Relative Entropy - Giving credibility to judgement

Relative Entropy seeks to improve the realism of econometric models by objectively including judgement during the modelling process. Alun Marriott of EMB explains.

Sensitivity to economic risk has rarely been more topical. Enterprise Risk Management, the ICAS regime and Solvency II all depend fundamentally on a framework that supports an assessment of economic risk in an uncertain world - not only to improve our understanding of variability in asset or liability valuation, but also the risk to the business strategy itself.

The financial crisis has proved a wake-up call to those charged with "forecasting" the economic future. It is apparent that, whichever approach we take, historic calibration alone is not sufficient to prepare for the socio-economic shocks that affect our economy. Foreign exchange models would have had great difficulty in foreseeing the collapse of Bretton Woods or the arrival of the EMU, yet the dynamic of the exchange rate process is clearly dependent on the regime in which it moves.

We need, therefore, to *coherently* overlay a wider source of opinion, or judgement, to support historically calibrated models of the economy (often known as an Economic Scenario Generators, or ESGs for short). This is non-trivial given the large number of parameters needed to feed models of multiple economies with multiple financial series within each; and coherence is crucial if we wish to retain the dependency structure we observe.

To capture these dependencies, we may either propose a formulaic link (i.e. one series explicitly refers to a term in another), or propose instead to capture non-formulaic dependencies by a copula approach that links the random innovations explicitly. In any case, it is clear that we not only wish to control unwanted effects from user-specified adjustments, but we must also ensure that the dependency structure is retained.

Practitioners may choose to use Bayesian statistics to introduce judgement into the modelling process. Here, the analyst proposes a “prior” distribution for each model parameter to ensure the economic projections behave as desired; but the difficulty with this approach is that the model parameters are not the quantities we directly observe. Hence, it is often difficult to determine how parameters should be changed to achieve a desired economic characteristic, especially when each economy and each series are intertwined.

In contrast, Relative Entropy is a technique that quantifies the information value in a statistical distribution and provides a framework that allows us to modify, in a statistically robust manner, the economic forecast distributions directly.

Relative Entropy is widely used in the scientific community, with applications as diverse as optical character recognition, spacecraft re-entry physics and derivative pricing [i]. Each of these areas possesses a problem domain with intractable variability and computational overload. Relative Entropy helps by allowing model specification to be simplified whilst still ensuring the models exhibit externally specified characteristics.

To return to the ESG, we wish to blend in, or overlay, our views of the global economy with a previously calibrated ESG. Traditionally, each simulation of an ESG is deemed equally likely.

However, let's instead propose we increase the relative likelihood of those simulations that make some statistic (such as the mean) more aligned with our "judgement"; but we need to do so in a manner that minimizes disturbance to the "a-priori" uniform distribution of the simulation weights and hence we need to define a measure of the "disturbance" we have just introduced.

Based on a paper by Shannon [ii] in 1948, Entropy provides us with just such a quantitative metric. A fully uninformative (i.e. uniform) distribution has the greatest entropy, whereas a distribution with a single certain event (i.e. a Dirac distribution) has an Entropy of zero. Applying this to our Monte-Carlo simulation, we start with a uniform distribution of simulation weights, and gradually move away from uniformity as we introduce additional information. To reiterate, we aim to disturb the original distribution the "least", i.e. tend to a uniform distribution of simulation weights.

Suppose this additional information, $\mathcal{X}$, is just the expectation of one of our modelled variables, X, i.e. $\mathbb{E}[X] = \mathcal{X}$, (for example, the mean level of the European equity market in 4 years time).

The "Informational Entropy", H, is defined to be

$$H(P) = -\sum_{j=1}^{n} p_j \ln p_j$$

where P represents the full probability distribution, and p_j the probability of simulation j occurring. i.e. as you change the weights of each simulation, you change the measure of Informational Entropy, $H(P)$.

Using the technique of Lagrangian multipliers, we aim to maximize Informational Entropy, $H(P)$, subject to two constraints, namely
$\Sigma_i(j = 1)^t n \; p_j = 1$ and $\Sigma_i(j = 1)^t n \; [x_j \; p_j] = ($

where each x_j represents the value of the chosen variable, X, and p_j the probability of simulation j respectively.

We wish to find the stationary points of the Lagrangian of the form

$$L(p_1, p_2, \ldots, p_n; \lambda_1, \lambda_2) = -\sum_{j=1}^{n} p_j \ln p_j + \lambda_1 \left(\sum_{j=1}^{n} p_j - 1 \right) + \lambda_2 \left(\sum_{j=1}^{n} x_j p_j - (\right)$$

where λ_1 and λ_2 are the Lagrangian multipliers.

Solving for λ_1 and λ_2, we can show that

$$p_j = \frac{e^{\lambda_2 x_j}}{\sum e^{\lambda_2 x_j}}$$

the "Gibbs" distribution, the "natural" alternative to a uniform distribution when we are ignorant of all but the expectation of X, namely (.

We now have a mechanism that enables us to "reweight" the Monte-Carlo simulations so that, rather than assuming each simulation of the global economy occurs with equal probability, each simulation occurs with a probability ascribed by the Gibbs distribution above.

Though it would be possible to apply the Gibbs expectation to the Monte-Carlo integration directly, many downstream applications still *assume* each simulation occurs with equal likelihood. Our next step is hence to resample the post-overlay simulations in proportion to the Gibbs simulation weights, creating an equally weighted post overlay ESG.

Before moving to a brief example, it is worth considering some key properties.

Firstly, Relative Entropy does *not* introduce new simulation paths, it simply reweights existing paths from the original simulation set. Hence, it is important to ensure that the source simulation set provides a rich universe of path possibilities. Secondly, we must consider simulation error when an overlay is applied, particularly as re-sampling can lead to duplication of simulations - more so when the overlays are extreme. This can be mitigated by increasing the number of simulations and ensuring overlays remain sensible.

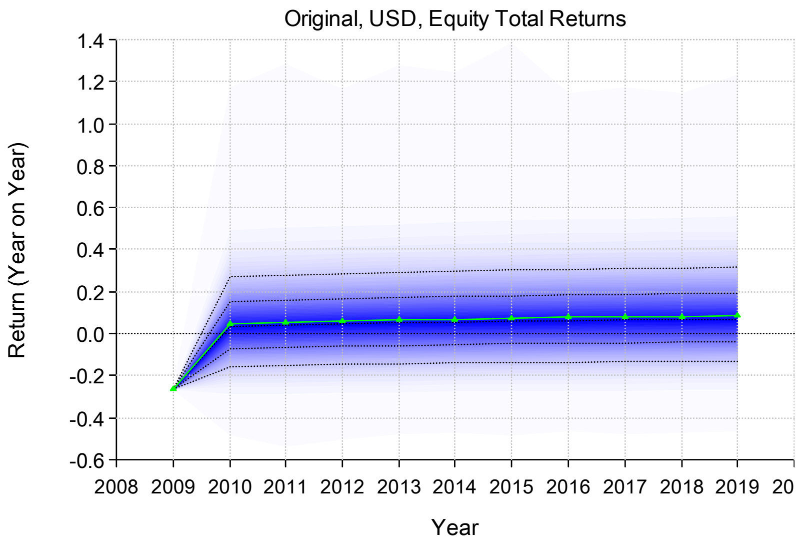
Lastly, we turn to the question of confidence. In a Bayesian world, one can express confidence in an overlay by specifying an appropriately “tight” prior distribution of the model parameters. However, since we do not adjust the parameters in this approach, we instead adopt an approach taken in a recent paper by Meucci [iii] which linearly blends the simulation weights from the pre and post overlay ESGs, allowing us to express our confidence measures consistently.

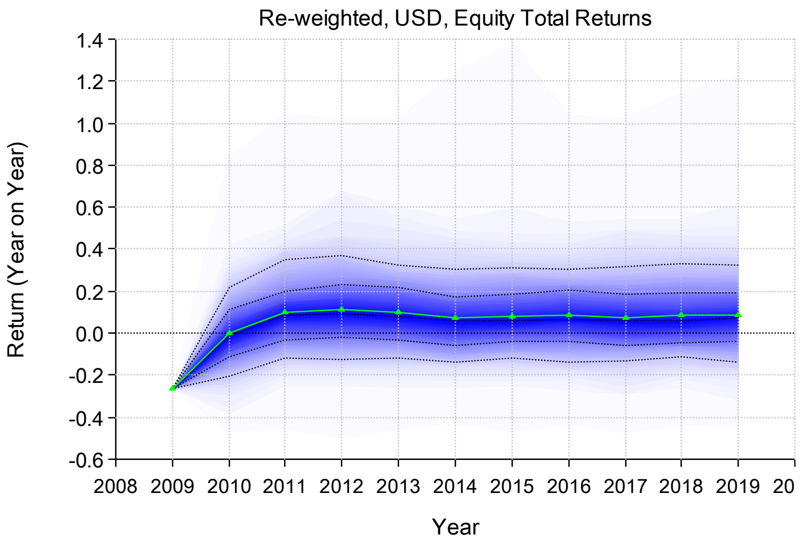
Applications and a brief example

In the field of actuarial finance, numeric complexity leads many to apply Monte-Carlo techniques widely and there will undoubtedly be situations where we may wish to adjust the probability distributions feeding the stochastic process; as, for example, we recently explore in an application to the determination of non-life reserves.

Returning to our ESG, a brief example below shows how we can overlay an externally provided forecast of price inflation, foreign exchange and cash rates onto a pre-calibrated ESG.

The first graph shows projections of the annual returns from the US equity market relying on historic calibration alone. The second shows the post-overlay projections with clearly visible structure to the evolution of returns - typically 4% higher in 2011 and 2012. It is interesting to note that we did not specify views as to the path of the equity market itself, but this followed given the dependency structure and overlays we supplied.





I hope this article has helped introduce the entropy concept, and proves timely given the imminent arrival of Solvency 2 and a corresponding focus on an improved understanding of the assumptions that drive the economic modelling process.

Alun Marriott runs the investment division of EMB.

Alun.Marriott@emb.com

---

i C. Neri, L Schneider, "Maximum Entropy Distributions Inferred from Option Portfolios on an Asset" arXiv:0903.4542v1 [q-fin.PR]
ii C. Shannon, "A Mathematical Theory of Communication", Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, July, October, 1948
iii Attilio Meucci, Fully Flexible Views: Theory and Practice (2008), Risk, 21, 10, 97-102