Practicalities of Pricing Exotic Derivatives

John Crosby

Grizzly Bear Capital

My website is: http://www.john-crosby.co.uk
If you spot any typos or errors, please email me.
My email address is on my website.

Lecture given 1st July 2014 for Module 4
of the M.Sc. course in Mathematical Finance
at the Mathematical Institute at Oxford University.

File date 11th June 2014 at 10.30

- Your course of studies has now shown you how to price several types of exotics analytically. You have studied generic pricing techniques such as Monte Carlo simulation and PDE lattice methods which can, potentially, price many different types of exotic options under many different types of stochastic processes.

- So pricing exotic options is easy, right?

- We are given the payoff of the exotic option by our trader. We choose our favourite stochastic process. Naturally, to show how clever we are, it has stochastic vol, stochastic skew, jumps, local volatility, stochastic interest-rates,...

- We calibrate our model to the market prices of vanilla (standard European) options. Naturally, the fit to vanillas is wonderful which just reinforces how great our model must be.

- We price the exotic option in question. The trader now has an arbitrage-free price - so nothing could possibly be wrong with it (!). The trader puts a bid-offer spread around our price (usually the spread is 10 per cent or more of the price). The sales person may add on a further sales margin (which may be another 10 per cent or more).

Motivation 2

- Everyone is happy.
- Nothing could go wrong, could it?

Motivation 3

- Everyone is happy.
- Nothing could go wrong, could it?
- Well, actually lots of things can go wrong.
- Pricing a book of exotics is not easy. It can sometimes be an art as well as a science.
- I will attempt to highlight a few of the problems - and a few possible solutions.

Possible problems

- Talk outline:
- Robust pricing of exotics.
- Choice of model.
- Vanilla and barrier fx options.
- Model calibration.

•

- Because the price of exotic options can often be highly sensitive to the chosen model, *robust* pricing methodologies are very attractive. A *robust* pricing methodology can, loosely, be defined as one which works for two or more (classes of) models. Because it works for two or more models, there is robustness against model error (either misspecification of the type of stochastic process and/or misestimation of the parameters of the process).

- An example of this is as follows:

- Consider a binary cash-or-nothing call option (sometimes called a European digital option) which pays one dollar if the stock price S_T at expiry T is greater than or equal to the strike K and it pays zero otherwise.

- Suppose we form a portfolio of long $1/(2\Delta K)$ vanilla call options with strike $K - \Delta K$ and short $1/(2\Delta K)$ vanilla call options with strike $K + \Delta K$, where $\Delta K > 0$.

- The payoff of the portfolio at expiry is:
  $1/(2\Delta K)[S_T - (K - \Delta K)] - 1/(2\Delta K)[S_T - (K + \Delta K)] = 1$ if $S_T \geq K + \Delta K$,
  0 if $S_T < K - \Delta K$,
  $0 < 1/(2\Delta K)[S_T - (K - \Delta K)] < 1$ if $K - \Delta K \leq S_T < K + \Delta K$.

- If we consider the limit as $\Delta K \to 0$, then we see that the payoff of the portfolio replicates that of the binary cash-or-nothing call option.

- Let us denote by $V(K, \sigma(K))$ the price of a vanilla call option with strike K and with implied Black and Scholes (1973) volatility $\sigma(K)$.

- Then, in the absence of arbitrage, the price of the binary cash-or-nothing call option is:

$$\begin{aligned}\lim \Delta K \downarrow 0 \;\; \frac{V(K-\Delta K, \sigma(K-\Delta K)) - V(K+\Delta K, \sigma(K+\Delta K))}{2\Delta K} &= -\frac{dV}{dK} \\ &= -\frac{\partial V}{\partial K} - \frac{\partial V}{\partial \sigma}\frac{\partial \sigma}{\partial K}\end{aligned}$$

- Note that we can compute the term $-\frac{\partial V}{\partial K}$ by noting that it is the discount factor $\exp(-rT)$ multiplied by the $N(d_2)$ term in the Black and Scholes (1973) formula, evaluated with strike K and volatility $\sigma(K)$:

$$-\frac{\partial V}{\partial K} = \exp(-rT)N(d_2)$$

- We can compute the term $\frac{\partial V}{\partial \sigma}\frac{\partial \sigma}{\partial K}$ by noting it is the “vega” multiplied by the “volatility skew”.

- This formula gives us a non-parametric methodology for pricing binary cash-or-nothing call options in the presence of a volatility skew or smile.

- If the volatility surface is negatively-skewed as it is for equities, then the term $\frac{\partial V}{\partial \sigma}\frac{\partial \sigma}{\partial K}$ is negative and so the presence of the skew increases the price of a binary cash-or-nothing call option relative to its price in a (constant volatility) Black and Scholes (1973) world.

- The vanilla call spread gives us the hedge. In practice, there is a problem in that, in the limit $\Delta K \to 0$, we have to simultaneously both buy and sell an infinite quantity of vanilla options (!). Recognising that this is not possible, a safer strategy is sub- or super-replication.

- Suppose we write a binary cash-or-nothing call option with strike K (an analogous strategy can be used if we are buying the option) and there are vanilla options, with strikes K_1 and K_2, which we can respectively buy and sell in the market at prices $V(K_1, \sigma(K_1))$ and $V(K_2, \sigma(K_2))$, where $K_1 < K_2 < K$.

- Then if we buy $1/(K_2 - K_1)$ of the options with strike K_1 and we sell $1/(K_2 - K_1)$ of the options with strike K_2, we can super-replicate the binary cash-or-nothing call option payoff (i.e. we never lose money). The closer to K that K_2 is, the cheaper this super-replicating portfolio is. Note that in the event that it turns out that the stock price is between K_1 and K_2 at expiry, then we always make money on our strategy.

- On the other hand, if we sold at the initial cost of the super-replicating portfolio, the option buyer has paid over the odds. A proportion of the business of investment banks in the area of exotic options appears to consist of trying to find poorly-informed or less technical investors who make this mistake.

- There are potentially economies of scale related to the analysis above in the sense that if we already have vanilla options with strike K on our books, we may be able to use these as part of our hedge.

- This illustrates a general principle of hedging a book of exotic options which is that it is much easier if you already have a large vanilla options book.

- Note that we can get the price of a binary cash-or-nothing put option by a "put-call parity" relation (namely, that a binary cash-or-nothing call option plus a binary cash-or-nothing put option equals the discount factor to expiry).

Choice of model

•

- Two papers by Schoutens, Simons and Tistaert illustrate a general feature of pricing exotic options.

- Specifically, in the first paper, they calibrated seven different stochastic models to the market prices of 144 different vanilla options on the Eurostoxx 50 index with maturities ranging from less than a month to just over five years. The models were sophisticated models (HEST, HESTJ, BNS, VGCIR, VGOU, NIGCIR, NIGOU) with between five and eight free parameters.

- Any model with this number of parameters should be able to give a decent fit to the market prices of vanilla options and, indeed, the paper demonstrates this. We leave aside for now important issues such as what happens if the error function has many local minima, potential parameter instability arising and the effect of parameter instability through time on hedging performance (see later).

- Using Monte Carlo simulation, a number of different exotic options were priced.

- A brief selection of prices are:

  Lookback call options.

Model	$HEST$	$HESJ$	BNS	$VGCIR$	$VGOU$	$NIGCIR$	$NIGOU$
Price	838.48	845.19	771.28	724.80	713.49	730.84	722.34

Cliquet options with local floor at -0.03, local cap at 0.05 and global floor at -0.05.

Model	$HEST$	$HESJ$	BNS	$VGCIR$	$VGOU$	$NIGCIR$	$NIGOU$
Price	0.0729	0.0724	0.0788	0.1001	0.1131	0.0990	0.1092

All prices are quoted from Schoutens, Simons and Tistaert ("A Perfect Calibration! Now What?", Wilmott magazine, 2004 or see Wim Schouten's website).

- We can see that there is a large variation in the prices of these exotic options.
- In the case of the lookback options, the variation is more than 18 per cent.
- In the case of the cliquet options, the variation is more than 56 per cent.
- In the case of barrier options, results in Schoutens, Simons and Tistaert show that the variation can be more than 200 per cent (especially if the spot price is close to the barrier level).
- Now it is worth remembering that the bid-offer spreads will typically be greater than ten per cent (even much more for barrier options when the spot price is close to the barrier level).
- However, the variation in the exotic option prices over the seven different models is still such that the highest bid will be greater than the lowest offer.

- This means that, for example, a hedge fund could, in principle, make money by buying from the lowest offer and simultaneously selling at the highest bid if it knew that these models were being used by particular market-making investment banks.
- This is potentially a very dangerous situation for either market-making bank.
- This is called “adverse price discovery risk”.
- It is the risk you do a lot of business because your prices are out of line with other banks.
- Totem prices from MarkIT can help identify this risk.
- A number of banks (for example, UBS, Barclays Capital and others) provide “live” screen-based prices for first and second generation exotics. These prices can also be used to identify this risk.

- It is important to see where your prices lie in relation to other banks' prices.
- This is a variation on the saying of John Maynard Keynes: "It is better to be wrong conventionally than right unconventionally".
- This, in turn, means that models have to be chosen not only so as to match the market prices of vanilla options but also, where possible, to match the observable prices of exotic options. This may rule out the use of some models and/or require the use of different models for different types of exotic options and/or require the development of new types of models with the built-in flexibility to fit the market prices of both vanilla and exotic options.

•

- A paper by Lipton and McGhee illustrates some of these features.

- In the foreign exchange (fx) options markets, barrier options are very actively traded. Many different types of barrier options (single barrier, double barrier, with or without rebates, partial, window, etc) trade but, by far, the most actively traded barrier options are double-no-touch (DNT) options. These pay one unit of domestic currency at expiry if the spot fx rate (quoted as the number of units of domestic currency per unit of foreign currency) never trades equal to or outside a lower barrier level nor an upper barrier level. If either the lower barrier level or the upper barrier level are touched or breached prior to expiry, the option expires worthless.

- The market prices of DNT options on major currency pairs are widely available to market-making banks - either through inter-dealer brokers or through “live” screen prices. Hence, it is desirable that any model can fit these market prices.

- The market-standard working assumption made in pricing these DNT options is that the spot fx rate is monitored continuously to see if the barrier levels have been hit (in practice, monitoring is continuous from Monday morning in Sydney until Friday close in New York).

- Lipton and McGhee make a number of observations:
- If a Dupire (1994) local volatility model is calibrated to the market prices of vanilla options, and then used to price DNT options, the resulting model DNT option prices are less than the market prices i.e. a local volatility model tends to under-price DNT options relative to the market prices. For example, the model price may be around 50 to 75 per cent of the market price.
- If a stochastic volatility (for example, Heston (1993)) model is calibrated to the market prices of vanilla options, and then used to price DNT options, the resulting model DNT option prices are greater than the market prices i.e. a stochastic volatility model tends to over-price DNT options relative to the market prices. For example, the model price may be around 125 to 150 per cent of the market price.
- In either case, the degree of mis-pricing is well in excess of the bid-offer spread (which is usually around 2 to 4 per cent of notional - eg. if the mid-market price is 0.48, then the bid-offer spread might be 0.465 to 0.495).

- The above observations are fairly robust to which type of stochastic volatility model is chosen and which currency pair is chosen (at least for major currency pairs).

- Adding a compound Poisson jump process may give a more realistic fit to the short-term skew/smile but it doesn't change the broad conclusion regarding DNT option prices.

- Lipton and McGhee did not consider pure jump Levy processes such as CGMY but my experience is that, if calibrated to vanilla options, the CGMY model over-prices DNT options relative to market prices.

- The above observations lead Lipton and McGhee to propose their "Universal Volatility" model.

- Let us denote the spot fx rate, at time t, by $S(t) \equiv S$. Domestic (respectively, foreign) interest-rates are denoted by r_d (respectively, r_f). Lipton and McGhee proposed the following risk-neutral dynamics:

$$\begin{aligned} \frac{dS}{S} &= (r_d - r_f - \lambda m)dt + \sigma_L(t, S(t))\sqrt{V(t)}dz_S(t) + (\exp(J) - 1)dN(t), \\ dV(t) &= \kappa(\theta - V(t))dt + \epsilon\sqrt{V(t)}dz_V(t),\ V(0) \equiv V_0, \end{aligned}$$

  where λ is the intensity rate of the Poisson process $N(t)$, the Brownian motions $z_S(t)$ and $z_V(t)$ have constant correlation ρ, $m = E[\exp(J) - 1]$ and where $V(t)$ is a stochastic volatility (or stochastic variance to be more exact) and $\sigma_L(t, S(t))$ is a local volatility function.

- If $\sigma_L(t, S(t)) \equiv 1$, then the model is the same as Heston (1993) (plus jumps).

- If $\epsilon \equiv 0$, then the model is the same as Dupire (1994) (plus jumps).

- Either Heston (1993) (plus jumps or without jumps) or Dupire (1994) (plus jumps or without jumps) could (on their own) be calibrated to the market prices of vanilla options. The key issue is that we can "mix" the local volatility and the stochastic volatility. Essentially, by "tweaking" the vol-of-vol parameter ϵ to some value which is between that value it would take if $\sigma_L(t, S(t)) \equiv 1$ (the value it would take in a wholly stochastic volatility model) and zero (the value it would take in a wholly local volatility model), we find the value which can approximately match the market prices of DNT options.

- To put it another way, a local volatility model tends to under-price DNT options relative to the market prices, a stochastic volatility model tends to over-price DNT options relative to the market prices, so we should be able to "mix" the two models by "tweaking" the vol-of-vol parameter ϵ in such a way that we match the market prices of DNT options. (In practice, "tweaking" ϵ will likely cause other parameters to change somewhat).

- A few comments are in order:

- Lipton and McGhee was published in 2002. Since publication, this model has become a kind of market standard model and is used by many banks. It is quite common - in fact, nearly standard - to NOT have a jump term as it complicates the solution but this is to the detriment of fitting short-dated options.

- Parameterizations of the local volatility function of the form $\sigma_L(t, S(t))S(t) = a + bS(t)$ i.e. displaced diffusion do NOT seem to fit the market at all. One typically parameterises the local volatility function $\sigma_L(t, S(t)) = \alpha + \beta(S(t) - S(0)) + \gamma(S(t) - S(0))^2$, where α, β and γ are independent of $S(t)$.

- No useful (by which we mean except in limited special cases) analytic formulae exist for either vanilla option prices or barrier option prices. Therefore, all options are priced by numerically solving the relevant two-factor (plus time) PDE.

- One invariably uses term structures of interest-rates in both currencies which poses no extra difficulty in solving the PDE numerically.

- With an efficient ADI implementation, a whole grid of options can typically be priced in around 2 seconds.

- The model has a number of attractive features. However, it is not perfect. Bid-offer spreads have declined and volatility skews have become more severe (especially dollar/yen and even before the credit crunch) since the model was first introduced. Now, in order to closely match both vanilla and DNT (or other barrier option) prices, one needs to introduce time-dependent parameters.

- This involves making the parameters α, β and γ of the local volatility function as well as, at least, the vol-of-vol parameter ϵ time-dependent.

- Often, calibrations to real market data show these parameters varying in a way that is not credible or believable from a financial perspective. By the time one has finished, one might ask whether one has a model or a giant interpolation machine.

- Much more recently, an alternative model has been proposed by Dherminder Kainth. It produces excellent fits without the need for time-dependent parameters and very realistic (risk-neutral) dynamics. We now describe the Kainth model in more detail.

- The Kainth model utilises what he calls a "double Heston" process. Let us denote the spot fx rate, at time t, by $S(t) \equiv S$. Domestic (respectively, foreign) interest-rates are denoted by r_d (respectively, r_f). The risk-neutral dynamics are:

$$
\begin{aligned}
\frac{dS}{S} &= (r_d - r_f)dt + \sum_{i=1}^{2} \sqrt{V_i(t)}dW_i(t), \\
dV_1(t) &= \kappa_1(\theta_1 - V_1(t))dt + \epsilon_1\sqrt{V_1(t)}dz_1(t), \; V_1(0) \equiv V_{0,1}, \\
dV_2(t) &= \kappa_2(\theta_2 - V_2(t))dt + \epsilon_2\sqrt{V_2(t)}dz_2(t), \; V_2(0) \equiv V_{0,2}.
\end{aligned}
$$

The parameters κ_1, θ_1, ϵ_1, $V_{0,1}$, κ_2, θ_2, ϵ_2, $V_{0,2}$ are positive constants. There are two stochastic variance processes $V_1(t)$ and $V_2(t)$, with initial (time zero) values $V_{0,1}$ and $V_{0,2}$ respectively. Both stochastic variance processes drive the process for the spot fx rate. The processes $W_1(t)$, $W_2(t)$, $z_1(t)$ and $z_2(t)$ are standard Brownian motions with the following correlation structure:

$$E[dz_1(t)dW_1(t)] = \rho_1, E[dz_2(t)dW_2(t)] = \rho_2.$$

All other correlations are identically equal to zero. In particular,

$$E[dz_1(t)dz_2(t)] = 0, E[dW_1(t)dW_2(t)] = 0, E[dz_1(t)dW_2(t)] = 0, E[dz_2(t)dW_1(t)] = 0.$$

- The way the correlation structure is set up is important. In particular, it is part of the model design that it is intended that ρ_1 and ρ_2 will have opposite signs (eg. ρ_1 is negative and ρ_2 is positive).

- If we define $\Sigma^2 \equiv V_1(t) + V_2(t)$, then the dynamics of $S(t)$ can be written in the form:

$$d(\log S) = (r_d - r_f - \frac{1}{2}\Sigma^2)dt + \Sigma dW(t).$$

  This shows that what we have is a three-factor model.

- Both vanilla options and DNT options (or other types of barrier options or exotic options) can be priced by numerically solving a three-factor PDE (i.e. three factors plus time).

- However, for vanilla options, (much faster) Fourier inversion methods can also be used because the characteristic function can be computed in closed-form.

- Exotic options can also be priced by Monte Carlo simulation.
- Using term structures of interest-rates in both currencies poses no extra difficulty.
- One could also easily allow some of the parameters (eg. ϵ_1 and ϵ_2) to be time-dependent.
- However, that seems to be unnecessary. Kainth shows that, at least, for major currency pairs and the data-set considered, this double Heston model can give an excellent fit to the market prices of both vanilla and DNT fx options with all parameters constant.
- The double Heston model also has two additional attractive features:

- Kainth (2007) (talk given at the ICBI Global Derivatives conference, Paris, May 2007) gives an example of parameter values obtained from a calibration:

Process	First	Second
i	1	2
$\sqrt{V_{0,i}}$	0.09135	0.06599
$\sqrt{\theta_i}$	0.08583	0.05800
κ_i	2.132	5.493
ρ_i	−0.8646	0.9612

- Notice how this seems to suggest two distinct processes, the first with a low mean reversion rate and the second with a high mean reversion. The idea of two characteristic time-scales for mean reversion is certainly appealing.

- Furthermore, the first correlation is large and negative and the second correlation is large and positive.

- If we consider the process for $\Sigma^2 = V_1(t) + V_2(t)$, we find (from Ito's lemma) that:

$$d\Sigma = [\ldots]dt + \frac{\epsilon_1\sqrt{V_1(t)}}{2\sqrt{V_1(t)+V_2(t)}}dz_1(t) + \frac{\epsilon_2\sqrt{V_2(t)}}{2\sqrt{V_1(t)+V_2(t)}}dz_2(t). \tag{1}$$

- Furthermore, the instantaneous correlation between $d(\log S)$ and $d\Sigma$ is:

$$\frac{\epsilon_1\rho_1 V_1(t) + \epsilon_2\rho_2 V_2(t)}{2((V_1(t)+V_2(t))(\epsilon_1^2 V_1(t) + \epsilon_2^2 V_2(t)))^{1/2}}, \tag{2}$$

  which (loosely speaking) is proportional to $\epsilon_1\rho_1 V_1(t) + \epsilon_2\rho_2 V_2(t)$. Clearly, this instantaneous correlation is stochastic since it changes as $V_1(t)$ and $V_2(t)$ change.

- We have noted that we expect ρ_1 and ρ_2 to have opposite signs (borne out by the calibration above). Also, ϵ_1, $V_1(t)$, ϵ_2, $V_2(t)$ are all non-negative.

- We see that the instantaneous correlation between $d(\log S)$ and $d\Sigma$ is not only stochastic but can be negative or positive and can actually switch sign through time (provided, to repeat, ρ_1 and ρ_2 are of opposite sign).

- This is linked to the idea of stochastic skew.
- Stochastic skew was a term first coined in a paper by Carr and Wu in connection with foreign exchange (fx) options. Consider the implied volatility of a call option with a delta of 0.25 (and a given maturity) and subtract from it the implied volatility of a put option with a delta of -0.25 (and the same maturity). This is called the 25-delta risk-reversal and is very actively traded in the fx options markets.
- Carr and Wu document that, not only the magnitude, but, also, the sign of the 25-delta risk-reversal changes through time for major currency pairs. This is in contrast to the equity options markets where risk-reversals are always negative. Since the risk-reversal is essentially a measure of the skewness of the (risk-neutral) distribution, we see that in order to capture this feature, we need the instantaneous correlation between $d(\log S)$ and $d\Sigma$ to be able to change sign.

- Most models cannot capture this. For example, in the Heston (1993) model (i.e. single Heston), the instantaneous correlation between $d(\log S)$ and $d\Sigma$ always has the sign of the ρ parameter and cannot change sign (since ρ is a constant).
- In their paper, Carr and Wu propose a class of stochastic processes which can capture stochastic skew of which the double Heston model is a special case.
- The ability of the double Heston model to capture stochastic skew is a very attractive feature.
- Taking everything into account, the double Heston model is very appealing.
- Its main drawback is that one is very reliant on solving a three-factor (plus time) PDE numerically or Monte Carlo simulation for the pricing of exotic options (both of which will be relatively slow).

-

- The typical calibration problem can be described as follows:
- We are given a set of N market option prices (usually, they are all vanilla options) $C_i(K_i, T_i)$ of different strikes K_i and maturities T_i, for $i = 1, \ldots, N$.
- We are given a stochastic processes with M parameters labelled $\overline{\theta} = \{\theta_1, \theta_2, \ldots, \theta_M\}$ which generates N model prices $G_i(\overline{\theta}, K_i, T_i)$, for $i = 1, \ldots, N$. Note $M \leq N$.
- Then we use a "solver" algorithm (simplex, simulated annealing, Newton-Raphson, Broyden's method, etc) to find the values of the parameters $\overline{\theta}$ which minimise:

$$\sum_{i=1}^{N} \omega_i (G_i(\overline{\theta}, K_i, T_i) - C_i(K_i, T_i))^2$$

(One does not have to use the L^2-norm but it is usual).

- This inverse problem is not well-posed. In other words, there may be many local minima, one of which is located by the "solver", as opposed to the global minimum, the results may be highly sensitive to the starting point of the "solver" algorithm and small changes in $C_i(K_i, T_i)$, for some $i = 1, \ldots, N$ may result in large changes in $\overline{\theta}$.

- There are suprisingly few papers or books which even acknowledge the problem here. Exceptions include chapter 13 of Cont and Tankov (2003) and Cont and Tankov (2004). Cont and Tankov use the concept of relative entropy, which, in fullness, is beyond the scope of this lecture, and they focus mostly on jump-diffusion models.

- I will be less formal than Cont and Tankov and try to suggest some "rules-of-thumb".

- Changing the target function (i.e. the function to be minimised) may sometimes help:
- Choose the vanilla options to be puts or calls according to which is out of the money.
- One could choose the weights ω_i such that they are inversely proportional to the square of the bid-offer spreads. This gives more weight to the most liquid options. Alternatively, choose the weights ω_i such that they are inversely proportional to the square of the Black and Scholes (1973) "vegas". This is roughly equivalent to minimising the sum of squares of differences in implied volatilities (which is more financially relevant and interesting).

- Cont and Tankov's idea of relative entropy could be summarised non-mathematically: "Use what you know from historical data eg. real-world physical measure parameter estimates or parameter estimates from the previous day's calibration".

- Suppose we are calibrating a jump-diffusion model with one Poisson process with either Merton (1976) lognormal jumps or with fixed jump amplitudes. Cont and Tankov explain how calibrating such models is particularly prone to parameter instability because as one increases the intensity rate and decreases the mean jump size, the target function can move along a flat valley.

- We expect jumps to be rare, extreme events. Therefore, we do not expect there to be 100 jumps per year on the average. We would expect a few (eg. less than 5) jumps per year on the average. This gives a very rough estimate of the real-world physical measure jump intensity rate. More refined estimates could be obtained from time-series data (eg. recursive filtering - which is roughly counting the number of jumps per year bigger than x standard deviations for some choice of x).

- Real-world and risk-neutral intensity rates can be completely different in theory and will typically not be the same in practice. However, with very rough estimates of risk-aversion parameters, we can roughly go from one to the other. For equities, this implies risk-neutral intensity rates will typically be higher than real-world intensity rates. For commodities, empirical evidence suggests that risk-neutral intensity rates and real-world intensity rates are typically quite close to one another.

- Therefore, it should be possible to estimate a value λ_{est} for the risk-neutral intensity rate λ, before starting the calibration. Then one can incorporate this estimate into the calibration by use of a penalty function. Specifically, we minimise:

$$(w_\lambda(\lambda_{\text{est}} - \lambda)^2) + \sum_{i=1}^{N} \omega_i (G_i(\overline{\theta}, K_i, T_i) - C_i(K_i, T_i))^2,$$

  for some positive weight w_λ (which can be obtained with the help of a little trial-and-error based on how "tightly" we wish to keep the risk-neutral intensity rate to λ_{est}).

- One could also use time-series data to estimate the mean jump size (which if the jumps are random in size may, of course, be different in the real-world and risk-neutral measures). Indeed, in a similar vein to above, we would probably expect the estimates of the mean jump size to be equivalent (in magnitude) to, say, at least 2 standard deviation daily moves (the sign can be determined from the slope of the implied volatility as a function of strike).

- By incorporating historical data into the calibration, we increase the curvature of the target function to be minimised and so, intuitively, increase the chance of a more stable calibration.

- If we are recalibrating a model daily (as is usual), in some cases, we might use the estimate of a parameter from the previous day in a penalty function. This can be useful if we are worried about parameter instability - though at risk of deteriorating calibrations if repeated over long time periods.

- We might expect mean-reversion rates to reflect characteristic time-scales. So for example, we might expect characteristic time-scales for fx option models with mean-reverting stochastic volatility to be 3 months to 3 years while, for interest-rate derivatives models, we might expect characteristic time-scales to be 3 years to 40 years. This might translate into mean-reversion rates (at least for single-factor models with a single mean-reversion rate - we might have to modify our reasoning if there are two mean-reversion rates) of the order of magnitude of 0.33 to 4.0 for fx options models with mean-reverting stochastic volatility and of the order of magnitude of 0.025 to 0.33 for interest-rate derivatives.

- Bear in mind that strictly speaking (depending on the specifications of the market price(s) of risk), mean-reversion rates can be different in the real-world measure compared to the risk-neutral measure but given the approximations inherent in our heuristics, this seems to be the least of our worries.

- With some models, it is not uncommon, in practice, in the markets for quants or traders to simply fix mean-reversion rates to plausible levels (which is just a special case of a penalty function with an infinite weight).

- Historical time-series data may be able to give rough estimates of "vol-of-vol" parameters.

- Such parameter estimates are very rough and approximate but, nonetheless, it should be possible to "guesstimate" many (if not all) parameter values (at least to the right order of magnitude) before the calibration takes place. We can then incorporate our "guesstimates" into penalty functions.

- The penalty function need not necessarily be of the form above. For example, if we believed that a parameter θ was very unlikely to be above some value θ_{high}, we could try a penalty function of the form: $(wI_{\theta>\theta_{\text{high}}}(\theta - \theta_{\text{high}})^2)$, where I denotes the indicator function, for some choice of the positive weight w.

- Similarly, we can add penalty functions if we believe that a parameter θ is very unlikely to be below some value θ_{low}.

- All penalty functions should be smooth.

- Consider the G2++ Gaussian HJM interest-rate model in which continuously-compounded instantaneous forward-rates $f(t,T)$, at time t, to tenor T follow the SDE:

$$df(t,T) = \ldots dt + \sigma_1 \exp(-\alpha_1(T-t))dz_1(t) + \sigma_2 \exp(-\alpha_2(T-t))dz_2(t),$$

  where the Brownian increments $dz_1(t)$ and $dz_2(t)$ have constant correlation ρ and where σ_1, α_1, σ_2, α_2 are positive constants (and for brevity we don't write down the drift term - it is easily derived from the HJM no-arbitrage condition).

- In the special case that $\alpha_1 = \alpha_2$, the model is degenerate and it would make financial sense to include a penalty function which penalises if α_1 and α_2 are too close together. In a similar vein, it may be helpful to include a penalty function which forces the choice of which of α_1 and α_2 is the larger.

- Humped volatility curves (which, in practice, are very common in the caps and swaptions market) are only possible if the parameter ρ is negative. So, if this is the case, we can penalise positive values of ρ. (Clearly, we can always penalise values of ρ outside of, say, $[-0.995, 0.995]$).

- Historical data tells us that instantaneous forward interest-rates $f(t, T_1)$ and $f(t, T_2)$ of different tenors T_1 and T_2 should be positively, but not perfectly, correlated. We can, from historical data (probably using forward LIBOR rates as proxies for continuously-compounded instantaneous forward interest-rates), estimate $\text{correl}(df(t, T_1), df(t, T_2))$ for, say, $T_1 = 1$ and $T_2 = 10$ (typical values will be around 0.4).

- We can compute $\text{correl}(df(t, T_1), df(t, T_2))$ analytically within this model and use this as a penalty function within the calibration.

- We can then use all this information when we calibrate our model parameters to the market prices of caps or European swaptions.

- The "information" contained within knowledge of $\text{correl}(df(t, T_1), df(t, T_2))$ is different from that contained within knowledge of the implied Black (1976) volatilities (which are essentially equivalent to variances). This, intuitively, should help the calibration.

- If we were very sure about the accuracy of our estimate of correl$(df(t,T_1),df(t,T_2))$, we could even use the available analytical results to eliminate one of the parameters from the calibration and in doing so fix correl$(df(t,T_1),df(t,T_2))$.
- Together with intuition on the mean-reversion rates, it is clear that we can and should give the calibration a significant amount of information in the hope of making it more stable.

- One often extends the G2++ Gaussian HJM model to allow for time-dependent (usually piecewise constant) volatilities so as to be able to, for example, exactly match the market prices of at-the-money-forward swaptions of different expiries.

- The best way to do this is to calibrate the model with constant parameters ρ, σ_1, α_1, σ_2, α_2. Then set $\beta \equiv \sigma_2/\sigma_1$. Then hold ρ, β, α_1, α_2 fixed. Then define the model:

$$df(t,T) = \ldots dt + \eta(t)(\exp(-\alpha_1(T-t))dz_1(t) + \beta \exp(-\alpha_2(T-t))dz_2(t)),$$

  where $\eta(t)$ is a time-dependent function of t.

- This parameterization preserves the stationary correlation structure between forward rates of different tenors, whereas a parameterization, for example, of the form:

$$df(t,T) = \ldots dt + \sigma_1(t)\exp(-\alpha_1(T-t))dz_1(t) + \sigma_2 \exp(-\alpha_2(T-t))dz_2(t),$$

  with $\sigma_1(t)$ time-dependent and σ_2 constant, does not.

- Then (keeping ρ, β, α_1, α_2 fixed), we solve for $\eta(t)$ by a series of one-dimensional searches matching market prices of increasing expiries sequentially.

- Consider the Heston (1993) model:

$$\begin{aligned}\frac{dS}{S} &= (r_d - r_f)dt + \sqrt{V(t)}dz_S(t),\\ dV(t) &= \kappa(\theta - V(t))dt + \epsilon\sqrt{V(t)}dz_V(t),\ \ V(0) \equiv V_0,\\ \text{correl}(dz_S(t), dz_V(t)) &= \rho.\end{aligned}$$

  It is not difficult to see that the value of the initial variance V_0 should be approximately equal to the square of the implied volatility of a very short-dated at-the-money-forward option. Hence, we can use a penalty function which penalises a value of V_0 a long way from this.

- Now $V(t)$ follows a mean-reverting diffusion process with long-run reversion level θ. If we calibrate the model and find that the value of θ is very significantly different from V_0, it would beg the question why? After all, $V(t)$ is a mean-reverting process which reverts to θ and so it should not move too far away from θ at any time, including at time 0.

- Hence, it is tempting to also use a penalty function which penalises if θ is too far from V_0.
- However, it is a little harder to define what "too far" might mean.
- It can be shown that if $2\kappa\theta < \epsilon^2$ (usually called the Feller condition) then the process for $V(t)$ can reach zero (and given that the long run mean level is θ, it must presumably be able to move significantly above θ also). Of course, we will not know what κ and ϵ are until after we have completed the calibration.

- The last slide demands a more considered approach. Suppose we have the last D consecutive working days data for the market prices of N options. We denote the market prices by $C_i(d)$, for each $d = 1, \ldots, D$ and each $i = 1, \ldots, N$.

- Now the parameter V_0 is different from the other parameters κ, θ, ϵ and ρ. Whereas the latter are supposed to stay constant, the parameter V_0 represents the current value of the instantaneous variance and it is supposed to vary through time. Suppose it has the values $V_0^{(1)}$, $V_0^{(2)}$, ..., $V_0^{(D)}$ on days $d = 1, \ldots, D$ respectively.

- Let us denote by $P(V(u), u, V(s), s)$ the transition probability density function for the probability of $V(s)$, at time s, transitioning to $V(u)$, at time u, for $u \geq s$. Denote by $G_i(\kappa, \theta, \epsilon, \rho, V_0^{(d)})$ the model price of the option on day d when the variance is $V_0^{(d)}$, for each $i = 1, \ldots, N$. Our calibration now minimises, by choice of the $D + 4$ parameters $\kappa, \theta, \epsilon, \rho, V_0^{(1)}, V_0^{(2)}, \ldots, V_0^{(D)}$:

$$\sum_{d=1}^{D} \sum_{i=1}^{N} \omega_i (G_i(\kappa, \theta, \epsilon, \rho, V_0^{(d)}) - C_i(d))^2 - \sum_{d=2}^{D} \log P(V_0^{(d)}, d, V_0^{(d-1)}, d-1).$$

- The rationale for the minus sign in front of the second summation is that we want to maximise the $P(V_0^{(d)}, d, V_0^{(d-1)}, d-1)$ terms (because that will mean that the successive values of the variance $V_0^{(d)}$ are more likely to be consistent with the underlying stochastic process for $V(t)$) which means minimising the negative of $P(V_0^{(d)}, d, V_0^{(d-1)}, d-1)$.

- A typical value of D might be between 2 and 10.

- What has this achieved?

- On the downside, we now have to value ND options rather than N and we have to minimize by choice of $D+4$ parameters rather than by choice of 5 parameters so we have made the calibration larger.

- On the upside, we can use data for D days and try to make sure that successive days calibrations are internally consistent with the stochastic process that we are trying to fit.

- Note that the final calibrated value of V_0 that we are interested in is simply $V_0^{(D)}$.

- The transition probability density function $P(V(u), u, V(s), s)$ is known in closed-form (in terms of Bessel functions) for the Heston (1993) process (see p607, chapter 13 of Lipton (2001)).

- We have illustrated this approach for the Heston (1993) model but the same approach could, for example, be used for the double Heston model or indeed any stochastic volatility model.

- It is better for this approach if the transition probability density function is known, in closed-form, for the model under consideration (as it is for Heston (1993) model) but, even if it is not known in closed-form, if the first few moments are known, then the transition probability density function could probably be satisfactorily approximated by a Edgeworth series expansion.

- Most of the possible methods described above for potentially improving the quality of calibrations could be described as ad-hoc or as fudges. Nevertheless, it is worth reflecting that the whole concept of calibrating a model by solving an ill-posed and unstable inverse problem is also unsatisfactory.

- Taken in this context, and given the importance of calibration, it seems reasonable to use every possible piece of information at our disposal.

Risk-neutral probabilities of large downward jumps are higher than real-world probabilities

- We claimed earlier that, at least for major equity indices, risk-neutral probabilities of large downward jumps are higher than real-world (historical) probabilities of the same large downward jumps.

- Why?

- The first point to emphasize is that option pricing theory is (nearly always) concerned with probabilities under a risk-neutral measure $\mathbb{Q}$. For example, we can extract risk-neutral probabilities from option prices (of a continuum of strikes). However, option pricing theory is not (with very few exceptions) concerned with probabilities under the real-world (historical) measure $\mathbb{P}$. Therefore, we need to go beyond the maths into economics in order to deduce relationships between real-world ($\mathbb{P}$) probabilities and risk-neutral ($\mathbb{Q}$) probabilities.

- For future reference, we define a utility function $u(c_t)$, where c_t is consumption, at time t.
- Note: We do not need (and will not specify) the parameters (eg. coefficient of risk-aversion) nor even the functional form (eg. log, quadratic, power, exponential) of the utility function (which is important as, in practice, it is difficult to come up with such parameters or functional form).
- A utility function is just a mapping function which defines preferences over future wealth or consumption. In a world with random outcomes (stochastic processes), they are preferences over **expected** future wealth or consumption.
- Investors want to maximise expected future utility of consumption.
- Sensible properties of a utility function: $u^{'}(c_t) > 0$ (investors prefer more consumption to less) and $u^{''}(c_t) < 0$ (but the rate at which they do so declines as consumption increases).

- Let $p_t \equiv p$ be the price of a security or asset at time t whose payoff at time $t+1$ is $x_{t+1} \equiv x$ in an economy in which the pricing kernel is m (m is also called the marginal rate of substitution or stochastic discount factor or the state-price density), where m is a random variable (or more generally a stochastic process) which will be defined shortly. Then:

$$p = E_t^{\mathbb{P}}[m\ x] \quad \text{or, with the time subscripts} \quad p_t = E_t^{\mathbb{P}}[m\ x_{t+1}]. \tag{3}$$

  $E_t^{\mathbb{P}}$ denotes real-world (**NOT** risk-neutral) expectations (i.e. under the real-world physical measure $\mathbb{P}$). This result does not assume complete markets or any particular form of the utility function $u(c_t)$.

- Proof (John Cochrane's book "Asset pricing"): An investor maximises by choice of ξ: $u(c_t) + E_t^{\mathbb{P}}[\overline{\beta}u(c_{t+1})]$ subject to $c_t = e_t - p_t\xi$ and $c_{t+1} = e_{t+1} + x_{t+1}\xi$. Here c_t denotes consumption, at time t, ξ is the amount of the security that the investor buys and e_t denotes the original consumption, at time t, (i.e. if the investor didn't buy the security at all).

- Hence, investor maximises: $u(e_t - p_t\xi) + E_t^{\mathbb{P}}[\overline{\beta}u(e_{t+1} + x_{t+1}\xi)]$ over choice of ξ:

- Note $\overline{\beta} > 0$ is some constant usually called the subjective discount factor (it measures the investor's "impatience" and is typically some number equal to or just slightly below one).

- To solve the maximization problem, differentiate with respect to ξ. The first-order condition is:

$$\begin{aligned} 0 &= -p_t u^{'}(c_t) + E_t^{\mathbb{P}}[\overline{\beta} u^{'}(c_{t+1})\, x_{t+1}] \Rightarrow \\ p_t &= E_t^{\mathbb{P}}[\frac{\overline{\beta} u^{'}(c_{t+1})}{u^{'}(c_t)}\, x_{t+1}] \equiv E_t^{\mathbb{P}}[m\ x_{t+1}],\ \text{where } m \equiv \frac{\overline{\beta} u^{'}(c_{t+1})}{u^{'}(c_t)}. \end{aligned} \tag{4}$$

- The result is (more or less) equivalent to:

$$p_t = E_t^{\mathbb{Q}}[\frac{1}{R_f}\ x_{t+1}], \tag{5}$$

  where R_f is one plus the risk-free rate (measured per period $t+1$ minus t - NOT per annum) so that $\frac{1}{R_f}$ is the risk-free discounting term / factor and $E_t^{\mathbb{Q}}$ denotes risk-neutral expectations (under some (non-unique in general, unique only in complete markets) measure $\mathbb{Q}$). Note that $E_t^{\mathbb{P}}[m] \equiv E_t^{\mathbb{P}}[m\,1]$ is the price of a payoff of one i.e. $E_t^{\mathbb{P}}[m]$ is the price, at time t, of a risk-free zero-coupon bond maturing at time $t+1$. Hence, $E_t^{\mathbb{P}}[m] = \frac{1}{R_f}$.

- Note that:

$$p_t = E_t^{\mathbb{Q}}[\frac{1}{R_f}\ x_{t+1}], \tag{6}$$

is a completely standard result based on no arbitrage (in words, prices are expected, under $\mathbb{Q}$, discounted payoffs).

- Note Girsanov's theorem hidden in here. m is also $\frac{1}{R_f}\frac{d\mathbb{Q}}{d\mathbb{P}}$.

- We said that a sensible property of a utility function is $u^{'}(c_t) > 0$. If this is so, then $m \equiv \frac{\overline{\beta}u^{'}(c_{t+1})}{u^{'}(c_t)} > 0$.

- Theorem: $p_t = E_t^{\mathbb{P}}[m\ x_{t+1}]$ and $m > 0$ imply no arbitrage.

- Proof: No arbitrage implies that if every possible payoff x_{t+1} is always non-negative and at least one possible (i.e. with strictly positive probability) payoff x_{t+1} is strictly positive, then the payoff has positive price i.e. $p_t = E_t^{\mathbb{P}}[m\ x_{t+1}] > 0$. But $m > 0$ and there is at least one possible payoff x_{t+1} where $x_{t+1} > 0$. Hence, $m\ x_{t+1} > 0$ for at least one state and $m\ x_{t+1} \geq 0$ in all possible states. Hence, $p_t = E_t^{\mathbb{P}}[m\ x_{t+1}] > 0$.

- We will henceforth assume $m > 0$.

- Now suppose there are N states i.e. x_{t+1} can take one of N possible values. Then:

$$p_t = \sum_{j=1}^{N} \pi(j)^{\mathbb{P}} x_{t+1}(j) m(j) \quad \text{and}$$

$$p_t = \sum_{j=1}^{N} \pi(j)^{\mathbb{Q}} \frac{1}{R_f} x_{t+1}(j), \tag{7}$$

where we have inserted the argument j for each of the $j = 1, \ldots, N$ states. Hence,

$$\Pi(j)^{\mathbb{P}} m(j) = \Pi(j)^{\mathbb{Q}} \frac{1}{R_f}$$

The term $\frac{1}{R_f}$ is the risk-free discounting term which can be viewed as a constant in a one period model. Hence, the relationship between $\Pi(j)^{\mathbb{P}}$ and $\Pi(j)^{\mathbb{Q}}$ is determined by $m \equiv m(j)$ which is determined by $u^{'}(c_{t+1})$.

- Now $m \equiv \frac{\overline{\beta} u'(c_{t+1})}{u'(c_t)}$ rises if marginal utility of consumption $u'(c_{t+1})$ rises but $u'(c_{t+1})$ rises if c_{t+1} and hence also wealth and also the market (proxied by a broad-based equity index) falls (since the market is held in aggregate).

- Hence, $\Pi(j)^{\mathbb{P}} m(j) = \Pi(j)^{\mathbb{Q}} \frac{1}{R_f}$ implies that risk-neutral ($\mathbb{Q}$) probabilities of large downward jumps in a broad-based equity index are greater than the real-world ($\mathbb{P}$) probabilities of the same large downward jumps.

- This provides support to the claim we made earlier.

- We can rewrite $p_t = E_t^{\mathbb{P}}[m\ x_{t+1}]$ in two equivalent forms: Using the definition of covariance, the first form is:

$$\begin{aligned} p_t &= E_t^{\mathbb{P}}[m]E_t^{\mathbb{P}}[x_{t+1}] + \text{Cov}_t[m, x_{t+1}] \text{ or since } E_t^{\mathbb{P}}[m] = \tfrac{1}{R_f} \\ &= \frac{1}{R_f}E_t^{\mathbb{P}}[x_{t+1}] + \text{Cov}_t[m, x_{t+1}] = \frac{1}{R_f}E_t^{\mathbb{P}}[x_{t+1}] + \text{Cov}_t[\frac{\overline{\beta} u^{'}(c_{t+1})}{u^{'}(c_t)}, x_{t+1}]. \end{aligned} \tag{8}$$

Hence, p_t consists of a term which somewhat resembles risk-free discounting and a risk-adjustment term $\text{Cov}_t[\frac{\overline{\beta} u^{'}(c_{t+1})}{u^{'}(c_t)}, x_{t+1}]$.

Now $\frac{\overline{\beta} u^{'}(c_{t+1})}{u^{'}(c_t)}$ rises if marginal utility of consumption $u^{'}(c_{t+1})$ rises but $u^{'}(c_{t+1})$ rises if c_{t+1} and hence also wealth and also the market (proxied by a broad-based equity index) falls. Hence, an asset's price is raised for assets whose payoffs are negatively correlated with consumption and hence with wealth and hence with the market. Put option payoffs are negatively correlated with the market - especially deep out-of-the-money puts - since they pay more when the market falls. Hence, we expect put option prices (and hence implied volatilities) are raised. This can explain the volatility smile.

- From the viewpoint of economics, investors pay a premium on the prices of deep out-of-the-money puts and they view this as a kind of insurance premium for protecting them from the unpleasant consequences of a fall in the market equity index.

- The second (completely equivalent) form is obtained by dividing through by p_t. We define $R \equiv x_{t+1}/p_t$ so that R is the return on the asset over the period t to $t+1$ (R is a number like eg. 1.03 or 1.07 - NOT like eg. 0.03 or 0.07). Then

$$\begin{aligned} p_t &= \frac{1}{R_f} E_t^{\mathbb{P}}[x_{t+1}] + \mathrm{Cov}_t[m, x_{t+1}] = \frac{1}{R_f} E_t^{\mathbb{P}}[x_{t+1}] + \mathrm{Cov}_t[\frac{\overline{\beta} u'(c_{t+1})}{u'(c_t)}, x_{t+1}] \Rightarrow \\ 1 &= \frac{1}{R_f} E_t^{\mathbb{P}}[R] + \mathrm{Cov}_t[m, R] = \frac{1}{R_f} E_t^{\mathbb{P}}[R] + \mathrm{Cov}_t[\frac{\overline{\beta} u'(c_{t+1})}{u'(c_t)}, R] \Rightarrow \\ E_t^{\mathbb{P}}[R] - R_f &= -R_f \mathrm{Cov}_t[m, R] = -R_f \mathrm{Cov}_t[\frac{\overline{\beta} u'(c_{t+1})}{u'(c_t)}, R]. \end{aligned} \tag{9}$$

- Written again:

$$E_t^{\mathbb{P}}[R] - R_f \;=\; -R_f \mathrm{Cov}_t[m, R] = -R_f \mathrm{Cov}_t[\frac{\overline{\beta} u'(c_{t+1})}{u'(c_t)}, R]. \tag{10}$$

Hence, assets whose returns are negatively correlated with the market have expected excess (i.e. over the risk-free rate) returns which are negative. This is in agreement with the argument two slides ago since high prices (respectively, low prices) imply low returns (respectively, high returns).

- Or, assets whose returns are positively correlated with the market (eg. (typically) individual company shares) have expected excess returns which are positive. Investors would like to smooth their consumption but positively correlated assets perform badly when the market does badly (i.e. in economic bad times). Hence, in the good times, they earn a positive excess (i.e. over and above the risk-free rate) return i.e. a risk premium to compensate them.

- To put it another way, investors demand a positive excess return, on average, to induce them to hold securities whose returns are positively correlated with consumption i.e. whose returns are good when they are already feeling wealthy and are bad when they are already feeling poor.

- Investors will accept a negative excess return, on average, to holding securities whose returns are negatively correlated with consumption. This is like insurance. Insurance involves paying premiums and most of the time getting nothing back. But when investors are feeling very poor because eg. their house has just burnt down, they get a large payoff eg. a big cheque from the insurance company.

- Finally, we can also write:

$$E_t^{\mathbb{P}}[R] - R_f = \left(\frac{\mathrm{Cov}_t[m, R]}{\mathrm{Var}_t[m]}\right)\left(\frac{-\mathrm{Var}_t[m]}{E_t^{\mathbb{P}}[m]}\right) \equiv \beta\lambda. \tag{11}$$

  Here, $\lambda \equiv \frac{-\mathrm{Var}_t[m]}{E_t^{\mathbb{P}}[m]}$ depends only on the market (i.e. only on m (and not on R which is security specific)) and is called the market price of risk while $\beta \equiv \frac{\mathrm{Cov}_t[m,R]}{\mathrm{Var}_t[m]}$ is the "beta" defined to be the regression coefficient of the return R on m. Note that β is different for each security or asset (it depends on R).

- What else can we do with the notion of a pricing kernel?
- Cochrane and Saa-Requejo (2000) introduce the idea of "no good deals" as a way of pricing and hedging derivatives or complex securities.
- No good deals cleverly use a little economics to price derivatives or complex securities or more accurately to compute lower and upper good deal bounds on their values.
- It is a very natural way of placing reservation prices (mark-to-market or mark-to-model values) on complex derivatives i.e. long positions are marked at the lower good deal bound and short positions are marked at the upper good deal bound.

- They say suppose we want to find a lower good deal bound $\underline{C}$ and an upper good deal bound $\overline{C}$ on the price, at time t, of a security C (eg a complex derivative) with payoff x_c at time $t+1$. There are N_b "basis assets" which are liquid and actively traded which pay x_i at time $t+1$ whose (market) prices are known to be p_i at time t, for $i = 1, 2, \ldots, N_b$. They solve:

$$\begin{aligned} \underline{C} = \min E_t^{\mathbb{P}}[m x_c] \quad \text{and} \quad \overline{C} &= \max E_t^{\mathbb{P}}[m x_c] \quad \text{by choice of } m\ , \\ \text{such that } p_i &= E_t^{\mathbb{P}}[m\ x_i], \text{ for all } i, \quad \text{and} \quad m > 0, \\ \text{AND such that } E_t^{\mathbb{P}}[m^2] &\leq \frac{(1+h^2)}{R_f^2}, \text{ for a constant } h,\ h > 0\ . \end{aligned} \tag{12}$$

The innovation of Cochrane and Saa-Requejo (2000) is all on the third (final) line. This is a restriction on the second moment and therefore on the volatility of the pricing kernel m.

- Note the change in notation. We are trying to place lower and upper bounds on the security C.

- The N_b "basis assets", with payoffs x_i, are liquid (we can observe their market prices) and actively traded (we can use them as hedges).

- The conditions $p_i = E_t^{\mathbb{P}}[m\ x_i]$, for each $i = 1, \ldots, N_b$, say we must reprice the basis assets. The condition $m > 0$ says we must be consistent with no arbitrage.

- **Without** the restriction on the second moment, the equations would read:

$$\underline{C} = \min E_t^{\mathbb{P}}[mx_c] \quad \text{and} \quad \overline{C} = \max E_t^{\mathbb{P}}[mx_c] \quad \text{by choice of } m \ ,$$
$$\text{such that } p_i \quad = \quad E_t^{\mathbb{P}}[m\ x_i], \qquad m > 0.$$

- These are the arbitrage bounds (Merton (1973) - solvable by a Linear Program). The problem with the arbitrage bounds is that (in incomplete markets) they are typically too wide to be useful even for vanilla options (because there are many (typically, an infinite number of) risk-neutral measures and hence many (infinite number of) pricing kernels). The innovation of Cochrane and Saa-Requejo (2000) is all in the extra condition $E_t^{\mathbb{P}}[m^2] \leq \frac{(1+h^2)}{R_f^2}$.

- Why the restriction $E_t^{\mathbb{P}}[m^2] \leq \frac{(1+h^2)}{R_f^2}$?

- By definition, a return R (NOTE: **any** return R) satisfies $E_t^{\mathbb{P}}[m\,R] = 1$ (since $R \equiv x/p$) which implies $1 = \text{Correl}_t[m\,, R]\,\text{st.dev}[R]\,\text{st.dev}[m] + E_t^{\mathbb{P}}[m]E_t^{\mathbb{P}}[R]$. But by definition of any correlation, $|\text{Correl}_t[m\,, R]| \leq 1$. Therefore, rearranging and using $\frac{1}{R_f} \equiv E_t^{\mathbb{P}}[m]$:

$$\frac{|E_t^{\mathbb{P}}[R] - R_f|}{\text{st.dev}[R]} \leq \text{st.dev}[m]R_f. \tag{13}$$

- But what is:

$$\frac{|E_t^{\mathbb{P}}[R] - R_f|}{\text{st.dev}[R]}\ ? \tag{14}$$

- $$\frac{|E_t^{\mathbb{P}}[R] - R_f|}{\text{st.dev}[R]} \quad \text{is the Sharpe ratio - lets call it SR.} \tag{15}$$

- Therefore, using the definition of standard deviation,

$$\begin{aligned} E_t^{\mathbb{P}}[m^2] \leq \frac{(1+h^2)}{R_f^2} &\iff R_f^2(E_t^{\mathbb{P}}[m^2] - \frac{1}{R_f^2}) \leq h^2 \\ &\iff \text{st.dev}[m]R_f \leq h \\ &\iff \text{SR} \leq h. \end{aligned} \tag{16}$$

- So the restriction is just equivalent to saying that the Sharpe ratio SR must be less than a fixed quantity h.

- (Footnote: Note how general this result is: In any economy, the maximum Sharpe ratio (of any security or asset or portfolio of these - no matter what its distribution) is bounded by st.dev$[m]R_f$ (this is called the Hansen and Jagannathan (1991) bound)).

- Note we are **not** pricing derivatives by appealing to CAPM type arguments in using Sharpe ratios.

- We are simply saying that no portfolio priced by m can have a Sharpe ratio SR greater than h (and also there can be no arbitrage opportunity (because $m > 0$)). Even more specifically, no portfolio of the N basis assets and/or the security C can have a Sharpe ratio greater than h. This places a restriction on m, beyond simply requiring $m > 0$. The extra restriction narrows the arbitrage bounds - and makes the good deal bounds useful.

- Note we are modelling prices under the real-world physical measure $\mathbb{P}$ (not under a risk-neutral measure $\mathbb{Q}$) so we need to model the real-world dynamics. The equation:

$$\begin{aligned} \underline{C} = \min E_t^{\mathbb{P}}[mx_c] \quad \text{and} \quad \overline{C} = \max E_t^{\mathbb{P}}[mx_c] \quad & \text{by choice of } m \ , \\ \text{such that } p_i \quad = \quad & E_t^{\mathbb{P}}[m\ x_i], \text{ for all } i, \quad \text{and} \quad m > 0, \\ \text{AND such that } E_t^{\mathbb{P}}[m^2] \quad \leq \quad & \frac{(1+h^2)}{R_f^2}, \text{ for a constant } h,\ h > 0\ . \end{aligned}$$

is a linear-quadratic program and can easily be solved numerically (and very occasionally analytically) - using the method of Lagrange multipliers.

- No good deals is NOT utility maximization. It is arbitrage pricing. But in incomplete markets, there are multiple (typically, an infinite number of) pricing kernels (or equivalent martingale measures). No good deals rules out pricing kernels which correspond to trading opportunities which are "too good to be true". Since if they are "too good to be true", then, like an arbitrage, they shouldn't be there in a market equilibrium.

- Alternative intuition:

- No good deals computes lower and upper good deal bounds which either
  (1) achieve the target level of the Sharpe ratio h, or
  (2) are the arbitrage bounds when the target level cannot be achieved.

- The no good deals methodology also computes optimal hedges (see exercise).

- Note that we are not appealing to strict assumptions about representative agents' utility functions here - although it turns out that the restriction $E_t^{\mathbb{P}}[m^2] \leq \frac{(1+h^2)}{R_f^2}$ can be (but does not have to be) linked to a quadratic utility function.

- This suggests that one can also solve for a whole class of no good deal bound problems replacing $E_t^{\mathbb{P}}[m^2]$ by $E_t^{\mathbb{P}}[f(m)]$ for a class of convex functions $f(m)$. For example, it turns out one can solve for "log good deal bounds" satisfying:

$$\begin{aligned} \underline{C}^{\log} = \min E_t^{\mathbb{P}}[mx_c] \quad \text{and} \quad \overline{C}^{\log} = \max E_t^{\mathbb{P}}[mx_c] \quad \text{by choice of } m \ , \\ \text{such that } p_i \quad &= \quad E_t^{\mathbb{P}}[m \ x_i], \ \text{for all } i, \quad \text{and} \quad m > 0, \\ \text{AND such that } E_t^{\mathbb{P}}[-\log m] \quad &\leq \quad \mathbb{M}, \ \text{for some constant } \mathbb{M}. \end{aligned}$$

- This problem can be simplified by the use of the method of Lagrange multipliers and then computed numerically.

- Pricing a book of exotic options is not easy. The perfect model does not exist - even for fx which I would say is the simplest-to-model asset class.

- A lot of lessons have been learnt by quants the hard way and a lot of models have been arrived at by trial and error.

- Carr P. and L. Wu (2007) "Stochastic skew in currency options" Journal of Financial Economics Vol. 86 (October 2007) p213-247.

- Cochrane J. and J. Saa-Requejo (2000) "Beyond arbitrage: Good deal asset price bounds in incomplete markets" Journal of Political Economy Vol. 108 p79-119

- Cochrane J. (2005) "Asset pricing" Princeton University Press

- Cont R. and P. Tankov (2003) "Financial modelling with jump processes" Chapman and Hall, CRC Press, 2003.

- Cont R. and P. Tankov (2004) "Nonparametric calibration of jump-diffusion option pricing models" Journal of Computational Finance, Vol. 7 p1-49.

- Hansen L.P. and R. Jagannathan (1991) "Implications of security market data for models of dynamic economies" Journal of Political Economy Vol. 99 p225-262.

- Kainth D. (2007) "Pricing FX options under multi-factor Heston" Presentation given at the ICBI Global Derivatives conference, Paris, 23rd May 2007.
- Lipton A. (2001) "Mathematical methods for foreign exchange" World Scientific
- Lipton A. and W. McGhee (2002) "Universal barriers" Risk Vol. 15 (May 2002) p81-85.
- Merton R. (1973) "The theory of rational option pricing" Bell Journal of Economics and Management Science" Vol. 4 p141-183
- Schoutens W., E. Simons and J. Tistaert (2004) "A Perfect Calibration! Now What?", Wilmott magazine, 2004 or see Wim Schouten's website
- Schoutens W., E. Simons and J. Tistaert (2005) "Model risk for exotic and moment derivatives" in "Exotic options and advanced Lvy models" A. Kyprianou, W. Schoutens and P. Wilmott (eds) Wiley