and 1.14 have clear parallels with general principles about optimal portfolio choice based on decompositions of variance: variance of a sum is sum of variances plus twice the covariance. Systematic investigation of these interrelationships and their implications for financial market analysis requires more space than we have here, but we note an important idea related to the optimal extent of diversification. If both the market and the agent regard S and T as stochastically independent, ***p(s,t)=p(s)p(t)*** and ***q(s,t)=q(s)q(t)*** and the double sums in 1.13 and 1.14 vanish. In this case the surplus (1.13) and risk premia (1.14) available to a CARA agent from trading in $n_s n_t$ markets in (S,T) contingent wealth *jointly* is equal to the sum of the surpluses and risk premia available trading in $n_s + n_t$ markets in S=*s* and T=*t* events *marginally*. Surplus value and risk premia are additive over marginal assets S and T when independence is presumed on both sides of a market transaction. From a consumer's standpoint, further diversification into cross product events other than simple bets on S and on T simply isn't optimal, even though it may be possible.

Where does variance of wealth enter this analysis? It doesn't, at least not directly. Since an optimal choice in 1.2 is a linear function of log score differences, *variance of optimally chosen wealth* as assessed from either the agent's or the market's beliefs is proportional to *variance of Log score differences* or to *variance in the weight of evidence.* Neither the indirect certainty equivalent (1.7) nor the optimal expected wealth (1.8) nor the indirect risk premium (1.10) depend directly on second (or higher order) moments of the distributions of log score differences for CARA agents.

Minimum variance portfolios may be used to *approximate* optimal choices and value functions for CARA agents. For example, Appendix D shows that a minimum variance portfolio chosen subject to constraints on expected wealth and affordability has differences in wealth between pairs of states *s* and *t* proportional to $\{\frac{p(t)}{q(t)} - \frac{p(s)}{q(s)}\}$. This expression is a first order Taylor series approximation to {Ln($\frac{q(s)}{p(s)}$)-Ln($\frac{q(t)}{p(t)}$)}, the difference in optimal wealth between states *s* and *t* (1.2). But why approximate when an exact expression is available? Or, on a deeper note, why use quadratic approximations to utility of wealth functions when CARA approximations may be superior (Brockett and Golden (1985))?

We have established that KL discrepancy is a natural measure of risk and of surplus for CARA agents trading in contingent claims markets. But can KL discrepancy be used to measure risk and surplus for non CARA agents? Yes, if we are dealing with *proper* utility functions as characterized by Pratt and Zeckhauser (1987) and Brockett and Golden (1985). Intuitively, a non increasingly risk averse utility of wealth function is proper whenever an agent who finds each of two independent risks unattractive taken one at a time, will also find these risks unattractive in combination, starting from either a certain or an uncertain endowment of risk. More technically, proper utility functions are smooth utility functions with alternating positive (odd) and negative (even) high order derivatives. Brockett and Golden (1985,pp 958, 963-964) show that all the common utility functions (linear, exponential, logarithmic[5], isoleastic) are proper and that proper utility functions can be represented exactly by mixtures of CARA utility of wealth functions and approximately by finite mixtures

These results are important for our analysis because the supporting prices and expenditure function (and therefore all other relevant economic variables) for an agent whose utility of wealth is a finite mixture of CARA utility functions are easy to compute and to interpret in terms of the supporting prices and expenditure functions for the implicit component CARA agents. Consider an EU function for a *mixture agent* defined as a weighted sum of two CARA agents with Pratt-Arrow risk *aversion* indices r_a and r_b for two positive constants α and β: $\mathrm{E}U_{ab}(r_a\, r_b, \boldsymbol{q}, \boldsymbol{z}) \equiv \sum_s q(s)\,(-\alpha\; e^{-r_a\, z(s)} - \beta\; e^{-r_b\, z(s)})$. Supporting prices $\boldsymbol{p_{ab}}$ for a risk $\boldsymbol{z}$ viewed as an optimal choice for the mixture agent are a weighted average of the supporting prices $\boldsymbol{p_a}$ and $\boldsymbol{p_b}$ for $\boldsymbol{z}$ as an optimal choice for each component agent, $p_{ab}(s) = \eta_a\, p_a(s) + \eta_b\, p_b(s)$, where weights are $\eta_i = \frac{r_i\, eu_i(z)}{r_a\, eu_a(z) + r_b\, eu_b(z)}$ and $eu_i(\mathrm{z}) = \mathrm{EU}(r_i, \boldsymbol{q}, \boldsymbol{z})$, *i*=,*a*,*b*. Weights can be calculated indirectly by finding the supporting prices and associated indirect certainty equivalent $ce_i(\boldsymbol{z})$ is (1.7) for the component agent i facing budget set $(\boldsymbol{p_i}, \boldsymbol{p_i} \cdot \mathbf{z}), i=a,b$. The corresponding expenditure function for the mixture agent is a weighted average of the expenditure function for each component agent:

1.15• $m(\boldsymbol{p_{ab}}, eu_{ab}, \boldsymbol{q}, r_a, r_b) = \eta_a\, m(\boldsymbol{p_a}, eu_a^o, \boldsymbol{q}, r_a) + \eta_b\, m(\boldsymbol{p_b}, eu_b^o, \boldsymbol{q}, r_b)$

$$= \eta_a\, ce_a(\mathbf{z}) + \eta_b\, ce_b(\mathbf{z}) - [\eta_a\, \tau_a\, D(\boldsymbol{p_a} \,||\, \boldsymbol{q}) + \eta_b\, \tau_b\, D(\boldsymbol{p_b} \,||\, \boldsymbol{q})].$$

Section III Applications

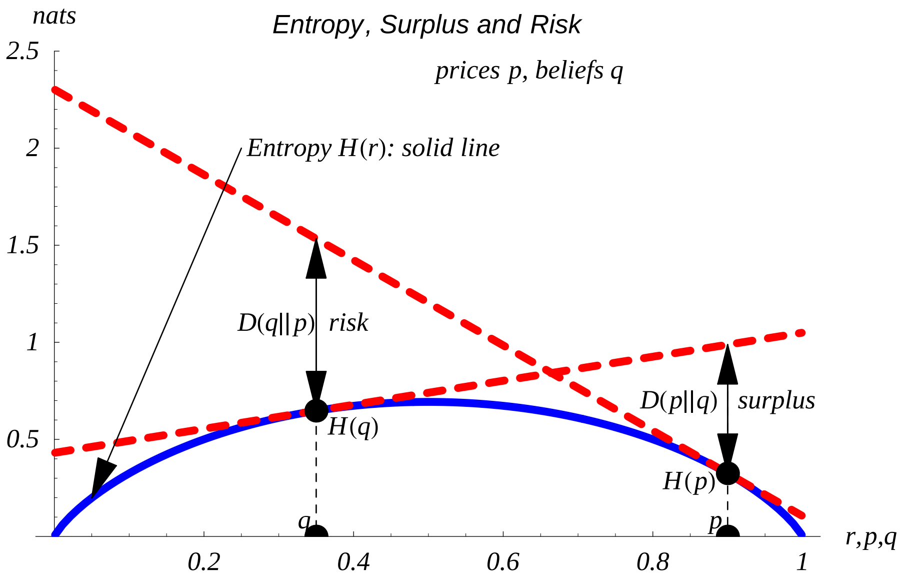
This section offers three suggestive applications of KL discrepancy measures of risk and surplus. The first analyzes the economic relevance of Shannon entropy as a measure of uncertainty and value of information. The second shows how KL discrepancy provides insights into the analysis of the costs, benefits, and risks, of trading in *incomplete* contingent claims markets. The third shows how KL discrepancy can be used to simplify the analysis of efficient risk sharing arrangements among heterogeneous agents by characterizing efficiency in terms of prices rather than quantities and by using the Gorman polar form of the certainty equivalent function to rationalise representative agent analysis.

Section III.A Entropy, risk, and surplus

Entropy based concepts have long been used by philosophers, statisticians and information theorists as measures of uncertainty (Jessop(1995),Cover and Thomas(1991), Bernardo and Smith(1994), Lad(1996)). The relevant uncertainty in these concepts, however, is formulated in terms of *variability in knowledge* expressed in probability distributions rather than in terms of *variability of wealth* or other objects of contingent consumption. Moreover, Arrow's (1985) analysis excepted, the analysis of optimal choice in contingent claims markets typically plays no explicit role in the interpretation of entropy as a measure of uncertainty. This section explains when entropy does, and does not, measure relevant uncertainty about trading in contingent claims markets for CARA agents.

Figure 2 provides a general insight into the quantitative relationships between Shannon entropies *H*(***p***) and *H*(***q***) and relative entropy, KL discrepancies D(***q***||***p***) and D(***p***||***q***). The graph is based on the fact that for any two pmfs ***a*** and ***b***, D(***a***||***b***)=*H*(***b***)+∇*H*(***b***)•(***a***-***b***)–*H*(***a***) where ∇*H*(***b***) is the vector of partial derivatives of the entropy function H(•) evaluated at ***b***. Hence, D(***r***||***b***), as a function of ***r***, is the difference between the linear support to the concave function *H*(•) at ***b***, *H*(***b***)+∇*H*(***b***)•(***r***-***b***), and *H*(***b***). As is evident from Figure 2, neither the *levels* nor the *differences* in simple entropies or *self referential* L-predictive abilities *H*(***p***) and *H*(***q***) are accurate indicators either standardized surplus D(***p***||***q***) or standardized risk D(***q***||***p***) when a CARA agent with beliefs ***q*** trades in a market with prices ***p***. The geometry reinforces a central message of this paper, that relative, not absolute differences in knowledge drive risks and benefits from trade in contingent claims markets. There are however two special cases.

Figure 2 here



Suppose first that market prices, ***p***, are uniform, denoted here by ***n***. Hence ***p*** in Figure 2 will be located at r=0.5 with corresponding maximum entropy point H(0.5) and the tangent line *H*(***p***)+∇*H*(***p***)•(***r***-***p***) at (0.5,H(0.5)) used to determine standardized risk D(***q***||***p***) will be horizontal. In this case the difference between maximum entropy *H*(***n***) and entropy *H*(***q***) in beliefs ***q*** measures the standardized Pratt-Arrow risk premium D(***q***||***n***) for CARA agents trading in contingent claims markets with uniform prices. When the agent becomes self referentially more knowledgeable in the sense of being better able to L-predict states , relative risk D(***q***||***n***) increases, as her beliefs are also getting farther away from the minimal self referential knowledge expressed in the uniform prices of her trading partner(s). It is also true that self referentially more informed traders obtain increased benefits in the form of larger standardized surpluses D(***n***||***q***) as beliefs ***q*** get farther away from uniform prices ***n***. To see this, imagine ***q*** in Figure 2 approaching r=0 so the tangent line *H*(***q***)+***∇****H*(***q***)•(***r***-***q***) at (***q***,H(***q***)) used to determine standardized surplus D(***n***||***q***) has a steeper slope, with an abscissa at r=0.5 farther away from H(0.5).

On the other hand, suppose the agent, rather than the market, has minimal self referential knowledge with uniform beliefs ***q***=***n***. Hence ***q*** in Figure 2 will be located at r=0.5 with corresponding maximum entropy point H(0.5), and the tangent line *H*(***q***)+***∇****H*(***q***)•(***r***-***q***) at (0.5,H(0.5)) used to determine standardized surplus D(***p***||***q***) will be horizontal. Ignoring endowment effects from price changes, the difference in self referential knowledge or simple entropies between the market and the agent , *H*(***n***)-*H*(***p***), *is fully captured as standardized surplus*, since D(***p***||***n***)=*H*(***n***)-*H*(***p***). But higher surplus D(***p***||***n***) as prices ***p*** get farther away from beliefs ***n*** may come at the cost of potentially large standardized risks D(***n***||***p***).[6] To see this, imagine ***p*** in Figure 2 approaching r=1 so the tangent line *H*(***p***)+***∇****H*(***p***)•(***r***-***p***) at (***p***,H(***p***)) used to determine standardized risk premium D(***n***||***p***) has a steeper slope, with an abscissa at r=0.5 farther away from H(0.5).

Section III.B *Risk bearing in incomplete markets*

This section develops KL discrepancy measures of gains from trade and risk premia in *incomplete* contingent claims markets. Contingent wealth markets can be incomplete in a number of ways and for a variety of reasons. The incompleteness examined here is for trades occurring on the basis of wealth contingent on *macro* state events rather than on the basis of wealth contingent on *micro* state events.

Suppose T is a function of S, T=g(S), with $n_t \leq n_s$ distinct possible values $\{t_1,...t_{n_t}\}$. Market prices ***p(s,t)*** and agent beliefs ***q(s,t)*** imply cohering prices and beliefs ***p(t)*** and ***q(t)*** on constituent events for T, $p(t) = \sum_{s:g(s)=t} p(s)$ and $q(t) = \sum_{s:g(s)=t} q(s)$. We call T a *macro state variable* and S a *micro state variable* since T-events are aggregates of S-events: (T=t)=$\sum_{s:g(s)=t} (S = s)$, where we are using our indicator function notation that an expression with brackets ("A") is the logical truth value, 0 or 1, of the enclosed proposition "A". Let m_t equal the number of states s in the level set defined by g(s)=t. Applying the Log Sum Inequality, $\sum_{i=1}^{m} p_i \log \frac{p_i}{q_i} \geq (\sum_{i=1}^{m} p_i) \log \frac{\sum_{i=1}^{m} p_i}{\sum_{i=1}^{m} q_i}$, (see appendix B) to the m_t prices *p(s)* and beliefs *q(s)* associated with each of the level sets g(s)=t in S, then summing the resulting inequalities across the n_t values of T, we find D(***p(s)***||***q(s)***)≥D(***p(t)***||***q(t)***). Analogously, D(***q(s)***||***p(s)***)≥D(***q(t)***||***p(t)***. Surplus value and relative risk both tend to decrease, and never increase, with aggregation. Hence, *dis*aggregation in contingent claims markets is a potential source of value, but also a potential source of risk for CARA agents.

Surplus D(***p(s)***||***q(s)***) from trading in wealth contingent on micro states S will *equal* surplus D(***p(t)***||***q(t)***) from trading in wealth contingent on macro states T when the agent and the market *share coherent conditional expectations* about the state variable S given T, ie p(*s*|*t*)=q(*s*|*t*) for every sϵS and tϵT. This proposition is an interesting application of the Chain Rule for relative entropy on joint state spaces S$\otimes$T, $s\epsilon S=\{s_1,...s_{n_S}\}$, $t\epsilon T=\{t_1,...t_{n_t}\}$

$$D(\boldsymbol{a(s,t)}||\boldsymbol{b(s,t)}) = D(\boldsymbol{a(t)}||\boldsymbol{b(t)}) + \sum_t a(t) D(\boldsymbol{a(s|t)}||\boldsymbol{b(s|t)})$$
$$= D(\boldsymbol{a(s)}||\boldsymbol{b(s)}) + \sum_s a(s) D(\boldsymbol{a(t|s)}||\boldsymbol{b(t|s)})$$

where ***f(•|•)*** denotes a conditional pmf derived from ***f(s,t)*** (see Appendix C). First, expand around the marginal for T, $D(\boldsymbol{p(s,t)}||\boldsymbol{q(s,t)})=D(\boldsymbol{p(t)}||\boldsymbol{q(t)})+\sum_t p(t)D(\boldsymbol{p(s|t)}||\boldsymbol{q(s|t)})=D(\boldsymbol{p(t)}||\boldsymbol{q(t)})$, with the last equality holding because every KL discrepancy D(***p(s|t)***||***q(s|t)*** between identical conditional pmfs p(*s*|*t*)=q(*s*|*t*) is zero (KL1). Using the Chain Rule again expand around the marginal for S, $D(\boldsymbol{p(s,t)}||\boldsymbol{q(s,t)})=D(\boldsymbol{p(s)}||\boldsymbol{q(s)})+\sum_s p(s)D(\boldsymbol{p(t|s)}||\boldsymbol{q(t|s)})=D(\boldsymbol{p(s)}||\boldsymbol{q(s)})$, with the last equality holding because the pmfs ***p(t|s)***|=***q(t|s)*** are identical with all probability mass on *t*=g(*s*) since T is a known function of S.

Analogously indirect risk premia D(***q(s)***||***p(s)***) from trading in wealth contingent on *micro* states S will equal indirect risk premia D(***q(t)***||***p(t)***) from trading in wealth contingent on *macro* states T under the same condition on inference strategies.

As a special case, imagine S consists of $n_S=2^k$ finite *k*-tuples of ones and zeroes and T is the summary statistic of the number of ones in a sequence, $t\epsilon\{0,1,2,....k\}$, with only $k+1< 2^k$ possible values. Suppose both the agent and the market regard the sequences in S finitely exchangeably, so that everyone believes any sequence *s*ϵS with *t* ones has the same probability as any other sequence with *t* ones, for all T=*t*. Then both the market and the agent must have identical hypergeometric conditional distributions ***q(s|t)***, ***p(s|t)*** (Lad(1996)). The Chain rule(KL3) then implies that standardized surplus in joint markets for (S,T) contingent claims equals the standardized surplus in the macro state market for T-contingent claims, D(***p(s,t)***||***q(s,t)***)= D(***p(t)***||***q(t)***. Similarly the standardized risk premium in joint markets for (S,T) contingent claims equals the standardized risk premium in the macro state market for T-contingent claims , D($\boldsymbol{q(s,t)}\|\boldsymbol{p(s,t)}$)=D($\boldsymbol{q(t)}\|\boldsymbol{p(t)}$). Hence trading in "incomplete" markets contingent on $n+1$ macro events (T=t) at marginal prices $\boldsymbol{p(t)}$ with marginal beliefs $\boldsymbol{q(t)}$ creates the same surplus value and relative risk characteristics as trade in the complete markets on underlying micro states. These ideas apply to trade in wealth contingent on sufficient statistics generally, rather than on the complete set of micro states underlying sufficient statistics.

Section III.C *Efficient risk sharing*

This section uses the exact aggregation properties of the certainty equivalent function (equation 1.7 is in Gorman normal form) to analyze efficient risk sharing arrangements for a group of CARA agents with diverse beliefs, risk tolerances, and endowments.

Consider a model with N=3 CARA agents from a population $POP\equiv\{$Q,R,P$\}$ whose beliefs are described by pmfs $\boldsymbol{r(s)}$, $\boldsymbol{q(s)}$ and $\boldsymbol{p(s)}$ respectively on constituent events (S=s), risk tolerances are τ_i, and endowment vectors are $\omega_i = (\omega_i(s), \omega_i(t))$, for each $i\epsilon POP$. Aggregate contingent wealth in state s is $W(s)=\Sigma_i\, \omega_i(s)$, with corresponding per capita wealth $\overline{W}(s)=\frac{W(s)}{N}$. Feasibility constraints $\Sigma_i\, \boldsymbol{z_i}(\boldsymbol{s})=\boldsymbol{W(s)}$ must be satisfied for non-negative allocation vectors $(\boldsymbol{z_r}(\boldsymbol{s}), \boldsymbol{z_q}(\boldsymbol{s}), \boldsymbol{z_p}(\boldsymbol{s}))$. Define *total risk tolerance for the group* as $\tau=\Sigma_i\, \tau_i$, *relative* risk tolerances $\frac{\tau_i}{\tau}$ for each agent i, and per capita risk tolerance $\overline{\tau}=\frac{\tau}{N}$.

Aggregate wealth risks $\boldsymbol{W(s)}$ are shared efficiently between members of the group when marginal rates of substitution between contingent wealth claims in pairs of states s,t are equalized for all agents (at an interior solution). This common marginal rate of substitution or *equilibrium odds* $\frac{\pi(s)}{\pi(t)}$, in Log-odds form is (detailed calculations available from author on request) :

$$1.17 \qquad \text{Ln}\,\frac{\pi(s)}{\pi(t)} = \frac{\tau_r}{\tau}\,\text{Ln}\,\frac{r(s)}{r(t)} + \frac{\tau_q}{\tau}\,\text{Ln}\,\frac{q(s)}{q(t)} + \frac{\tau_p}{\tau}\,\text{Ln}\,\frac{p(s)}{p(t)} - \frac{1}{\tau}\,[W(s)-W(t)]$$

Efficient allocations of wealth risks are those that would be demanded in a competitive market in state contingent claims (1.2) with equilibrium Log-odds prices given by equation1.17. For each agent these efficient variations in wealth are:

1.18q $z_Q(s)\text{-}z_Q(t) = \frac{\tau_Q}{\tau}\,[\tau_R\,(\text{Ln}\frac{q(s)}{q(t)} - \text{Ln}\frac{r(s)}{r(t)}) + \tau_P\,(\text{Ln}\frac{q(s)}{q(t)} - \text{Ln}\frac{p(s)}{p(t)}) + W(s)\text{-}W(t)]$

1.18r $z_R(s)\text{-}z_R(t) = \frac{\tau_R}{\tau}\,[\tau_Q\,(\text{Ln}\frac{r(s)}{r(t)} - \text{Ln}\frac{q(s)}{q(t)}) + \tau_P\,(\text{Ln}\frac{r(s)}{r(t)} - \text{Ln}\frac{p(s)}{p(t)}) + W(s)\text{-}W(t)]$

1.18p $z_P(s)\text{-}z_P(t) = \frac{\tau_P}{\tau}\,[\tau_Q\,(\text{Ln}\frac{p(s)}{p(t)} - \text{Ln}\frac{q(s)}{q(t)}) + \tau_R\,(\text{Ln}\frac{p(s)}{p(t)} - \text{Ln}\frac{r(s)}{r(t)}) + W(s)\text{-}W(t)]$

If everyone shares the same beliefs, the expressions involving Logs in 1.18 vanish, implying that more risk tolerant agents bear proportionately more aggregate wealth variation across states *s* and *t* in an efficient allocation. An allocation rule that shares aggregate wealth proportionately, $\frac{\tau_i}{\tau}\,W(s)$, as derived using Taylor series approximation methods in Paul Milgrom and John Roberts (1992,p 213), is sufficent for efficiency in this case, but not necessary. Only *differences* in aggregate wealth across states, not *levels*, need be shared proportionately for efficiency between CARA agents with homogeneous beliefs. There are a variety of efficient allocations of wealth *levels* to individuals that will satisfy the feasibility and non-negativity constraints and the patterns of wealth *differences* across states described in 1.18.

1.18 indicates that L-predictive abilities, Log-Bayes factors, or weights of evidence, matter for efficient allocation of risk. Consider first predictive abilities as assessed by the Log scoring rule. A term like $\frac{\tau_Q\,\tau_R}{\tau}\,(\text{Ln}\frac{q(s)}{q(t)} - \text{Ln}\frac{r(s)}{r(t)})$ in 1.18q compares Q's differential L-predictive ability for state *s* relative to state *t*, $\text{Ln}\frac{q(s)}{q(t)}$, to R's, $\text{Ln}\frac{r(s)}{r(t)}$. If the difference in Q's Log scores predicting state *s* relative to state *t* is larger than than R's, Q should bear more wealth risk in *s* than in *t* and R correspondingly less (1.18r), The absolute predictive abilities of Q and R aren't relevant here. Only relative L-predictive abilities as assessed by the Log scoring rule matter, but they do matter. *Efficient allocations reflect pairwise differences in L-predictive* abilities *between agents.*

Another way to interpret this idea is in terms of weights of evidence. *Efficient allocations follow the weight of the evidence for and against the competing hypotheses of diverse agents in a risk sharing group.* Rearranging $\frac{\tau_Q\,\tau_R}{\tau}\,(\text{Ln}\frac{q(s)}{q(t)} - \text{Ln}\frac{r(s)}{r(t)})$ in 1.18q to $\frac{\tau_Q\,\tau_R}{\tau}\,(\text{Ln}\frac{q(s)}{r(s)} - \text{Ln}\frac{q(t)}{r(t)})$, one can see that if *ex post* the weight of the evidence $\text{Ln}\frac{q(s)}{r(s)}$ in favour (±) of Q's hypothesis about *s* over R's exceeds the weight of the evidence $\text{Ln}\frac{q(t)}{r(t)}$ in favour (±) of Q's hypothesis about *t* over over R's, Q should bear more wealth risk in *s* than in *t* and R correspondingly less (1.18r).

Of course with multiple hypotheses, weights of evidence or differences in L-predictive abilities computed pairwise between the competing hypotheses held by diverse agents will vary and possibly have to be combined or traded off against one another. 1.18 indicates that those tradeoff rates should occur at rates reflecting relative risk tolerances. For example, using 1.18q, if the weight of evidence in *s* relative to *t* favours Q over the R, so $(\text{Ln}\frac{q(s)}{q(t)} - \text{Ln}\frac{r(s)}{r(t)}) > 0$, but P over Q, so $(\text{Ln}\frac{q(s)}{q(t)} - \text{Ln}\frac{p(s)}{p(t)}) < 0$, the differences should be combined linearly using weights equal to the risk tolerances τ_R and τ_P respectively and traded off dollar for dollar with aggregate wealth differences.

When beliefs differ, Pareto efficiency demands that agents "put their money where their mouth is". This conclusion will no doubt gladden the hearts of sports betting bookies and their clients everywhere. It also buttresses critics of "scientific management" by requiring, on efficiency grounds, that expert economists, policy analysts, and statisticians back up their policy forecasts by putting their own wealth at stake, in amounts that reflect the *ex post* weights of evidence for or against their hypotheses. Standing on the sidelines sniping at the differing beliefs or predictive hypotheses of others and being part of an efficient risk sharing group don't mix.

Note also that the *source* of individual differences in predictive hypotheses about the publicly observable and verifable state S is no more a concern of efficiency than are the psychological or social sources of differing risk attitudes. One agent might have volumes of historical data deemed relevant to predicting S as well as sophisticated computers for processing this information via regressions and sensitivity analysis, while another agent might rely on a completely different, looser set of cues and their own personal hunches. These differing technologies and information endowments for producing predictive knowledge no doubt are associated with different actual levels of wealth risks experienced, gross and net. But these distributional considerations are irrelevant for efficient risk sharing between CARA agents. Predictive skill only matters here through the final predictive distributions agents hold, not how they were formed from various component parts.

Equilibrium Log-odds in 1.17 can be rearranged into a useful general form by defining a numeraire state (S=1) and expressing the beliefs of agent i in terms of odds $o_i(s)$ relative to this numeraire state, eg for i=R, $o_R(s)=\frac{r(s)}{r(1)}$. The terms involving individual agent Log-odds in 1.17 can be combined to form a *weighted* geometric mean $\Pi_i^N [o_i(s)]^{\frac{\tau_i}{\tau}}$ of the agents' odds. In turn, the Log of this geometric mean can be expressed as a *simple* geometric mean $\overline{o}(s)=\Pi_i [o_i(s)]^{1/n}$ plus a covariance, COV($\frac{\tau_i}{\tau}$,Ln $o_i(s)$), the simple covariance across individuals in the population between relative risk tolerances and their Log odds. Setting population numbers at N, equilibrium Log-odds 1.17 can be written in a compact general form as:

1.19 $$\text{Ln}\,\frac{\pi(s)}{\pi(1)} = \text{Ln}\,\overline{o}(s) + N\,\text{COV}\left(\frac{\tau_i}{\tau}, \text{Ln}\,o_i(s)\right) - \frac{1}{\tau}\,[W(\text{s}) - W(1)]$$

The general equilibrium pricing formula 1.19 has many potentially intriguing applications in economics, only some of which are familiar. For example, we have the standard implication that, other things equal, when aggregate wealth in a specific state $W(s)$ is in excess supply relative to wealth in state t, units of wealth in state s will have a relatively lower equilibrium price while units of wealth in state t, scarce in supply, will rise (Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green (1995, p 693)). Consider next a homogeneous risk sharing club case where all agents have identical beliefs Ln$\overline{o}(s)$ and identical risk tolerances $\overline{\tau}$. Then the equilibrium Log-odds in 1.19 reduce to the supporting relative prices Ln$\overline{o}(s)-\frac{1}{\overline{\tau}}[\overline{W}(\text{s})-\overline{W}(1)]$ for *Everyman* who shares the common beliefs, has the common average risk tolerance, and is endowed with the group's per capita wealth $\overline{W}$(s) in every state. There are no diversification benefits (against a given set of publicly observed states S) if all identical people have identical endowments since equilibrium prices will support those endowments as optimal choices.

If beliefs and risk attitudes are homogeneous but endowments differ, then an efficiently organised group can realise diversification benefits by trading at equilibrium prices equal to *Everyman*'s supporting prices. The precise magnitude of the benefits of participating in a homogeneous risk sharing club can be calculated indirectly in terms of budget set changes via equation 1.11 using the supporting prices for agents' endowments as an initial price vector and *Everyman*'s prices as the terminal price vector. Some special cases of efficiently organised risk sharing clubs are familiar. With pure risk sharing, one individual, say Q, facing substantial wealth risk variation across states $\omega_Q(s)–\omega_Q(1)$ joins with N-1 others who have no wealth variation. Per capita wealth differences $[\overline{W}(s)-\overline{W}(1)]=\frac{\omega_Q(s)-\omega_Q(1)}{N}$ decline as the number of other individuals increase and Log-odds equilibrium prices $\text{Ln}\overline{o}(s)-\frac{1}{\overline{\overline{\tau}}}[\overline{W}(s)-\overline{W}(1)]$ approach fair odds prices as N increases. Individuals with no wealth risk will be motivated to take on wealth risks by the gains from trade at equilibrium relative prices different to their beliefs. The individual Q with the wealth risk is motivated to engage in such trade by the opportunity of trade at prices close to fair odds prices. In a pure hedging case, sets of individuals can be found with precisely offsetting individual wealth risks in the aggregate. Per capita wealth risks become zero, *Everyman*'s supporting relative prices now are simply his beliefs, fair-odds prices, and all individuals benefit from risk sharing.

Once homogeneity of beliefs and risk attitudes is abandoned, all of these intuitive results become suspect. For example, consider the idea that in states where aggregate wealth W(*s*) is in excess supply equilibrium prices of wealth in that state should be lower. This is a supply side argument and ignores demand side considerations, which 1.19 details. While greater supply W(*s*) depresses equilibrium prices for units of S=s wealth, either higher population mean odds Ln $\overline{o}(s)$ and/or positive covariance in the population between relative risk tolerances and odds on this state, COV($\frac{\tau_i}{\tau}$,Ln $o_i(s)$)>0, are demand side forces increasing relative prices. When agents who are risk tolerant also tend to believe that S=*s* is relatively more likely, unit prices in S=*s* will rise relatively, offsetting and perhaps overturning any tendency for unit prices in S=*s* to fall because of relatively larger aggregate endowments of wealth in that state.

Next, consider pure risk sharing. In the homogeneous risk sharing club case the benefits to the agent Q with the primary source of risk derive from favourable endowment effects associated with the movement of equilibrium prices towards fair odds prices for Q and away from the supporting prices for his risky initial endowment. But with diverse membership, equilibrium relative prices will be an average of the Log-odds beliefs of those in the club, 1.19. As club size N increases, equilibrium prices will approach the average of the Log-odds beliefs of others in the group, not Q's . Moreover, these beliefs if correlated with risk tolerances of new members could turn out to lead to equilibrium prices and associated patterns of efficient risk sharing that are quite surprising for Q. With a group of risk tolerant others who are much more pessimistic about S=*s* than Q, Q could find herself efficiently taking on more, not less, wealth risk in state S=*s*. Even in the case of a pure hedge, if the agent with equal sized but opposite signed risks happens to have different beliefs and risk tolerances than you, perfect certainty, while an option for the both of you, is not efficient.

In general adding other homogeneous members (ie identical risk tolerances and beliefs) to the group may either harm or help incumbents depending on the possibilities for redistributing aggregate wealth so as to make everyone, incumbents and new additions, better off. The crucial test is whether at the new per capita wealth levels *Everyman* is better off or not. It turns out that we don't need homogeneity in beliefs and risk attitudes to use such a representative agent test, since the Gorman polar form of the indirect utility function for CARA agents satisfies sufficient conditions for exact aggregation(Cornes(1992,pp192-194).

Define an indirect utility function for the group by aggregating indirect certainty equivalents for a group of individuals with endowments, beliefs and risk attitudes ($\boldsymbol{\omega_i}, \boldsymbol{q}_i, \tau_i$) facing external contingent claims prices $\boldsymbol{p}$, Σ_i CE($\boldsymbol{p}, \omega_i, \boldsymbol{q}_i, \tau_i$) =$\boldsymbol{p}$•($\Sigma_i$ $\boldsymbol{\omega_i}$) + $\tau\Sigma_i \frac{\tau_i}{\tau}$ D($\boldsymbol{p}||\boldsymbol{q}_i$). The associated aggregate expenditure function for a given distribution of certainty equivalents ce_1 ... ce_N is the sum of the expenditure functions (1.3) for each agent $\Sigma_i\, ce_i$ - $\Sigma_i\, \tau_i$ D($\boldsymbol{p}||q_i$). By 1.4 the aggregate compensated demands for this group are Z(s)=$\Sigma_i\, ce_i$ + $\Sigma_i\, \tau_i \text{Ln}\frac{q_i(s)}{p(s)}$), so that aggregate wealth differences between states *s* and *t* are Z(*s*)–Z(*t*)= $\Sigma_i\, \tau_i\, [\text{Ln}\frac{q_i(s)}{p(s)} - \text{Ln}\frac{q_i(t)}{p(t)}]$. These aggregate wealth variations can be expressed as the differences in demands of a single CARA agent Z(s)–Z(t)=$\tau[\text{Ln}\frac{b(s)}{p(s)} - \text{Ln}\frac{b(t)}{p(t)}]$, where the agent has risk attitude $\tau=\Sigma_i\, \tau_i$ and beliefs ***b(s)*** defined in Log-odds form as $\text{Ln}\frac{b(s)}{b(t)} = \frac{1}{\tau}\Sigma\tau_i \text{Ln}\frac{i(s)}{i(t)}$, and certainty equivalent $\Sigma_i\, ce_i$. When a group of *N* CARA agents with diverse beliefs and risk attitudes is efficiently pooling risks, it behaves *as if* its aggregate demands in external markets are generated by a representative CARA agent whose preferences are described by these indirect utility functions: group risk tolerance is the sum of the risk tolerances of all individuals and beliefs for the group, expressed as odds relative to a numeraire state, are pooled as a weighted geometric mean $\Pi_i^N \, o_i(s)^{\frac{\tau_i}{\tau}}$ of indiviudal agent log-odds beliefs,with weights equal to each agent's relative risk tolerance $\frac{\tau_i}{\tau}$.

D Concluding remarks

Our paper has used duality theory of consumer choice to show that *relative knowledge*, operationally defined as relative predictive ability as indicated by log scoring rules, Log Bayes factors, and the weight of the evidence for and against competing hypotheses, matters decisively for optimisation and for Pareto efficiency in contingent claims markets. Expectations of these useful quantities from Bayesian statistics, in the form of Kullback Leibler discrepancies, are natural measures of risk and economic surplus for agents trading in contingent claims markets. These interesting and analytically tractable concepts have many potentially rich applications. Their useful aggregation properties across states and across individuals with heterogeneous beliefs, as well as their relevance for approximating the measurement of risk and surplus for a wide class of common utility functions, should make them productive additions to the economists' box of analytical tools.

■ Endnotes

* Economics Department, University of Canterbury, PO Box 4800, Christchurch, New Zealand; please address all correspondence to j.fountain@econ.canterbury.ac.nz

1. For example, if $\sum_s p(s)>1$ [<1] an agent with $1 can sell [buy] a unit of wealth in each state to [from] the market for a positive non-contingent profit $\sum_s p(s)-1$, $[1-\sum_s p(s)]$.

2. The general form of the Log scoring rule, $A\bullet\text{Ln}(r(s))+b(s)$ for a reported pmf ***r(s)***={p(1)...p(n)} on n discrete states S with A>0 and *b(s)* some arbitrary function on states (Proposition 2.29, Bernardo and Smith (1994,p. 73)). The Log scoring rule rewards entire pmfs on the basis *only* of the probability assessed for the constituent event that occurs. Neither penalty nor reward is assessed for probabilities for events that might have, but do not, actually occur. Bernardo and Smith(1994,p 72-74) show that the Log scoring rule (modulo positive affine transforms) is the unique, smooth, proper scoring rule with this property and argue that such a rule is especially appropriate for "pure" scientific inference problems.

Scoring rule methods for assessing predictive performance have also been designed (Lad, 1996, Ch 6) for forecasts that are less than full pmfs, eg cohering means, medians, modes, quartiles, etc. Scoring rules can also be viewed normatively as methods of assessing the "goodness" of probability forecasts, rather than as incentive structures to elicit truthful reports.

3. To avoid mis-interpretation,nothing directly is said here about mean values of the state variable S (it may in fact be categorical), nor does a $D(\boldsymbol{p}||\boldsymbol{q})$ spread of prices or a $D(\boldsymbol{q}||\boldsymbol{p})$ spread of beliefs refer directly to Rothschild-Stiglitz mean preserving spreads of probability mass in *wealth* space.

4. These cost savings are bounded above and below. When prices $\boldsymbol{p}$ are equal to beliefs $\boldsymbol{q}$ -so called *fair odds* prices- no costs savings are available as $D(\boldsymbol{q}||\boldsymbol{q})=0$. Standardized cost savings are at a maximum, $\text{Ln}(q(s_*))$ when the market believes that one of the constituent events ranked least likely by the agent, $S=s_*$ where $s_* \epsilon \text{Argmin}_s((q(s))$, is a certainty and prices it and other contingent wealth levels accordingly.

5. Smart money follows the weight of the evidence or Bayes factors for agents with Logarithmic utility as well as for CARA agents. Letting the value of endowment $\boldsymbol{p}\bullet\omega=m$, ordinary demand functions are $z^o(s,\boldsymbol{p},m,\boldsymbol{q})=m\,\frac{q(s)}{p(s)}$, indirect utility is $EU(\boldsymbol{p},m,\boldsymbol{q})=\mathrm{Ln}(m)+D(\boldsymbol{q}||\boldsymbol{p})$, the indirect certainty equivalent CE is $CE=me^{D(\boldsymbol{q}||\boldsymbol{p})}$,the expenditure function is $m(\boldsymbol{p},\boldsymbol{q},eu)=e^{eu-D(\boldsymbol{q}||\boldsymbol{p})}=CEe^{-D(\boldsymbol{q}||\boldsymbol{p})}$ and Hicksian compensated demands are $z^h(s,\boldsymbol{p},\boldsymbol{q},eu)=e^{eu-D(\boldsymbol{q}||\boldsymbol{p})}\,\frac{q(s)}{p(s)}$. Since $\frac{q(s)}{p(s)}=e^{\mathrm{Ln}\,\frac{q(s)}{p(s)}}$, Log transforms of optimal choices are an increasing function of actual and expected relative predictive abilities of the agent and the market as assessed by the Log scoring rule. Indirect certainty equivalents for budget sets $me^{D(\boldsymbol{q}||\boldsymbol{p})}$ are determined by the agent's expected Log score advantage over the market $D(\boldsymbol{q}||\boldsymbol{p})$. Expected wealth at optimal choice $\Sigma q(s)z^o(s)$ as a multiple of income m is the expected Bayes factor in favour of the agent's predictive hypothesis $\Sigma\frac{(q(s))^2}{p(s)}$, an increasing convex function of subjective beliefs $\boldsymbol{q}$ with a minimum value of 1 when q(s)=p(s). Note that the linear scoring rule $\frac{m}{p(s)}r(s)$ for fixed m and $p(s)=\frac{1}{n}$, which pays out *$nm* in proportion to report *r(s)* made, while not proper, will elicit beliefs ***q(s)*** honestly from an agent with Logarithmic utility of wealth. Reports r(s) simply determine quantities of contingent wealth *z(s)=r(s)nm* that can be traded off across states dollar for dollar, ie at equal relative prices. With such a budget line any other choice than truth telling is not expected utility maximising for this agent.

6. Relative risk is unbounded from above, since $D(\boldsymbol{n}||\boldsymbol{p})=\tau\{-\Sigma\frac{\mathrm{Ln}(p(s))}{n}-\mathrm{Ln}(n)\}$, the difference between an unbounded from above, non-negative, convex function $-\Sigma\frac{\mathrm{Ln}(p(s))}{n}$ and its minimum value, Ln(n).

Appendices

Appendix A: Optimal Choice:

This appendix explains the derivation of equations1.1→1.4. Consider first expenditure minimisation:

AA.1• $\min_{\boldsymbol{z}} \boldsymbol{p(s)} \cdot \boldsymbol{z}$ subject to $\mathrm{EU}(\boldsymbol{z},\boldsymbol{q},\tau) \equiv \sum_s q(s)\left(-e^{-\frac{1}{\tau} z(s)}\right) = eu$

The first order conditions for AA.1 are:

AA.2a• $p(s) = \lambda\, q(s)\, \frac{1}{\tau}\, e^{-\frac{1}{\tau} z(s)}$, where $\lambda = \frac{\tau \sum p(s)}{(-eu)} = \frac{\tau}{-eu}$ when $\sum p(s) = 1$.

or in terms of relative prices:

AA.2b• $\frac{p(s)}{p(t)} = \frac{q(s)}{q(t)}\, e^{-\frac{1}{\tau}(z(s)-z(t))}$

Inverting AA.2a, using $ce=-\tau \mathrm{Ln}(-eu)$, $\mathrm{CE}(\boldsymbol{z})=-\tau \mathrm{Ln}(\mathrm{EU}(\boldsymbol{z},\boldsymbol{q},\tau))$, and solving for $s=1...n$ Hicksian compensated demands $\boldsymbol{z}^h$ (equation 1.4):

AA.3• $z^h(s,\boldsymbol{p},eu,\boldsymbol{q},\tau) = \tau[-\mathrm{Ln}(-\mathrm{eu})) + \mathrm{Ln}(\sum_t p(t)) - \mathrm{Ln}(\frac{p(s)}{q(s)})]$

$= \mathrm{CE}(\boldsymbol{z}^h) + \tau \mathrm{Ln}(\frac{q(s)}{p(s)/\Sigma p(t)})]$

$= \mathrm{CE}(\boldsymbol{z}^h) + \tau \mathrm{Ln}(\frac{q(s)}{p(s)})]$ when $\sum_t p(t) = 1$

Summing $p(s) z^h(s,\boldsymbol{p},eu,\boldsymbol{q},\tau)$ across all states and using $\frac{\boldsymbol{p}}{\boldsymbol{\Sigma p}}$ for the n-tuple of prices ***p(s)*** normalized by $\Sigma p(s)$, we obtain the expenditure function (equation 1.3):

AA.4• $m(\boldsymbol{p},eu,\boldsymbol{q},\tau) = \boldsymbol{p} \cdot \boldsymbol{z}^h(\boldsymbol{p},eu,\boldsymbol{q},\tau)$

$= \tau[(\Sigma p(t))\{-\mathrm{Ln}(-eu) + \mathrm{Ln}(\Sigma p(t))\} - \Sigma_s\, p_s\, \mathrm{Ln}\left(\frac{p(s)}{q(s)}\right)]$

$= \Sigma p(t)[\mathrm{CE}(\boldsymbol{z}^h) - \tau\, \mathrm{D}(\frac{\boldsymbol{p}}{\boldsymbol{\Sigma p}} \| \boldsymbol{q})]$

$= \mathrm{CE}(\boldsymbol{z}^h) - \tau\, \mathrm{D}(\boldsymbol{p} \| \boldsymbol{q})$ when $\Sigma p(t) = 1$

Solving AA.4 for *eu* and substituting in $m = \boldsymbol{p} \cdot \boldsymbol{\omega}$, we find the indirect utility function (equation 1.1)

AA.5 $\mathrm{EU}(\boldsymbol{p},\omega,\boldsymbol{q},\tau) = -e^{-\frac{1}{\tau(\sum_S p(s))}[\boldsymbol{p} \cdot \omega + \tau D(\boldsymbol{p} \| \boldsymbol{q})] + \mathrm{Ln}(\sum_S p(s))}$

$= -e^{-\frac{1}{\tau}[\boldsymbol{p} \cdot \omega + \tau D(\boldsymbol{p} \| \boldsymbol{q})]}$ when $\sum_s p(s) = 1$

The ordinary demand functions, AA.7 below, are the solutions to the expected utility maximisation problem: $\underset{\boldsymbol{z}}{\text{Max}} \sum_s q(s)\left(-e^{-\frac{1}{\tau} z(s)}\right)$ subject to constant wealth $\boldsymbol{p}\bullet\boldsymbol{z} = m$

AA.6 $z^o(s,\boldsymbol{p},\boldsymbol{q},m,\tau) = \frac{1}{\sum_s p(s)}[m + \tau D(\boldsymbol{p}\,||\,\boldsymbol{q})] - \tau \text{Ln}\,\frac{p(s)}{q(s)}$

Equation 1.2 in the text is AA.6 with $\sum_s p(s) = 1$ and income $m=\boldsymbol{p}\bullet\boldsymbol{\omega}$.

From AA.5 the certainty equivalent for a trading opportunity ($\boldsymbol{p},\boldsymbol{\omega}$) (equation 1.7) is AA.7 with income m= $\boldsymbol{p}\bullet\boldsymbol{\omega}$ using that $\sum_s p(s)=1$

AA.7 CE($\boldsymbol{p}$,m)= $\frac{1}{\Sigma_s p(s)}[m + \tau D(\boldsymbol{p}\,||\,\boldsymbol{q})] - \tau \text{Ln}\,(\sum_s p(s))$

Appendix B: Log Sum Inequality.

For non negative numbers $p_1 \ldots p_m$ and $q_1 \ldots q_m$

$$\sum_{i=1}^m p_i \log \frac{p_i}{q_i} \geq \left(\sum_{i=1}^m p_i\right) \log \frac{\sum_{i=1}^m p_i}{\sum_{i=1}^m q_i}$$

with equality iff $\frac{p_i}{q_i} = constant$ for $i=1\ldots m$

Proof: Continuity implies the conventions 0 log 0=0, a log $\frac{a}{0} = \infty$ for $a>0$, and 0 log $\frac{0}{0}=0$. Hence, without loss of generality, assume strictly positive p_i and q_i. Note that the function f(t)= t log t is strictly convex. Using $\alpha_i = q_i / \sum_{j=1}^m q_j$ and $t_i = p_i / q_i$ apply Jensen's inequality, $\sum\alpha_i\, f(t_i) \geq f(\sum\alpha_i\, t_i)$ for $\alpha_i \geq 0$ and $\sum \alpha_i = 1$, to obtain the log sum inequality.§

Appendix C: Chain rule for relative entropy.

$$D(\boldsymbol{p(s,t)}||\boldsymbol{q(s,t)})=D(\boldsymbol{p(t)}||\boldsymbol{q(t)})+\sum_t p(t)D(\boldsymbol{p(s|t)}||\boldsymbol{q(s|t)})=D(\boldsymbol{p(s)}||\boldsymbol{q(s)})+\sum_s p(s)D(\boldsymbol{p(t|s)}||\boldsymbol{q(t|s)})$$

Proof: The first equation follows directly from the following expansion:

$$\sum_s\sum_t p(s,t)\text{Ln}\,\frac{p(s,t)}{q(s,t)}$$
$$=\sum_s\sum_t p(s,t)\text{Ln}\,\frac{p(s|t)\,p(t)}{q(s|t)\,q(t)}$$
$$=\sum_t p(t)\text{Ln}\,\frac{p(t)}{q(t)} + \sum_t p(t)\sum_s p(s|t))\text{Ln}\,\frac{p(s|t)}{q(s|t)};$$

A similar expansion around s first establishes the second equation§

Appendix D: Variance minimization

Mimimizing variance $\sum_s q(s)(z(s) - \mu)^2$ subject to a constraint on mean wealth, $\Sigma_s q(s)z(s)=\mu$, and the budget constraint $\Sigma_s\, p(s)z(s)=\varpi<\mu$ has first order condition $2q(s)\{z(s)-\mu\}=\lambda_1\, q(s)+\lambda_2\, p(s)$ with multipliers λ_i at an interior solution. Summing across states s and using that ***q*** and ***p*** are pmfs, $\lambda_1+\lambda_2=0$, so the first order condition reduces to $z(s)-\mu=\frac{\lambda_1}{2}\{1-\frac{p(s)}{q(s)}\}$.

Using the budget constraint and summing across states

$\lambda_1=2\frac{\left(\Sigma\,\frac{p(s)^2}{q(s)}-1\right)}{\mu-\varpi}>0$ since $\Sigma\,\frac{p(s)^2}{q(s)}$ is a convex function in p(s) with a minimum value of 1 at p(s)=q(s), a generalisation of the repeat rate for a distribution Lad(1996),

■ References

Arrow, Kenneth "Informational Structure of the Firm." American *Economic Review* May 1985 (*Papers and Proceedings*) 75(2), pp. 303-397

Bernardo,Jose M and Adrian F M Smith. *Bayesian Theory*. New York:Wiley,1994

Brockett, Patrick L and Golden, L.L. "A Class of Utility Functions Containing all the Common Utility Functions." *Management Science*,1987, (33), pp. 955-969

Cover, Thomas M. and Thomas, J *Elements of Information Theory.* New York:Wiley,1991

Cornes, Richard, *Duality and Modern Economics.* New York:Cambridge University Press, 1992

Good, I. J. (1998) "Subjective Probability" in John Eatwell, Murray Milgate, Peter Newman eds., *New Palgrave : A Dictionary of Economics* (4), pp. 537-542, Reprint edition,1998

Holt, Charles A."Scoring rule Procedures for Eliciting Subjective Probability and Utility Functions." in Prem K Goel and A ZelLner eds. *Bayesian Inference and Decision Theory: Essays in Honour of Bruno de Finetti* Amsterdam : North-Holland : Elsevier, 1986,pp. 279-290

Hirshleifer, Jack and Riley, John. *The Analytics of Uncertainty and Information.*New York:Cambridge University Press,1992

Jessop, Alan. *Informed Assessment: An Introduction to Information,Entropy, and Statistics.*New York*:* Ellis Horword Press, 1995

Kass, R.E. and A.E. Raftery. "Bayes Factors" *Journal of The American Statistical Association*,June 1995, 90 (430), pp. 773-795,

Lad, Frank. *Operational Subjective Statistical Methods: A Mathematical Philosophical and Historical Introduction*. New York:Wiley, 1996

Leamer, Edward E,(1986), "Bid-Ask Spreads for Subjective Probabilities" n Prem K Goel and A ZelLner eds. *Bayesian Inference and Decision Theory: Essays in Honour of Bruno de Finetti* Amsterdam : North-Holland : Elsevier, 1986

Mas-Colell,Andreu , Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory* New York : Oxford University Press, 1995.

Pratt, John W and Zeckhauser,Richard J. "Proper Risk Aversion" *Econometrica*,1987, 55(1), pp. 143-154

Salanie, Bernard *The Economics of Contracts: A Primer*. Cambridge:MIT Press,(1987) Savage, Leonard J. "Elicitation of Personal Probabilities and Expectations" *Journal of American Statistical Association.* December 1971, 66(336), pp. 783-801

Von Winterfeldt, Detlof. and Ward Edwards. *Decision Analysis and Behavioral Research.* New York : Cambridge University Press,,1986