Optimal hedging strategies for Australian electricity retailers

Jeff Dewynne
University of Wollongong and Oxford University

Pam Davy
University of Wollongong

1 Introduction

Electricity retailers, such as Integral Energy, provide electricity to domestic consumers at a price known as the transfer-price. It varies in a deterministic manner with the time of day and the day of the year. This variation occurs on long time scales (the order of a number of years), which is of interest in this report.

Retailers obtain their electricity from a pool and the price of electricity from the pool, the pool-price, varies randomly and at half-hourly intervals. Often the variations can be dramatic and it is not unusual for there to be spikes during peak periods, where the half-hourly pool-price varies by a multiplicative factor of the order of one hundred.

The amount of electricity the retailer has to supply to their customers is known as the load and it too is a process which varies randomly at half-hourly intervals. Figure 1 shows typical (simulated) loads and prices as time series and figure 2 shows a time series of typical average load and average log(pool-price), along with a rough indication of their variability. Note that in the data provided by Integral Energy, monthly loads are normalised so that the average monthly load is unity.

At each half-hourly interval the retailer must sell a random quantity of electricity at a fixed price but must buy it at a random (and fairly volatile) price. Therefore the retailer is exposed to significant market risk. Two relevant measures of this risk are Value-at-Risk (VaR) and Earnings-at-Risk (EaR). Value-at-Risk is the more common of these measures; its estimation is often required by regulatory authorities. It may be defined as the maximum loss at a given confidence level and over a relatively short, fixed, time period that can occur when liquidating a portfolio. Earnings-at-risk is a related concept and, in this context, it is defined to be the maximum loss incurred at a given confidence level over a fixed time interval if the current committed portfolio is fully exposed to the volatility of the pool price. In effect, it is the largest loss (at the given confidence level and over the fixed time interval) that can occur as a result of selling electricity to the consumer at the fixed price and buying from the pool at the varying price.

In order to protect against the risk resulting from the varying pool price and load, an electricity retailer typically enters into derivative contracts, such as forwards, swaps and caps. These contracts usually cover the retailer for a small number of days (for example, 48 half-hourly periods) and must be set up well in advance, often up to a year before they are used. They are also relatively expensive so, on the one hand, it is highly inefficient to buy

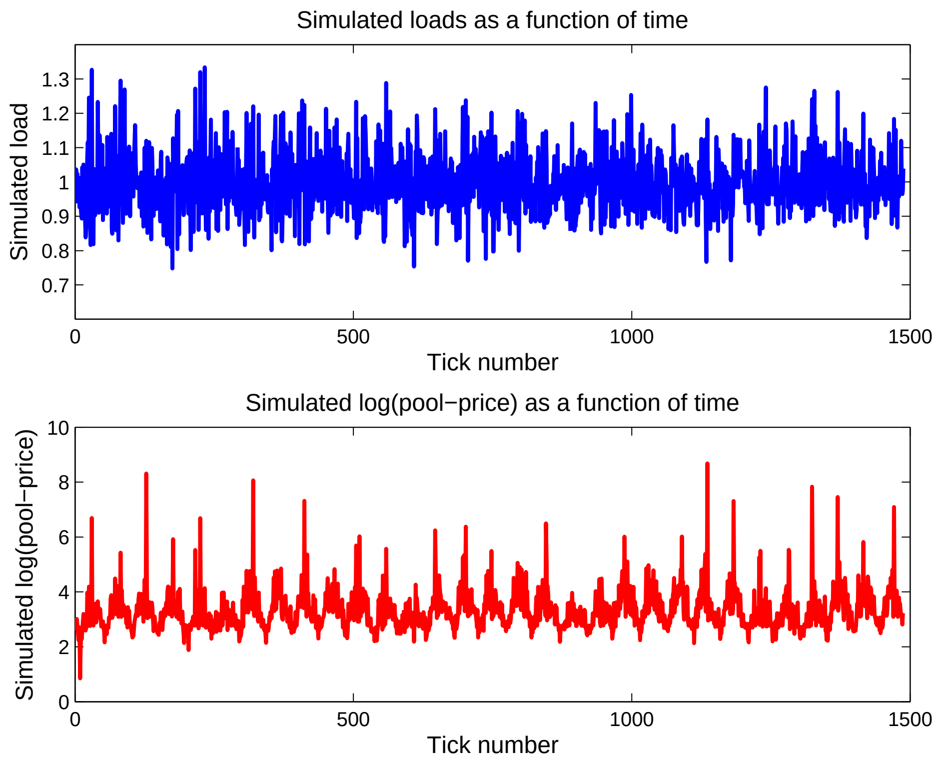

Figure 1: Typical simulated time series for load and log(pool-price) as functions of time (from simulated data for January 2008). Note that load is normalised over each month so that the average monthly load is unity.

more contracts than necessary. On the other hand, it is very dangerous to leave this risk completely unhedged.

The problem involves two main random processes; the load, which is the number of units of electricity consumers use, and the pool price, which is the cost of the electricity (per unit) to the retailer. The bivariate distribution of load and pool price is highly non-Gaussian, as illustrated in Figure 3.

Most retailers have models which predict (with varying degrees of accuracy) the expected values and random variations about these expected values of both the load and the pool-price. These can, and usually do, take factors such as holiday periods, expected weather patterns and so forth, into account when producing forecasts. Although both these random processes have statistical properties that vary with time, we may assume that the retailer's model for these variations are sufficiently accurate and, for our purposes, do not require any further refinement. As already noted, in our price-load data the load is normalised so that the monthly average is unity.

For an unhedged position, the earnings (over any particular half-hourly period) are simply half the product of the load and the difference between the transfer price and the pool-price; the factor of one half arises because of the half-hourly time increment. The earning may be changed by introducing derivate contracts, in which case the earnings are the unhedged

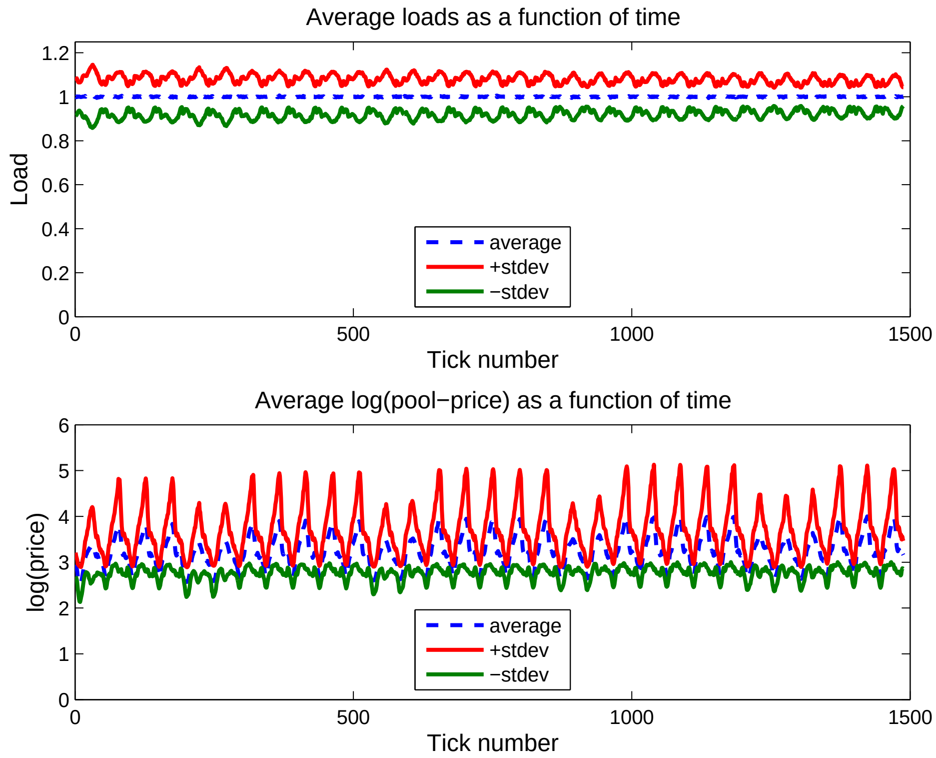

Figure 2: Typical simulated time series for average load and average log(pool-price) as functions of time (from simulated data for January 2008). Note that load is normalised over each month so that the average monthly load is unity.

earnings, minus the cost of the derivative contracts involved, plus the cash-flows from these derivative contracts. The monthly earnings, which are of particular interest, are just the sum of all half-hourly earnings over each half-hourly interval in the month.

The aim of the project is to find portfolios of derivative contracts which, when added to the retailers unhedged position, produce earnings which have statistical properties that are, in some sense, optimal. That is, the aim is to choose numbers of derivative contracts, from a given set of available contracts, that results in a position that is, in some sense, optimal.

To this end, there are three particular quantities turn out to have special significance:

- the expected earnings, i.e., the average earnings — with the average being taken over a calendar month;
- the lower five percent quantile of the earnings, i.e., the level of monthly earnings which will be exceeded ninety-five percent of the time, again taken over one month; and
- the (monthly) earnings at risk, which is the difference between the two previous quantities.

The first and third of these are particularly important because management requires the (monthly) expected earning to be positive and sets an upper bound on the (monthly) earnings

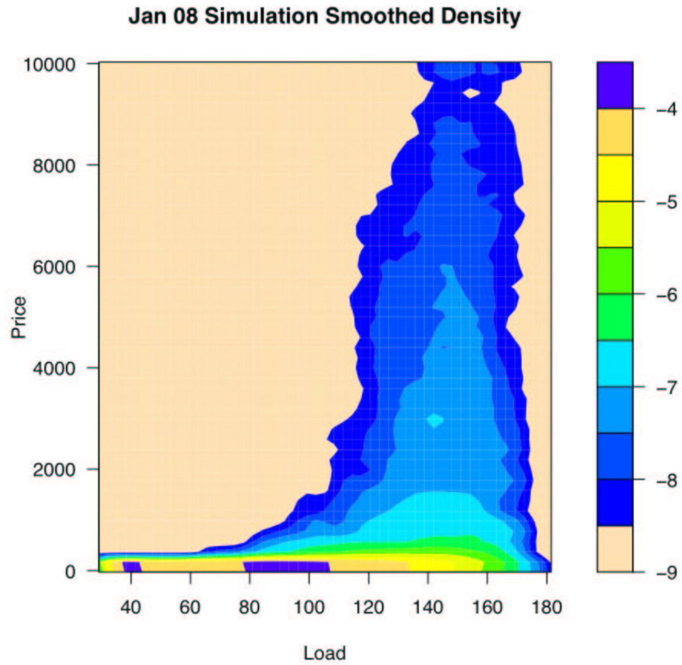

Figure 3: Bivariate price and load distribution function from the January 2008 simulation. The shading is on a logarithmic scale. Note that this distribution is computed using 3,000 simulations of 1,488 half-hourly price-load pairs and therefore represents the distribution of price-load over the entire month, rather than at any particular single half-hourly tick.

at risk. Therefore, any hedging strategy proposed *must* ensure that these two requirements are satisfied.

During the study group we looked at very simple portfolios, consisting of the unhedged position plus:

- a number of swaps; a swap effectively allows the retailer to buy electricity from the pool for a fixed price (we took both peak and off-peak swaps, which we assumed to be par swaps – the fixed price is chosen so that it costs nothing to enter into the contract when it is set up — although modifying this to allow for non-par swaps is a simple matter);
- a number of caplets; a caplet effectively sets an upper bound on the price that the retailer must pay to buy over one half-hourly period and is therefore essentially a call option on a single unit of power with strike equal to the upper bound set on the price (we used both prices inferred from market prices of caps – see below – and prices based on the expected value of the caplet under the physical measure, plus a premium);
- a number of caps; a cap is effectively a series of caplets (we assumed a constant strike for the caplets although there seems to be no obvious impediment to allowing the strike to vary across caplets, apart from computational time. We priced that cap as a series of caplets, see above).

Although the distribution of the load and spot price is very far from a bivariate normal distribution, we find that the distribution of earnings does not seem any "less" normal than the distribution of profit and loss associated with many other (sensible) financial positions in general. Therefore, it seems appropriate to try the standard techniques used in analytic VaR (e.g., assume normality of changes in earnings and use delta-normal and delta-gamma approximations when derivatives are involved). We find, possibly surprisingly, reasonable agreement in parameter space with the results obtained from direct simulation using simulated price-load and data — the difference between the two are comparable with the differences in VaR estimates one finds between the various flavours of VaR, for example.

2 Nomenclature and basic results

In what follows we use:

t	as our time index (with units of half-hours); it takes integer values and particular values of t are referred to as "ticks";
Q	to denote the load in general (a positive random process);
P^f	to denote the pool-price in general (a positive random process);
P^w	to denote the transfer price in general (positive and deterministic);
Q_t	to denote the load at time t (a positive random variable);
P_t^f	to denote the pool-price at time t (a positive random variable);
P_t^w	to denote the transfer price at time t (a positive number).

In what follows, we do not attempt to generate Q_t, P_t^f or P_t^w, rather we use values provided to us (produced by simulation in the case of the load and pool-price, Q_t and P_t^f, and either from regulatory data or, for simplicity, as a constant in the case of the transfer price, P_t^w).

The unhedged earnings at time t are given by

$$E_t = \tfrac{1}{2}\, Q_t \left(P_t^w - P_t^f \right),$$

where the factor of a half is necessary because Q has units of power and P^w and P^f have units of price per unit of energy (and therefore the half implicitly has units of hours).

The unhedged monthly earnings are given by[1]

$$E_M = \sum_{t=1}^{m} E_t$$

where m is the number of half-hourly periods in the particular month in question (typically $m = 1,440$ or $1,488$). For any particular month, this is a random variable, with a distribution similar to that shown in Figure 4. Note that although the simulated distribution is unimodal, it is quite strongly leptokurtotic, so is significantly different from a normal distribution with

[1] We do not attempt to correct for the time value of money here; partly because both the effect of doing so and the necessary modifications are relatively minor but mainly because the corrections depend crucially on the billing and accounting procedures, to which we are not privy.

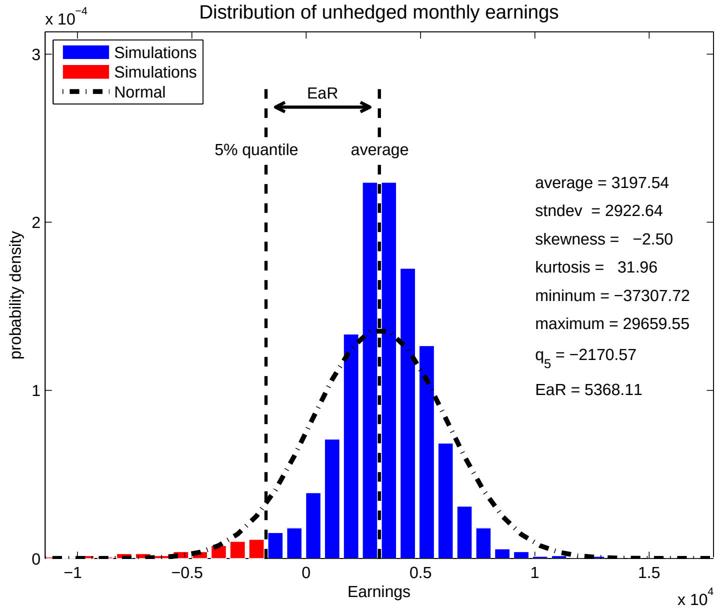

Figure 4: Distribution of monthly unhedged-earnings from the January 2008 set of simulations, containing 3,000 simulations of 1,448 half-hourly price-loads and pairs. For simplicity we used a constant transfer price of $P_t^w = 48.00$. Note that the 5% quantile is based on the simulated density function rather than the normal density and that returns below this level are shaded red and those above are shaded in blue in the figure.

the same mean and variance, as it has fat tails. In fact, the distribution is qualitatively very similar to the distribution of returns on equities and indices.

As noted, a constraint imposed by management is that the expected value of the monthly earnings must not be negative,

$$\mathbb{E}[E_M] \geq 0. \tag{1}$$

Also illustrated in Figure 4 is the 5%-quantile (of the simulated earnings distribution), which we shall denote by q_5, the 5%-quantile of the distribution of the monthly earnings. It is important because the Earnings at Risk, EaR, are defined to be the difference between the mean monthly earnings and the 5%-quantile,

$$\text{EaR} = \mathbb{E}[E_M] - q_5.$$

As mentioned earlier, management impose a restriction on the earnings at risk, of the form

$$\text{EaR} \leq \epsilon, \tag{2}$$

where $\epsilon > 0$ is prescribed.

As it stands, there is no control variable available in order to satisfy constraints (1) and (2). In order to gain some measure of control over the distribution of monthly earnings, an electricity retailer must hedge its exposure with derivative contracts. The two most important, for the purposes of this report, are swaps and caps.

2.1 Swaps

The effect of a swap is to allow the retailer to buy one unit from the pool at a known price, for a small number of half hour intervals. In practice this known price may be a function of time, but for simplicity we shall assume that it is a constant which we denote by P^S. For a single half hourly tick, the swap pays out an amount

$$\tfrac{1}{2}\left(P_t^f - P^S\right)$$

per unit of load. Typically the swap only applies to one unit of load, so if the retailer wishes to hedge n units of load, they must buy n swaps. Note that because our load data is normalised so the monthly average is unity, what is referred to as n swaps in this report does *not* correspond to n swaps in reality; the n referred to in this report must be rescaled to determine the actual number of swaps necessary in the real world.

If the retailer holds a single swap then they pay P_t^f for one unit of load, if they buy it, and receive $P_t^f - P^S$ from the swap for that unit of load and so the net cost of that unit of load over the half-hour tick is

$$\tfrac{1}{2}\left(P_t^f - (P_t^f - P^S)\right) = \tfrac{1}{2}P^S,$$

so in effect they pay P^S for the unit of load. If the retailer holds a swap but doesn't buy any power from the pool, they simply receive

$$\tfrac{1}{2}\left(P_t^f - P^S\right),$$

a quantity which may be negative (in which it represents a cost).

The price of the swap depends on the swap level, P^S. A par swap (which is the most common type of swap in most areas of financial engineering) is a swap which costs nothing to enter into (in which case P^S is uniquely determined as the "fair swap price" and typically turns out to be the expected price under some probability measure). In general, we shall assume that swaps are *not* par swaps and that they have a non-zero price. So we denote V^S as the up-front price of a swap.

A par swap is achieved by setting P^S so that this value is zero, $V^S = 0$. It is self-evident that as the swap level increases the cost of the swap decreases (indeed, if we assume the absence of arbitrage, this can be proved as a "theorem"). We shall also assume that a swap is available to cover an entire month's worth of ticks.

If the retailer holds N_S swaps with swap level P^S then their monthly earnings become

$$E_M = \sum_{t=1}^{m} \tfrac{1}{2}\left(Q_t(P_t^w - P_t^f) + N_S(P_t^f - P^S)\right) - N_S V_S.$$

Allowing the swap level to vary with time[2] amounts to changing P^S to P_t^S, where P_t^S is a known deterministic function of t, in this equation. Furthermore, allowing for the fact that, in reality, a sequence of swaps is necessary to hedge an entire month amounts to replacing N^S by N_t^S in the above where N_t^S is again a known deterministic function of t. In this case

[2] As examples, we would need to do this to model a sequence of peak and off-peak swaps or to model a sculptured swap.

the term $N_S V_S$ must be replaced by the total cost of all swaps involved. In view of this, there is no particular generality to be gained by assuming that more than one swap is involved: if this is the case then all that need be done is to replace N^S and P^S by the appropriate N_t^S and P_t^S.

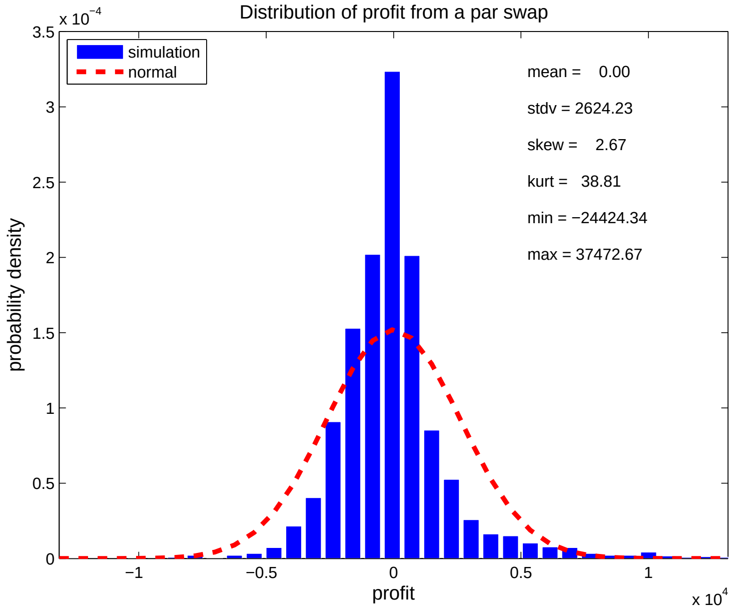

Figure 5: Distribution of monthly income for a single par swap on one unit of load and with swap level of $P^S = 42.4$, from the January 2008 set of simulations. Note that this swap is *not* being used to hedge.

Figure 5 illustrates the distribution of monthly profit and loss for a month-long par swap with constant swap level $P^S = 42.4$, which is the average monthly pool-price under the obvious measure. That is, the figure shows the distribution of profit and loss from buying one unit of power at the swap level and selling it at the pool-price at each half-hourly interval of the entire month. It is unimodal, very leptokurtotic and slightly skewed to the left.

2.2 Caps

A cap is a means of limiting the maximum price paid for a sequence of units of load purchased from the pool, and is analogous to a sequence of European call options, each with a different expiry date.

A caplet, which is the basic unit underlying a cap, is analogous to a single European call option. It gives the retailer the *right* to buy a unit of load for the specified cap level, P^C, at a given tick, t. The retailer does not have to exercise this right and would be foolish[3] to do so if the pool-price is below the cap level. A sensible retailer who holds a caplet which may be exercised at tick t receives

$$\tfrac{1}{2}\max\left(P_t^f - P^C, 0\right)$$

[3] Assuming there are no tax advantages in doing so, for example.

aer unit of load at tick t.[4]

As this can only ever be of benefit to the retailer, they must pay a positive premium for the caplet (usually in advance of the expiry tick, t).

If used to hedge a unit of power at tick t, the net effect is to limit the price of a unit of load to the caplet level, P^C. If the retailer holds a caplet and the pool-price is below the caplet-level they pay the pool-price for a unit of load and ignore the caplet, but if the pool-price is above the caplet-level they pay the pool-price for the unit of load, exercise the caplet and receive the difference between the pool price and the caplet level, which is the same things as paying the caplet level for the unit of load.

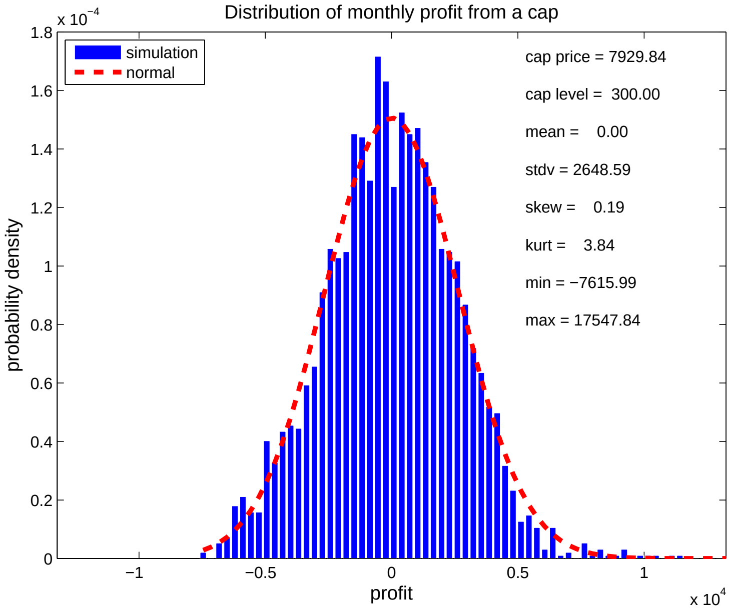

Figure 6: Distribution of monthly income for a single cap on one unit of load and with cap level of $P^S = 300$, from the January 2008 set of simulations. Note that this cap is *not* being used to hedge. The price is chosen so that the expected profit is zero and does not affect the *shape* of the distribution.

A cap is a sequence of caplets which can be exercised at different ticks, typically consecutive ticks. In this case the cap level may vary as a deterministic function of the tick and has the form P_t^C. For the purposes of this report, however, we shall assume that the cap level is constant in time. We shall also assume the cap consists of caplets for every tick in the month. If the retailer holds N^C caps in their portfolio the effect is that their monthly earnings become

$$E_M = \sum_{t=1}^{m} \tfrac{1}{2} \left(Q_t(P_t^w - P_t^f) + N^C \max(P_t^f - P^C, 0) \right) - N^C V^C,$$

where V^C is the cost of the cap (the sum of the costs of the individual caplets). There is no such thing as a par cap, in the sense that a cap necessarily has a positive price (if not there is a fairly obvious arbitrage opportunity). It is also obvious that as the cap-level P^C increases, the value of the cap decreases.

[4]The comments made about rescaling the number of swaps also apply to the numbers of caplets and caps.

The modifications to allow for tick dependent cap levels and caps with shorter lives are analogous to those for swaps, essentially all we need do is replace P^C by a deterministic function P^C_t and the constant number of caps, N^C, by a deterministic function N^C_t. The cost of the caps, $N^C V^C$ is then replaced by the total cost of all the caps (or caplets) involved.

Figure 6 shows the distribution of monthly profit and loss from a month-long cap with a constant cap level of 300. The price of the cap is chosen so that the expected profit, under the obvious measure, is zero; this is to some extent arbitrary but the only effect is on the location of the average. The distribution is slightly leptokurtotic and very slightly skewed relative to a normal distribution with the same mean and variance.

3 Strategy

In both this section and the next, we assume that the retailer is free to vary the number of swaps and number of caps, but can not vary the swap level, P^S, or cap level, P^S. That is, we assume that P^S and P^C are given and that N^S and N^C are our control variables. The swap and cap prices are, of course, uniquely determined by the swap and cap levels, respectively, and can not be varied independently of these.

For the purposes of illustration, we will assume that a single swap and a single cap are available and that these have constant swap and cap levels for the entire month.

This assumption is not particularly restrictive, in the sense that the number of available swaps and caps is limited and, unless the retailer is prepared to purchase over-the-counter contracts which are tailored to their specific requirements, the retailer must choose from this limited set. We can easily generalise to the situation where the choice is from a limited set by making P^S and P^C fixed functions of t, that is P^S_t and P^C_t, and then allow their numbers N^S and N^C to be tick dependent. That is, we can make our control variables N^S_t and N^C_t where t ranges of all ticks — in this case there may also be constraints on the values of N^S_t and N^C_t implied by the available contacts.[5] Needless to say, as the number of control variables increase so too does the dimension the problem, as does the computational complexity of solving it.

This assumption means that the monthly earnings can be decomposed into a linear combination of the numbers of swaps and caps[6]

$$E_M = E_{\mathrm{U}} + N^S E_{\mathrm{S}} + N^C E_{\mathrm{C}}, \tag{3}$$

where

- E_{U} denotes the monthly earnings from the unhedged position,
- E_{S} denotes the monthly income of a swap on a unit of load, and
- E_{C} denotes the monthly income from a cap on a unit of load.

The advantage of this is that E_{U}, E_{S} and E_{C} need only be calculated *once* per price-load simulation and then the total monthly earnings can computed by varying N^S and N^C without having to rerun the simulation. This is not true if we vary the cap-level in our cap, for instance.

[5]For example, if the shortest tenor for a swap is two days, it may be necessary for all of N^S_t, N^S_{t+1}, …, N^S_{t+47} to take the same value.

[6]This remains the case in the time dependent case, although in that case the linear function involves more terms.

This means that computing the distribution of the monthly earnings reduces to finding the distribution of a linear combination of distribution that are already know. From a numerical point of view, this amounts to constructing a linear combination of known vectors rather than a recalculation involving the entire simulation procedure.

3.1 The feasible region

The first task is to check that it is indeed possible to find any values of N^S and N^C which are consistent with *both* of the constraints (1) and (2), namely

$$\mathbb{E}[E_M] \geq 0 \quad \text{and} \quad \text{EaR} \leq \epsilon.$$

In our simple case, this is relatively easy because we can plot the contours of $\mathbb{E}[E_M]$ and EaR in our two dimensional parameter space. The result of this calculation for the January 2008 data set, a par swap with swap level $P^S = 37$ and a cap with cap level $P^C = 300$ and price $V^C = 5000$ is shown in Figure 7.[7]

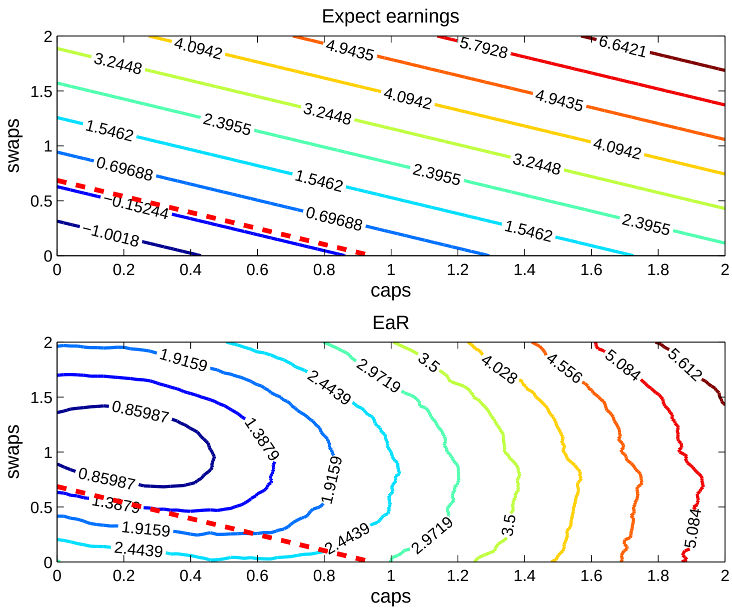

Figure 7: Contour plots of $\mathbb{E}[E_M]$ and EaR in (N^S, N^C) parameter space. The red dashed line represents the contour $\mathbb{E}[E_M] = 0$, the region to the right and above it represents the region in which $\mathbb{E}[E_M] > 0$ and the region to left and below it represent the region where $\mathbb{E}[E_M] < 0$. In this example we took a par-swap with $P^S = 37$ and a cap with $C^S = 300$ and $V^S = 5,000$.

[7] For this data set, there are 3,000 simulations of 1,484 half-hourly pool-price and load pairs. This results in vectors with 3,000 simulated values for each of the unhedged monthly earnings, swap income and cap income. The monthly earnings correspond to linear combinations of these vectors and its distribution is represented by a 3,000 element vector. The average value is easily computed. The 5%-quantile is computed by sorting the 3,000 element vector representing the simulated monthly earnings into ascending order and then looking, carefully, at the 150$^{\text{th}}$ element and its neighbours. The EaR is simply the difference between the average and the 5%-quantile.

It is clear that the curve $\mathbb{E}[E_M] = 0$ is a straight-line in parameter space, this follows by taking expectations of (3) to find that

$$\mathbb{E}[E_M] = \mathbb{E}[E_{\text{U}}] + N^S \mathbb{E}[E_{\text{S}}] + N^C \mathbb{E}[E_{\text{C}}].$$

In the more general case where we take N_t^S and N_t^C, the surface $\mathbb{E}[E_M] = 0$ becomes a hyperplane separating a half-space in which $\mathbb{E}[E_M] > 0$ form a half-space in which $\mathbb{E}[E_M] < 0$. Note that, in general, this result does *not* remain true if we include the swap and cap levels in our control variables; in that case the surface $\mathbb{E}[E_M] = 0$ becomes a nonlinear surface.

It is also apparent that the level curves of EaR are closed curves. This is almost certainly a consequence of the unimodal nature of the distributions involved and, if necessary, it should be possible to prove that this is indeed the case. Note that as q_5 is, by definition, the 5% quantile and is computed from a sample of $3,000$ simulations it is effectively based on a sample size of 150. Thus, it would be a mistake to regard the EaR derived from it, and in particular the level curves in Figure 7b, as having a high degree of precision.

It is also clear that there is a minimum value of EaR. If the earnings distribution is unimodal, the only way the EaR can be zero is if our portfolio is perfectly hedged and there is no risk at all: for the EaR to be zero, the 5%-quantile and the mean of the distribution must be identical and if the distribution is unimodal this implies that it is a delta-function, which in turn means that there is only one possible monthly earning.

Therefore, if we assume that a perfect hedge is not possible *and* try to impose the constraint EaR $\leq \epsilon$, with too small a value of ϵ, there will be no solutions, i.e., there is no feasible region if we try and restrict the EaR to too small a value.

4 Multivariate-normal approximation

We can make some analytic progress if we assume that E_{U}, E_{S} and E_{U} are drawn from a trivariate normal distribution. This assumption is to some extent justified by the marginal distributions shown in Figures 4, 5 and 6 which appear to be at least approximately normal.[8] Under this assumption, the distribution of

$$E_M = E_{\text{U}} + N^S\, E_{\text{S}} + N^C\, E_{\text{C}}$$

is itself normal, allowing us to proceed with the same sort of methods used in classical portfolio theory and analytic VaR. A similar comment applies if we allow time dependent N_t^S and N_t^C and assume that the corresponding earnings are drawn from a multi-variate normal distribution.

The distribution of E_M is then uniquely determined by

$$\bar{E}_M = \mathbb{E}[E_M] = \mathbb{E}[E_{\text{U}}] + N^S \mathbb{E}[E_{\text{S}}] + N^C \mathbb{E}[E_{\text{C}}]$$

[8] These figures, of course, only show us the marginal distributions of the trivariate distribution and it involves a great leap of faith to assume that because these marginal distributions are (very) approximately normal it follows that the joint distribution is also approximately normal. This obvious problem notwithstanding, it may also be possible to regard this approximation as the leading order term in an asymptotic approximation based on a skewness-kurtosis correction for the quantiles of the total earnings distribution, along the lines of a Cornish-Fisher expansion. This is a possibility that remains to be investigated. These analytic solutions may also serve as initial guesses for a numerical optimization based purely on simulated price-load data rather than the assumption of a multivariate normal distribution.

and

$$
\begin{aligned}
\sigma_M^2 &= \mathrm{var}[E_M] \\
&= \mathrm{var}[E_{\mathrm{U}}] + 2\,N^S\,\mathrm{covar}[E_{\mathrm{U}}, E_{\mathrm{S}}] + 2\,N^C\,\mathrm{covar}[E_{\mathrm{U}}, E_{\mathrm{C}}] \\
&+ \left(N^S\right)^2\,\mathrm{var}[E_{\mathrm{S}}] + \left(N^C\right)^2\,\mathrm{var}[E_{\mathrm{C}}] + 2\,N^S\,N^C\,\mathrm{covar}[E_{\mathrm{S}}, E_{\mathrm{C}}].
\end{aligned}
$$

The expectations, variances and covariances on the right-hand-sides of these expressions may be easily computed from simulated data and, under our assumptions, they only need be computed once as they do not depend on N^S or N^C.

For the purposes of illustration in what follows, we arbitrarily take a transfer price of 42, a par-swap level of 37, a cap level of 300 and a cap price of $5,000$ and use the January 2008 simulation set for the pool-price and load values. With these parameters we find we find that the expected monthly earnings and the covariance matrix of the expected monthly earnings are[9]

$$
\mathbb{E}\begin{pmatrix} E_{\mathrm{U}} \\ E_{\mathrm{S}} \\ E_{\mathrm{C}} \end{pmatrix} = \begin{pmatrix} -1.8511 \\ 2.7023 \\ 1.9690 \end{pmatrix},
$$

$$
\begin{pmatrix} \sigma_{\mathrm{U}}^2 & \sigma_{\mathrm{U\,S}} & \sigma_{\mathrm{U\,C}} \\ \sigma_{\mathrm{S\,U}} & \sigma_{\mathrm{S}}^2 & \sigma_{\mathrm{S\,C}} \\ \sigma_{\mathrm{C\,U}} & \sigma_{\mathrm{C\,S}} & \sigma_{\mathrm{C}}^2 \end{pmatrix} = \begin{pmatrix} 3.8584 & -3.3931 & -1.8381 \\ -3.3931 & 3.1103 & 1.2013 \\ -1.8381 & 1.2013 & 3.1683 \end{pmatrix}.
$$

In view of what follows, and its generalisation to higher dimensions, it is convenient to introduce the notation

$$
\bar{E}_i = \mathbb{E}[E_i] \quad \text{and} \quad \sigma_{i\,j} = \mathrm{covar}[E_i, E_j],
$$

where i and j range over U, the unhedged portfolio, S, the swap, and C, the cap, and then define the vectors $\boldsymbol{E}$, $\boldsymbol{N}$ and $\boldsymbol{\sigma}_1$ by

$$
\boldsymbol{E} = \begin{pmatrix} \bar{E}_{\mathrm{S}} \\ \bar{E}_{\mathrm{C}} \end{pmatrix}, \quad \boldsymbol{N} = \begin{pmatrix} N^S \\ N^C \end{pmatrix} \quad \text{and} \quad \boldsymbol{\sigma}_1 = \begin{pmatrix} \sigma_{\mathrm{US}} \\ \sigma_{\mathrm{UC}} \end{pmatrix}
$$

and the symmetric matrix $\boldsymbol{\Sigma}$ by

$$
\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{\mathrm{SS}} & \sigma_{\mathrm{SC}} \\ \sigma_{\mathrm{SC}} & \sigma_{\mathrm{CC}} \end{pmatrix}.
$$

With this notation we can write

$$
\bar{E}_M = \bar{E}_{\mathrm{U}} + \boldsymbol{E}\cdot\boldsymbol{N} \quad \text{and} \quad \sigma_M^2 = \sigma_{\mathrm{U}}^2 + 2\,\boldsymbol{N}\cdot\boldsymbol{\sigma}_1 + \boldsymbol{N}\cdot\boldsymbol{\Sigma}\,\boldsymbol{N}. \tag{4}
$$

Note that as $\boldsymbol{\Sigma}$ is a covariance matrix, it is strictly positive definite (unless one or more of the hedging instruments are redundant and we shall assume that is not the case). It follows that σ_M^2 can be expressed as a positive-definite quadratic form in $\boldsymbol{N}$ and therefore has a unique minimum. Indeed, we can write (4) as

$$
\sigma_M^2 = \sigma_{\mathrm{U}}^2 - \boldsymbol{\sigma}_1\cdot\boldsymbol{\Sigma}^{-1}\boldsymbol{\sigma}_1 + (\boldsymbol{\sigma}_1 + \boldsymbol{\Sigma}\,\boldsymbol{N})\cdot\boldsymbol{\Sigma}^{-1}\,(\boldsymbol{\sigma}_1 + \boldsymbol{\Sigma}\,\boldsymbol{N}),
$$

[9] Again, it is important to recall that because our data set has the load normalised to give a monthly average of unity, one of the swaps or caps here applies to the average monthly load, not a single unit of real load.

from which it follows that $\sigma_M^2 \geq \sigma_{\min}^2$ where the minimum possible variance, $\sigma_{\min}^2$, is given by

$$\sigma_{\min}^2 = \sigma_{\mathrm{U}}^2 - \boldsymbol{\sigma}_1 \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_1. \tag{5}$$

We have $\sigma_M^2 = \sigma_{\min}^2$ if and only if $\boldsymbol{N} = \boldsymbol{N}_{\min}$, where $\boldsymbol{N}_{\min} = -\boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_1$. This is important in what follows. Given our assumption that E_{U}, E_{S} and E_{S} are drawn from a trivariate normal distribution, the distribution of E_M is normal with mean $\bar{E}_M$ and variance σ_M^2, i.e.,

$$E_M \sim \mathcal{N}(\bar{E}_M, \sigma_M^2).$$

Trivially, the expected value of this distribution is $\bar{E}_M$. The 5%-quantile is determined by

$$\frac{1}{\sqrt{2\pi\, \sigma_M^2}} \int_{-\infty}^{q_5} e^{-(x-\bar{E}_M)^2/2\sigma_M^2}\, dx = 0.05,$$

which, on making the change of variables $y = (x - \bar{E}_M)/\sigma_M$ reduces to

$$\Phi\left(\frac{q_5 - \bar{E}_M}{\sigma_M}\right) = 0.05,$$

where $\Phi(\,\cdot\,)$ denotes the standard normal cumulative function. The solution is

$$q_5 = \bar{E}_M + \Phi^{-1}(0.05)\; \sigma_M \approx \bar{E}_M - 1.6449\; \sigma_M \tag{6}$$

so $\mathrm{EaR} = \bar{E}_M - q_5$ may be approximated by

$$\mathrm{EaR} \approx 1.6449\; \sigma_M. \tag{7}$$

As before, the constraint $\mathbf{E}[E_M] \geq 0$ reduces to the linear constraint

$$\bar{E}_U + \boldsymbol{N} \cdot \boldsymbol{E} \geq 0$$

while the constraint $\mathrm{EaR} \leq \epsilon$ reduces to a constraint on the variance of the monthly earnings distribution,

$$\sigma_M^2 \leq \hat{\epsilon}^2 = \left(\frac{\epsilon}{1.6449}\right)^2. \tag{8}$$

In view of (4) and (5), we must insist that[10]

$$\hat{\epsilon}^2 > \sigma_{\min}^2 \quad \text{or} \quad \epsilon > 1.6449\, \sigma_{\min}, \tag{9}$$

or there will be no feasible region at all.

Looked at from a geometric point of view, since $\boldsymbol{\Sigma}$ is strictly positive definite, the level curves $\sigma_M^2 = c^2$ are ellipses in $\boldsymbol{N}$-parameter space and therefore if the region $\sigma_M^2 \leq \hat{\epsilon}^2$ is non-empty it consists of the boundary and interior of an ellipse. The feasible region in parameter space is non-empty *provided* that this ellipse exists *and* at least some of it lies in the half-plane $\bar{E}_M \geq 0$. This idea is illustrated in Figure 8b.

[10]The reason for strict inequality here is that if we have equality then there is only one value of $\boldsymbol{N}$ consistent with the EaR constraint, namely $\boldsymbol{N} = \boldsymbol{N}_{\min}$. So, *any* optimization problem we formulate in this case reduces to a trivial check to see whether or not the condition

$$\bar{E}_M = \bar{E}_{\mathrm{U}} - \tfrac{1}{2}\, \boldsymbol{E} \cdot \Sigma^{-1} \boldsymbol{\sigma}_1 \geq 0$$

is satisfied.

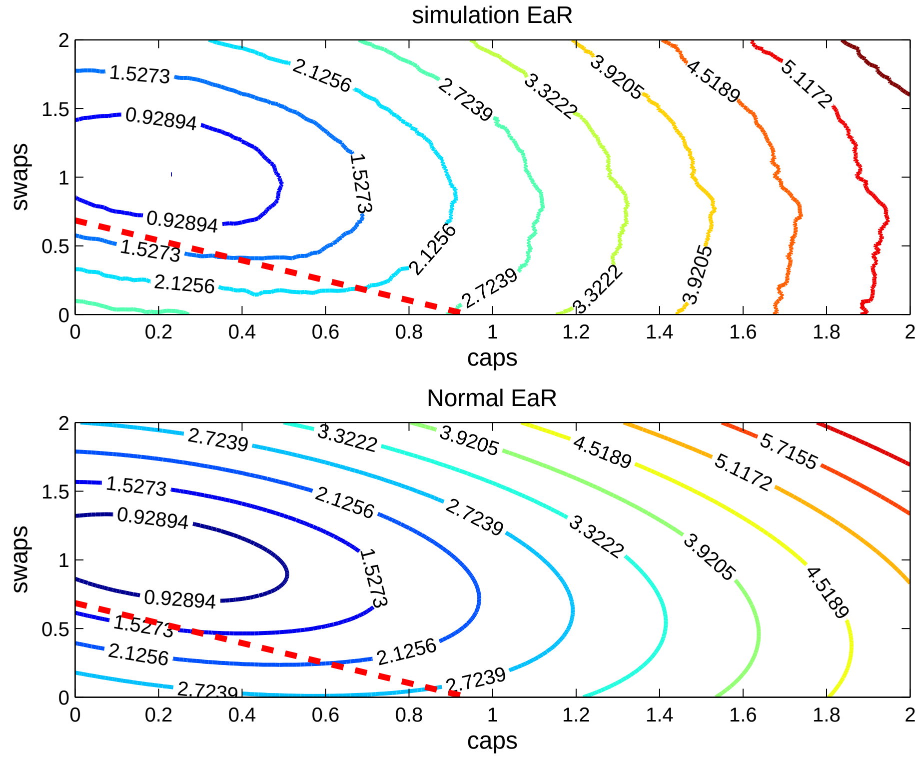

Figure 8: Comparison of the level curves of EaR computed directly from simulated data and with the assumption of normality together with (4) and (7). The dashed red line corresponds to zero expected monthly earnings, the region to the right and above it corresponds to positive expected monthly earnings, while to the left and below we have negative expected monthly earnings.

In N-parameter space, the constraint that $\mathbb{E}[E_M] \geq 0$ is unaffected by our assumption about the joint distribution of the earnings, but the constraint on the earnings at risk is affected. If Figure 8 we compare the level curves of the EaR derived directly from the simulated data (for January 2008) with the level curves of the EaR derived from (4) and (7). On the one hand, there is clearly qualitative agreement between the two sets, although there are noticable quantitative differences; given the unimodal but strongly leptokurtotic nature of the simulated distributions, this is hardly surprising. On the other hand, the simulated EaR contours depend strongly on the 5% quantiles and these are computed from a sample of size $3,000$ — therefore it would be unwise to regard the contours in Figure 8a as being particularly accurate.

5 Optimisation under the multi-variate normal approximation

There are many definitions of optimal in general use. It is not the purpose of this report to suggest which particular definition best meets a retailer's needs. Rather, we present two common alternatives and illustrate their implementation in our analytic framework.

5.1 Maximise expected monthly earnings

In this approach, we attempt to find the number of hedging instruments that will maximise the expected monthly earnings consistent with the constraint on the EaR. The problem is to maximise

$$\bar{E}_M = \bar{E}_{\mathrm{U}} + \boldsymbol{N} \cdot \boldsymbol{E},$$

subject to the constraint that

$$\sigma_M^2 = \sigma_{\mathrm{U}}^2 + 2\,\boldsymbol{N} \cdot \boldsymbol{\sigma}_1 + \boldsymbol{N} \cdot \boldsymbol{\Sigma}\,\boldsymbol{N} \leq \hat{\epsilon}^2.$$

As the objective function, $\bar{E}_M$, here is linear in the control variables, $\boldsymbol{N}$, and the constraint defines a convex region in $\boldsymbol{N}$-space it follows that the maximum occurs on the boundary of the convex region. Therefore, we can replace the constraint by

$$\sigma_M^2 = \sigma_{\mathrm{U}}^2 + 2\,\boldsymbol{N} \cdot \boldsymbol{\sigma}_1 + \boldsymbol{N} \cdot \boldsymbol{\Sigma}\,\boldsymbol{N} = \hat{\epsilon}^2.$$

This allows us to introduce a Lagrange multiplier, λ, and a Lagrange function $\mathcal{L}_1$ defined as

$$\mathcal{L}_1 = \bar{E}_M + \lambda(\sigma_M^2 - \hat{\epsilon}^2).$$

The stationary points of the Lagrange function are determined by

$$\nabla_N\,\mathcal{L}_1 = \mathbf{0} \quad \text{and} \quad \frac{\partial \mathcal{L}_1}{\partial \lambda} = 0$$

where ∇_N represents the $\boldsymbol{N}$-gradient operator. Since we have

$$\nabla_N\,\bar{E}_M = \boldsymbol{E}, \quad \nabla_N\,\sigma_M^2 = 2\,(\boldsymbol{\sigma}_1 + \boldsymbol{\Sigma}\,\boldsymbol{N}), \tag{10}$$

the first of these gives the linear system

$$\boldsymbol{E} + 2\,\lambda\,(\boldsymbol{\sigma}_1 + \boldsymbol{\Sigma}\,\boldsymbol{N}) = \mathbf{0}$$

and the second reduces to our constraint (with equality) $\sigma_M^2 = \hat{\epsilon}^2$. Solving the linear system gives

$$\boldsymbol{N} = -\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\sigma}_1 + \tfrac{1}{2\lambda}\,\boldsymbol{E}\right)$$

and substituting this into the constraint (with equality) gives

$$\frac{1}{2\,\lambda} = \pm\sqrt{\frac{\hat{\epsilon}^2 - \sigma_{\mathrm{min}}^2}{\boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1}\boldsymbol{E}}}\,.$$

As we have

$$\boldsymbol{E} \cdot \boldsymbol{N} = -\left(\boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1}\,\boldsymbol{\sigma}_1 + \tfrac{1}{2\lambda}\,\boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1}\,\boldsymbol{E}\right),$$

and $\boldsymbol{\Sigma}^{-1}$ is positive definite, we must take the *negative* value of λ in order to find the maximum expected monthly earnings, which after some algebra turns out to be

$$\bar{E}_M^{\mathrm{max}} = \bar{E}_{\mathrm{U}} - \boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1}\,\boldsymbol{\sigma}_1 + \sqrt{(\boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1}\,\boldsymbol{E})\,(\hat{\epsilon}^2 - \sigma_{\mathrm{min}}^2)}\,. \tag{11}$$

There is, of course, no guarantee that this value of $\bar{E}_M^{\mathrm{max}}$ is positive. This must be checked *a posteriori*. If it is not positive, then the constraint on the EaR is too tight and is not consistent with the constraint that the expected earnings be positive; i.e., there is no feasible region. In this case, in order to obtain positive monthly expected earnings the constraint on the EaR *must* be relaxed.

The EaR for this portfolio is the maximum allowed value, of ϵ^2, which also corresponds to the largest variance portfolio consistent with the EaR constraint. If we quantify risk by standard deviation then, put loosely, this shows that in order to obtain the maximum (expected) return we have to take the largest possible risk.

It is clear from (11) that increasing the bound on the EaR, ϵ, increases the expected monthly earnings. This is to be expected and comes at the price of increasing the standard deviation of the monthly earnings — put roughly, the more risk that is taken, the higher the expected earnings.

For our particular example, we find that

$$\bar{E}_{\mathrm{U}} = -1.8511, \qquad \sigma_{\mathrm{U}}^2 = 3.8584,$$

$$\boldsymbol{E} = \begin{pmatrix} 2.7023 \\ 1.9690 \end{pmatrix}, \qquad \sigma_1 = \begin{pmatrix} -3.3931 \\ -1.8381 \end{pmatrix}$$

and

$$\boldsymbol{\Sigma} = \begin{pmatrix} 3.1103 & 1.2013 \\ 1.2013 & 3.1683 \end{pmatrix},$$

from which it follows that

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} 0.3767 & -0.1428 \\ -0.1428 & 0.3698 \end{pmatrix},$$

and $\sigma_{\min}^2 = 0.0539$. This implies that our upper bound on the EaR, ϵ, must satisfy $\epsilon > 0.3819$. If we choose a bound of $\epsilon = 1$, that is if we insist that EaR ≤ 1 then

$$\hat{\epsilon}^2 = \left(\frac{\epsilon}{1.6449}\right)^2 = 0.6080,$$

and we find that $\lambda = -1.4526$ and that the optimal number of swaps and caps[11] is

$$\boldsymbol{N} = \begin{pmatrix} N^S \\ N^C \end{pmatrix} = \begin{pmatrix} 1.2691 \\ 0.3129 \end{pmatrix},$$

which give a maximum expected monthly earning of $\bar{E}_M^{\max} = 2.1946$ with a standard deviation of monthly earnings and EaR of $\sigma = 0.6080$ and EaR $= 1$. These numbers, of course, are based on the assumption of a trivariate normal distribution for the unhedged, swap and cap monthly earnings. This is an approximation. It is instructive to compare the monthly earnings distribution and its statistics with those we obtain if we use the price-load simulation directly. If we assume a portfolio of the composition given above and compute the corresponding

[11] As noted earlier, as the load is normalised to have a monthly average of unity, these should be interpreted as the percentage of monthly load that needs to be hedged with swaps and caps, respectively. The numerical values of these quantities should not be taken too seriously in any event, as we have assumed that constant swap and cap levels must be applied for *every* tick of the entire month and therefore the only way to hedge against a high pool-price coinciding with an above average load is to have more than the monthly average load covered by a swap or a cap.

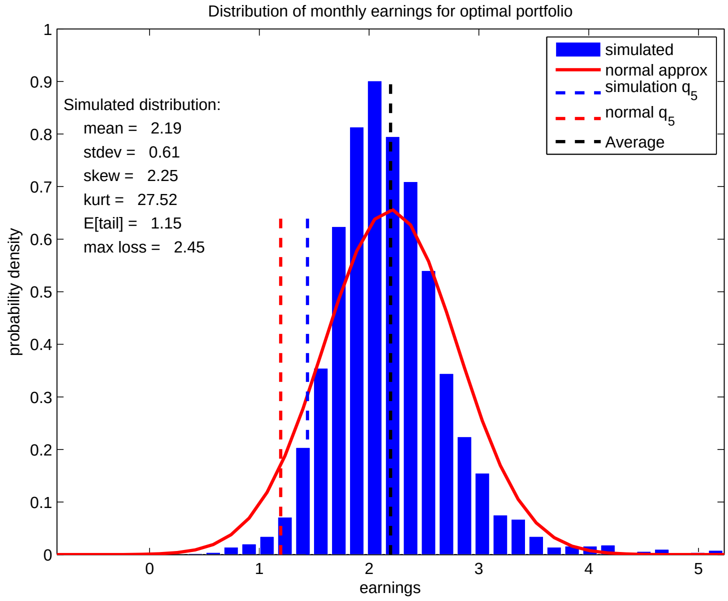

Figure 9: Normal and simulated distributions of the monthly earnings for the portfolio which maximises expected monthly earnings on the basis of the multivariate normal approximation.

statistics using the simulated price-load data (for January 2008), rather than assuming a trivariate normal distribution, we find

$$\bar{E}_M^{\text{sim}} = 2.1946, \quad \sigma^{\text{sim}} = 0.6080 \quad \text{and} \quad \text{EaR}^{\text{sim}} = 0.7562.$$

Given the leptokurtotic nature of the simulated distributions, one would expect that the EaR based on the trivariate-normal approximation is greater than the EaR computed from the simulated data, as it indeed is.

In Figure 9 we show the distribution of earnings based on a hedged portfolio using these numbers of swaps and caps and the January 2008 data set. Note that we have also reported two statistics associated with the tail (that is, below the 5%-quantile); the maximum loss, which is simply minus the smallest simulated monthly earning, and the expected monthly income in the tail, which is the expected value of the monthly earnings conditional on those earnings being below the 5%-tail. These, and similar statistics, should be considered before accepting the portfolio composition computed above, or by any other means that involves an optimization of a small set of statistics.

In this simple example, with only two control variables (N^S and N^C), we can also use the simulated data directly to estimate the composition of the portfolio that maximises the expected monthly earnings subject to the constraint that the EaR ≤ 1. In (N^S, N^C)-parameter space, the expected monthly earnings are represented by straight lines and using the simulated data we find that the level curves EaR $= c$ are closed and approximately convex curves. Therefore, all we need do to find the optimal portfolio is find the straight line that is tangent to the level curve EaR $= 1$ (at the top, the one that is tangent at the bottom represents the *worst* monthly expected earnings consistent with the EaR constraint). The coordinates of the point of tangency then gives us the composition of the optimal portfolio. This is illustrated

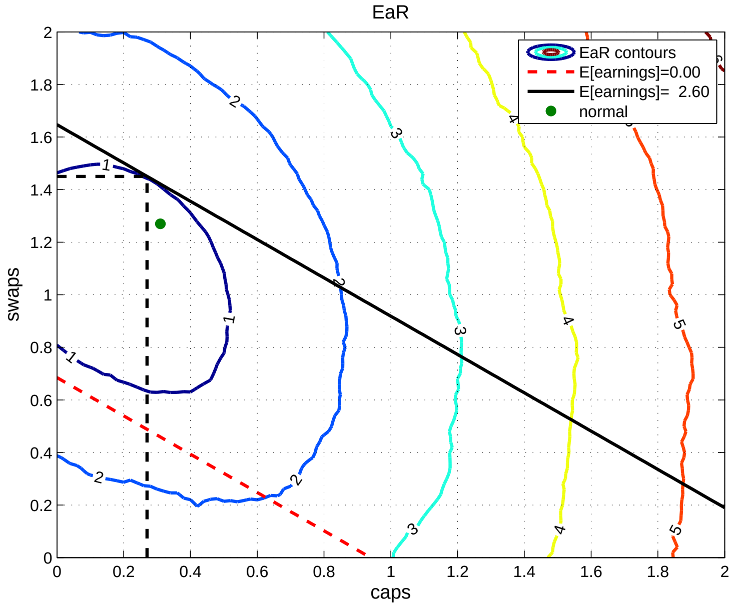

Figure 10: Level curves of EaR $= c_1$ and $\bar{E}_M = c_2$. The portfolio which maximises expected earnings corresponds to the topmost straight line $\bar{E}_M = c$ that is tangent to the the level curve EaR $= \epsilon$, where ϵ is the upper bound on the EaR. The composition of the portfolio corresponds to the coordinates of the point of tangency. The composition of the optimal portfolio calculated assuming a multivariate normal distribution is shown as a (green) dot.

in Figure 10 and we find that, using the simulated data without any assumptions about the normality of the distribution

$$N^S_{\text{opt}} \approx 1.45 \quad \text{and} \quad N^C_{\text{opt}} \approx 0.27 \;\; \text{corresponding to} \;\; \bar{E}^{\text{opt}} \approx 2.60 \quad \text{and} \quad \text{EaR} \approx 1.$$

Although the two approaches give distinct answers, they are reasonably close to each other. In particular, the analytic results derived here on the basis of a multivariate normal distribution could be regarded as either a starting point for a numerical optimization based entirely on simulated data (see later) or as the leading-order term in an asymptotic expansion, along the lines of a Cornish-Fisher expansion.

The distribution of expected monthly earnings corresponding to this composition of the hedging portfolio is shown in Figure 11, which should be compared with Figure 9.

5.2 Mean-variance approach

This is essentially the approach used in classical Markowitz portfolio theory. In this approach we attempt to find the minimum variance portfolio for a *given* level of expected return. This does *not* produce a unique optimal portfolio, rather it gives the optimal portfolio composition as a function of the given level of expected monthly earnings.

The constraint that the expected monthly earnings is positive implies that we only ever specify positive expected earnings. Slightly less trivially, the constraint on the EaR does not need to be explicitly imposed, as the EaR is proportional to the portfolio variance, and

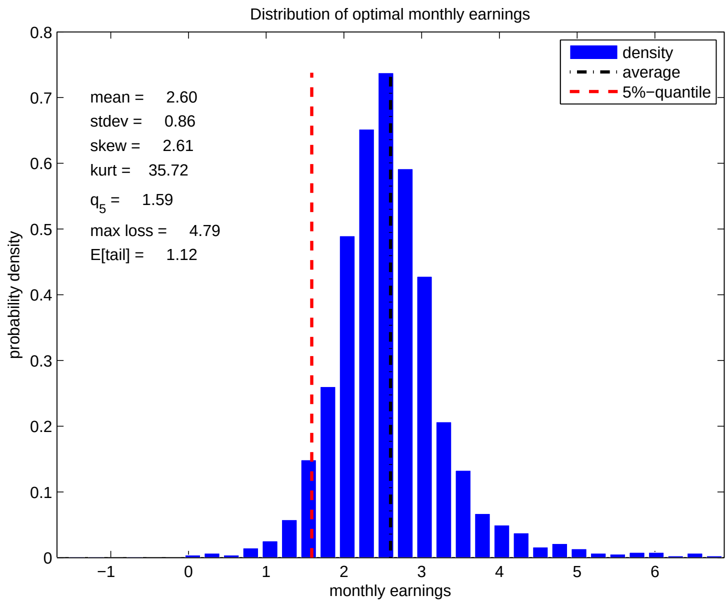

Figure 11: Simulated distribution of the monthly earnings for the portfolio which maximises expected monthly earnings computed directly from the simulated data (January 2008). Compare this with Figure 9.

therefore all we need do is check, *a posteriori,* whether or not the variance of the minimum variance portfolio is consistent with this constraint or not — this of course is only valid for the multivariate normal approximation we are using here. If, however, we implement a numerical optimization based only on price-load simulations, we can still check whether or not the EaR condition is met *a posteriori*, and as it seems plausible that the EaR is an increasing function of the variance so long as the distribution remains unimodal this would still imply a limit on largest minimum variance. Alternatively, in such an optimization we could eliminate the variance altogether and find the minimum EaR portfolio for a given level of expected returns; as noted above, these two possible approaches are identical under our multivariate normal assumption.

The analysis is similar to that for the maximum expected earnings analysis given above. In this frame work, we regard $\bar{E}_M > 0$ as prescribed, say $\bar{E}_M = \mu$ where μ is given, and apply the condition

$$\mu = \bar{E}_M = \bar{E}_{\mathrm{U}} + \boldsymbol{N} \cdot \boldsymbol{E} \tag{12}$$

as a constraint. The objective is to minimise the portfolio variance

$$\sigma_M^2 = \sigma_{\mathrm{U}}^2 + 2\,\boldsymbol{N} \cdot \boldsymbol{\sigma}_1 + \boldsymbol{N} \cdot \boldsymbol{\Sigma}\,\boldsymbol{N} \tag{13}$$

by varying $\boldsymbol{N}$, subject to the constraint (12).

We first introduce a Lagrange multiplier, λ, and define our Lagrange function $\mathcal{L}_2$ by

$$\begin{aligned} \mathcal{L}_2 &= \sigma_M^2 + \lambda\,(\bar{E}_M - \mu) \\ &= \sigma_{\mathrm{U}}^2 + 2\,\boldsymbol{N} \cdot \boldsymbol{\sigma}_1 + \boldsymbol{N} \cdot \boldsymbol{\Sigma}\,\boldsymbol{N} + \lambda\left(\bar{E}_{\mathrm{U}} + \boldsymbol{N} \cdot \boldsymbol{E} - \mu\right). \end{aligned}$$

We then find the stationary points of $\mathcal{L}_2$, with respect to $\boldsymbol{N}$ and λ, by setting

$$\nabla_N \mathcal{L}_2 = \mathbf{0} \quad \text{and} \quad \frac{\partial \mathcal{L}_2}{\partial \lambda} = 0.$$

It is relatively easy to show that provided the covariance matrix of the entire system,

$$\begin{pmatrix} \sigma_{\mathrm{U}}^2 & \boldsymbol{\sigma}_1^T \\ \boldsymbol{\sigma}_1 & \boldsymbol{\Sigma} \end{pmatrix},$$

is strictly positive definite, there is a unique stationary point and it corresponds to a minimum. The covariance matrix of the entire system will be strictly positive definite provided none of the hedging instruments is redundant (i.e., provided none of the hedging instruments can be exactly replicated using the others).

Setting the $\boldsymbol{N}$-gradient of $\mathcal{L}_2$ to zero gives the linear system

$$\lambda \boldsymbol{E} + 2\left(\boldsymbol{\sigma}_1 + \boldsymbol{\Sigma}\, \boldsymbol{N}\right) = \mathbf{0},$$

from which it follows that $\boldsymbol{N}$ is given by

$$\boldsymbol{N} = -\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\sigma}_1 + \tfrac{1}{2}\lambda \boldsymbol{E}\right). \tag{14}$$

The covariance matrix $\boldsymbol{\Sigma}$ is invertible provided none of the hedging instruments is redundant.

Setting $\partial \mathcal{L}_2 / \partial \lambda = 0$ simply reproduces the constraint (12), which then becomes

$$\mu = \bar{E}_{\mathrm{U}} - \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\sigma}_1 + \tfrac{1}{2}\lambda \boldsymbol{E}\right) \cdot \boldsymbol{E},$$

from which it follows that λ is determined by

$$\tfrac{1}{2}\lambda = \frac{\bar{E}_{\mathrm{U}} - \mu - \boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_1}{\boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{E}}. \tag{15}$$

Note that unless $\boldsymbol{E} = \mathbf{0}$, the denominator in this expression is positive and that if $\boldsymbol{E} = \mathbf{0}$ then the hedge is useless anyway as the hedging instruments have, on average, no effect on the expected earnings.

Therefore, for a prescribed level of expected earnings, the optimal portfolio's composition is determined by (14) and (15) and is

$$\boldsymbol{N}^*(\mu) = \left(\frac{\mu + \boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_1 - \bar{E}_{\mathrm{U}}}{\boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{E}}\right) \boldsymbol{\Sigma}^{-1} \boldsymbol{E} - \boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_1. \tag{16}$$

The variance of this portfolio is determined from (13) with $\boldsymbol{N} = \boldsymbol{N}^*(\mu)$, and is

$$\begin{aligned} {\sigma_M^*}^2(\mu) &= \sigma_{\mathrm{U}}^2 - \boldsymbol{\sigma}_1 \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_1 + \frac{\left(\mu + \boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_1 - \bar{E}_{\mathrm{U}}\right)^2}{\boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{E}} \\ &= \sigma_{\min}^2 + \frac{\left(\mu + \boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_1 - \bar{E}_{\mathrm{U}}\right)^2}{\boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{E}}. \end{aligned} \tag{17}$$

This shows, incidentally, that the minimum possible variance portfolio occurs when $\bar{E}_M^* = \mu^* = \bar{E}_{\mathrm{U}} - \boldsymbol{E} \cdot \boldsymbol{\Sigma}^{-1} \boldsymbol{\sigma}_1$, consistent with the results in the previous subsection. It would, of course, be foolish to specify $\mu < \mu^*$ as this would in effect give the *worst* possible expected return consistent with the variance ${\sigma_M^*}^2(\mu)$ — for any $c > \sigma_{\min}^2$ there are two possible values of μ, say μ_1 and μ_2, which give

$$\sigma_M^*(\mu) = c,$$

and they satisfy $\mu_1 < \mu^* < \mu_2$. To *maximise* the expected return, we always choose μ_2 for a given value of c. It is plausible that this situation persists if we carry out the variance (or EaR) minimization using only the simulated price-load data, without making the assumption of a multivariate normal distribution and therefore the specified level of expected earnings needs some thought.

A plot of ${\sigma_M^*}^2(\mu)$ as a function of μ, for the same data as the previous subsection, is shown in Figure 12. This figure also indicates the region consistent with the constraints $\bar{E}_M = \mu \geq 0$ and EaR ≤ 1 and shows the corresponding portfolio composition, $(N^S(\mu), N^C(\mu))$. As in the previous subsection, these results are based on the assumption of a multivariate normal distribution for the unhedged, swap and cap earnings and are therefore only approximations.

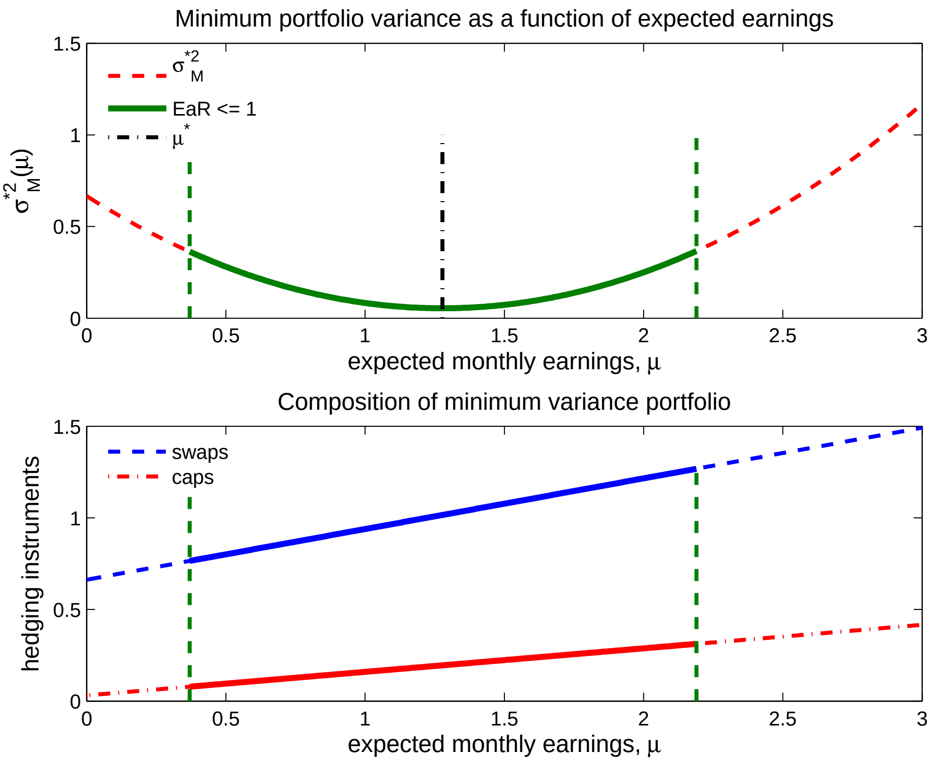

Figure 12: Minimum variance consistent with a given level of expected return, μ, and the portfolio compositions that produces it, also as functions of μ. The results are based on the assumption of a multivariate normal distribution for earnings and are therefore only approximate.

5.3 Expected utility of earnings

It is, in principle, possible to repeat the analyses given above using the expected utility of the earnings, instead of the expected earnings, assuming a utility of wealth function which satisfies the usual conditions of non-satiation, $U'(W) > 0$, and risk-aversion[12] $U''(W) < 0$ and a multivariate normal distribution of unhedged, swap and cap earnings.

The main problems with this approach are:

- it requires a utility-of-wealth function $U(W)$ and the results will depend on it;
- it is only analytically tractable, for the multivariate normal distribution, if the utility-of-wealth function is of the constant-absolute-risk-aversion (CARA) form

$$U(W) = \frac{1}{\gamma}\left(1 - e^{-\gamma W}\right),$$

  where $\gamma > 0$. Some progress can be made for other forms, but various transcendental equations involving multidimensional integrals need to be solved numerically in these cases. The overhead in doing so negates the advantages of analytic tractability stemming from the multivariate normal approximation and one may as well do the whole problem numerically, without making the assumption of normality, in this case.

The details for the CARA function are not given in this report as they do not add anything new to the methodology.

6 Numerical methods

Provided we optimise the portfolio by varying the *numbers* of hedging instruments chosen from a given set of available instruments, in principle it should not be difficult to perform a numerical optimization using only simulated price-load data, *provided the level surfaces of the EaR are convex.* It seems plausible that the level surfaces will be convex provided all the earnings distributions are unimodal. The main numerical problem likely to be encountered is that, as a result of sampling error, the EaR level surfaces inferred from a finite number of price-load simulations may not be convex — see Figures 7b, 8a and 10 where what is clearly sampling error in the estimation of the quantiles is producing wobbles in the EaR level surfaces, for example. These spurious wobbles mean the level curves are not convex and this implies that there may be spurious internal optima for an EaR-constrained problem. This may make it necessary to smooth these level surfaces before attempting to optimise.

Attempts to optimise over the parameters of the hedging instruments (such as swap-level or cap-level) will necessarily be far less computationally efficient; partly because the total monthly earning cease to be linear functions of the control variables, partly because the earnings from an instrument needs to recomputed each time one of its parameters is changed, partly because even if the EaR level surfaces remain convex it is by no means clear that optima can not occur in the interior of these if the monthly earning are not a linear function of control variables (or, more correctly, if they are not a monotonic function of a linear combination of the control variables) and partly because it is also by no means clear that the level surfaces will be convex in this case.

[12] This condition is necessary in order that a maximum expected utility exists. If, for example, we replace it with the "risk-seeking" condition, $U''(W) > 0$, we find that the stationary points of the expected utility turn out to be local minima or saddles and that no maximum exists.

7 Conclusion

Restricting the optimization to the choice of the number of already available contracts makes the problem tractable. The constraints on expected monthly earnings and the EaR need to be considered carefully, both to ensure they are consistent with the hedging products available and to ensure that they do not force suboptimal choices.

Some care needs to be taken when deciding on a definition of optimal and, whatever definition is chosen it would be wise to check tail distributions before proceeding. The tail, by definition, contains the most extreme events and, although relatively improbable, if they occur these are the events that can cause *serious* problems. The risk inherent in any portfolio can only be fully described by the *distribution* of earnings, not by a single statistic, whether or not that single statistic is optimised.