

ONE FOR ALL

The potential approach to pricing and hedging

L. C. G. Rogers

University of Cambridge

Based on a talk given at the European Congress on Mathematics in Industry, TU Eindhoven, June 2004, and to appear in the proceedings of that meeting.

L. C. G. Rogers
Statistical Laboratory
Wilberforce Road
Cambridge CB3 0WB
Great Britain
phone = +44 1223 766806, fax = +44 1223 337956
e-mail = lcgr1@cam.ac.uk

ONE FOR ALL

The potential approach to pricing and hedging

L. C. G. Rogers

University of Cambridge

1 Introduction

Any academic working in mathematics, physics or engineering will know examples of bright PhD students who have turned their backs on the profession that nurtured them, and sought out instead the more uncertain but more lucrative world of financial services. What do they find themselves doing when they get there? This largely depends on the job they get (for a large investment bank is a diverse employer, with many different roles and required skills), but most often banks are employing people with outstanding quantitative skills because those are the skills they need.

In particular, a big part of the business of an investment bank is in making and selling various *derivatives*¹. A very simple example is a (zero-coupon) *bond*, where the seller undertakes to pay 1 at a fixed time T in the future, in return for a payment (typically less than 1!) made by the buyer at time 0. In more complicated examples, the amounts to be paid can be random, as in a European put option, which gives the buyer the right to sell one unit of a named stock at a specified time (the *expiry* of the contract) for a specified price (the *strike* price). However, the holder is not compelled to exercise this right, so he will clearly do so at expiry if and only if the stock is trading for less than the strike, because he is then able to buy the stock, and pocket the difference between the current price and the strike price he gets by selling it to the option writer. The timings of payments can also be random, as in an American put, where the holder of the option is allowed to sell the stock for the strike price at *any* time before the expiry of the option.

What should be the price the bank charges for a derivative? At one level, the answer is ‘As much as the market will bear!’ but competition precludes arbitrary

¹A derivative is a well-defined financial deal between two parties where the timings and amounts of payments to be made are specified in the contract.

profit, and the bank needs to have some idea of how cheaply they can sell the derivative and still make a profit. Quite viable models and methods exist to help banks answer such questions, and the job of a *quant* (as your former PhD student is now known) is to use such models to come up with prices, and to extend them to deal with the novel derivatives that are often requested by clients. Such models will be used to find lower bounds for the prices the banks will charge. In very liquid (active) markets, the actual price charged may be very close to that computed in the model, but in less liquid markets it will typically be a lot higher. This is because the bank does not simply sell a derivative and then wait for events to determine what their part of the deal is going to cost them; they engage in *hedging*, which is to say they take some offsetting position in other financial assets so as to cancel out (as best they can) any gains and losses to be made on the derivative. If the market is highly liquid, they are better able to do this than in an illiquid market, and the price charged reflects this. Indeed, it is largely true that the price a bank will charge for a derivative is the cost they face in hedging it, rather than anything that a model tells them.

So pricing and hedging forms a large part of the work your former PhD student will be doing. In this introductory paper, we will see in Section 2 some of the basic notions of pricing developed from a perfectly plausible (though slightly unconventional) axiomatic standpoint; and we will see in Section 3 how the resulting expressions suggest an approach to modelling asset prices. This approach (known as the *potential approach*) is again rather unconventional, but has overwhelming advantages in the modelling of complex cross-currency derivatives (for example), that more conventional approaches struggle with; Section 5 explains why. The potential modelling approach is actually extremely general, and specific choices have to be made to apply it in practice - these are discussed in Sections 4 and 6. We discuss how such models can be calibrated in Section 7, and present the results of such a calibration in Section 8. As befits an unconventional approach, the form that hedging takes (explained in Section 9) is also unconventional, but perfectly tractable. Section 10 concludes and presents further directions for research.

Very little that is in this paper is new; indeed, most of it is in one or more of [5], [6], [7], [8], [9]. I have said that the potential approach is unconventional, and it is; for fine expositions of the conventional approach to derivative pricing and hedging you can consult Baxter & Rennie [1], Karatzas & Shreve [4], Dana & Jeanblanc [2] (for an account with more emphasis on the economic origins), or many others - there are plenty. These accounts will tell you a lot about how pricing and hedging is done today; this account will tell you a little about how it may well be done tomorrow.

2 Generalities about pricing

We put ourselves in a filtered probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and consider *market pricing* operators $(\pi_{tT})_{0 \leq t \leq T}$ for contingent claims²:

$$\pi_{st} : L^\infty(\mathcal{F}_t) \rightarrow L^\infty(\mathcal{F}_s) \quad (0 \leq s \leq t).$$

The idea is that if Y is some bounded random variable which is \mathcal{F}_t -measurable, then the time- s market price of Y is $\pi_{st}(Y)$, again a random variable (because what we had observed up to time s would affect what we thought this contingent claim was worth), and again bounded - obviously.

We shall assume that the pricing operators $(\pi_{st})_{0 \leq s \leq t}$ satisfy certain axioms:

(A1) Each π_{st} is a bounded positive linear operator from $L^\infty(\mathcal{F}_t)$ to $L^\infty(\mathcal{F}_s)$;

(A2) If $Y \in L^\infty(\mathcal{F}_t)$, $Y \geq 0$, then

$$\pi_{0t}(Y) = 0 \iff P(Y > 0) = 0.$$

(*no arbitrage*)

(A3) For $0 \leq s \leq t \leq u$, $Y \in L^\infty(\mathcal{F}_u)$, $X \in L^\infty(\mathcal{F}_t)$,

$$\pi_{su}(XY) = \pi_{st}(X\pi_{tu}(Y))$$

(*intertemporal consistency*)

(A4) If $(Y_n) \in L^\infty(\mathcal{F}_t)$, $|Y_n| \leq 1$, $Y_n \uparrow Y$ then $\pi_{st}(Y_n) \uparrow \pi_{st}(Y)$ (*continuity*)

REMARKS. Axiom (A1) says that the price of a non-negative contingent claim will be non-negative, and the price of a linear combination of contingent claims will be the linear combination of their prices - which are reasonable properties for a market price. Axiom (A2) says that a contingent claim that is almost surely worthless when paid, will be almost surely worthless at all earlier times (and conversely) - again reasonable. The third axiom, (A3), is a ‘consistency’ statement; the market prices at time s for XY at time u , or for X times the time- t market price for Y at time t , should be the same, for any X which is known at time t . The final axiom is a natural ‘continuity’ condition which is needed for technical reasons.

Theorem 1. *Assuming Axioms (A1)–(A4), there exists a strictly positive process $(\zeta_t)_{t \geq 0}$ such that the pricing operators π_{st} can be expressed as*

$$\pi_{st}(Y) = \frac{E_s[\zeta_t Y]}{\zeta_s} \quad (0 \leq s \leq t). \quad (2.1)$$

²A contingent claim is local jargon for a random variable.

If we also assume

(A5) For all $0 \leq s \leq t$, $\pi_{st}(1) \leq 1$

(where 1 denotes the constant function identically equal to 1) then ζ is a positive supermartingale:

$$\zeta_s \geq E_s \zeta_t \quad (0 \leq s \leq t)$$

REMARKS. This result is the famous ‘Fundamental Theorem of Asset Pricing’ (FTAP) - or at least its *conclusion* is the same as that of the FTAP, though its hypotheses are quite different. The FTAP is proved from the hypothesis that the market does not admit any arbitrage³. This is a perfectly sensible axiomatic starting point, but not the only one possible, and the approach taken here shows that if we adopt the equally-sensible axioms (A1)–(A4) (which need no subtle modification for a continuous-time setting), then we harvest the conclusion of the FTAP using little more than basic measure theory. Notice how short the proof is!

Proof. Firstly, for any $T > 0$, the map

$$A \mapsto \pi_{0T}(I_A)$$

defines a non-negative measure on the σ -field \mathcal{F}_T , from the linearity and positivity (A1) and the continuity property (A4). Moreover, this measure is absolutely continuous with respect to \mathbb{P} , in view of (A2). Hence there is a non-negative \mathcal{F}_T -measurable random variable ζ_T such that

$$\pi_{0T}(Y) = E[\zeta_T Y]$$

for all $Y \in L^\infty(\mathcal{F}_T)$. Moreover, $\mathbb{P}[\zeta_T > 0] > 0$, because of (A2) again. Now we exploit the consistency condition (A3); we have

$$\pi_{0t}(X \pi_{tT}(Y)) = E[X \zeta_t \pi_{tT}(Y)] = \pi_{0T}(XY) = E[XY \zeta_T].$$

Since $X \in L^\infty(\mathcal{F}_t)$ is arbitrary, we deduce that

$$\pi_{tT}(Y) = E_t[Y \zeta_T] / \zeta_t,$$

as claimed. The final statement that ζ is a positive supermartingale under (A5) is now immediate. ■

³The intuitive idea of an arbitrage as ‘something for nothing’ is easy to formalise mathematically in discrete time, but the exact definition for a continuous-time setting was elusive and subtle, and was indeed a key part of the difficulty experienced in proving this celebrated result. See Delbaen-Schachermayer [3] for the definitive form.

REMARKS. (i) The form (2.1) shows that if we write $Y_t \equiv \pi_{tT}(Y)$ for some fixed $Y \in L^\infty(\mathcal{F}_T)$ then

$$\zeta_t Y_t = E_t[\zeta_T Y] \text{ is a martingale.}$$

Conventionally, the process ζ (known as the *state-price density process*) is represented as

$$\zeta_t = \exp\left(-\int_0^t r_s ds\right) Z_t,$$

where r_t is the instantaneous rate of interest at time t , and Z_t is a positive martingale, which is interpreted as a change of measure, from the reference probability \mathbb{P} to some new ‘pricing’ probability, also referred to as an *equivalent martingale measure*, because it is equivalent to the original measure (both have the same null sets), and because in the new measure the discounted prices of all traded assets become martingales.

(ii) Though we have looked at pricing systems which are linear in the contingent claim, there is good reason not to restrict exclusively to this property, because individual agent’s prices for contingent claims are generally concave - you might be prepared to pay \$2 for 1l of icecream, but does this mean you would be prepared to pay \$200 for 100l of icecream?! Taking this into account leads us into ideas of economic equilibrium; often, the analysis of an equilibrium can be enormously complicated, and the equilibrium prices arrived at will depend on the nature of all the agents in the market. However, the equilibrium prices, being *marginal* prices, will be linear in the contingent claim.

3 The potential approach

Theorem 1 and the form (2.1) of the price of a contingent claim suggests a simple and natural approach to modelling (and pricing) in a financial market: *model ζ !*

The process ζ is a positive supermartingale (if we make the natural further assumption A5). The expression (2.1) allows us to write down the price $B(t, T)$ at time t of a zero-coupon bond maturing at later time T ; in this case, $Y \equiv 1$, so we have

$$B(t, T) = E_t[\zeta_T] / \zeta_t. \tag{3.2}$$

If we make the further assumption (financially very natural) that as the maturity T of the bond tends to infinity the current value of it tends to zero, we see that the positive supermartingales ζ that we are considering have to satisfy the further condition

$$\lim_{T \rightarrow \infty} E\zeta_T = 0; \tag{3.3}$$

a positive supermartingale satisfying this condition is called a *potential*, whence the name of this approach. Under a mild further condition⁴ a potential can be represented as

$$\zeta_t = E_t[A_\infty - A_t] = E_t(A_\infty) - A_t \quad (3.4)$$

for some previsible increasing integrable process A . The potential approach therefore requires us to find tractable forms of previsible increasing process to build models. We do not need to look very far; by (2.1), prices are to be expressed in terms of conditional expectations of random variables whose values are yet to be revealed, and for tractability we will want such conditional expectations to be expressible simply in terms of a few variables. Thus we are inevitably drawn towards modelling in the context of Markov processes.

4 Markov processes and potentials

If (X_t) is a Markov process on a statespace \mathcal{X} , and $f : \mathcal{X} \rightarrow [0, \infty)$, then for any $\alpha > 0$ we may consider the increasing process

$$A_t = \int_0^t e^{-\alpha s} f(X_s) ds. \quad (4.5)$$

This is adapted and continuous (therefore previsible), and under mild conditions on f (uniform boundedness will be sufficient but far from necessary) it will also be integrable. From the discussion of Section 3 we can use this to build a pricing model; we find that

$$\zeta_t = E_t \left[\int_t^\infty e^{-\alpha s} f(X_s) ds \right] = e^{-\alpha t} R_\alpha f(X_t), \quad (4.6)$$

where $(R_\alpha)_{\alpha>0}$ is the so-called *resolvent*⁵ of the Markov process.

Though this is not by any means the only way that we could use a general Markov process to build a potential pricing model (see [5] for other ideas), it is sufficiently explicit for us to appreciate immediately how flexible and simple this modelling methodology will be:

(i) We can choose *any* non-negative function f on the statespace, and *any* positive α .

(ii) The decomposition (3.4) of ζ into a martingale less an increasing process takes a very simple form. If we make the usual interpretation of the supermartingale ζ

⁴Explicitly, that the process ζ should be of class (D) - see, for example, [10].

⁵Equation (4.6) is the definition of the resolvent. This is an important and familiar concept from the theory of Markov processes; see, for example, [10].

as the product of a positive change-of-measure martingale Z times the discount factor $\exp(-\int_0^t r_s ds)$, then we have two decompositions of ζ using Itô's formula:

$$\begin{aligned} d\zeta_t &= \zeta_t(dM_t - r_t dt) \\ &= dN_t - e^{-\alpha t} f(X_t) dt, \end{aligned}$$

where M and N are two local martingales, so equating the finite-variation parts gives us

$$r_t = \frac{f(X_t)}{R_\alpha f(X_t)}, \quad (4.7)$$

an explicit expression for the spot-rate process r as a function of the underlying Markov process X .

(iii) There are few examples where the resolvent of a Markov process can be written in closed form (though see Section 6). Nevertheless, using the relation

$$R_\alpha = (\alpha - \mathcal{G})^{-1}$$

between the resolvent and the infinitesimal generator \mathcal{G} of the Markov process, we may build examples by firstly choosing $g \equiv R_\alpha f$ and then recovering f by the recipe $f = (\alpha - \mathcal{G})g$. There is no guarantee that the f so constructed will be non-negative, but choice of α allows considerable leeway here. See [5] for this approach in use in a number of examples.

Before we take up the theme of explicit construction of models based on Markov chains, we digress to point out how simply foreign exchange (and more general asset classes) can be incorporated in the potential approach.

5 Foreign exchange in the potential approach

Suppose now that we wish to consider the pricing of assets in many countries at once, each asset's price being expressed in the currency of its home country. This kind of problem arises quite frequently in practice; we may be asked to price a swap which swaps floating USD interest payments for fixed EUR interest payments. In a conventional approach to such a problem, one would firstly build a model for the interest rates in the US, then a model for the interest rates in Euroland, and then try to model the USD/EUR exchange rate. Even using extremely simple models, a conventional approach would need one driving Brownian motion for the USD yield curve, one for the EUR yield curve and one for the exchange rate - three Brownian motions in total. Bearing in mind that a pricing calculation is in effect an integration, we are beginning to hit problems of dimensionality; a pricing calculation is an integration over three dimensions, and

a pricing calculation for an American-style option is an optimal stopping problem in three dimensions. Add to this the facts that no-one would use a one-factor model to model the USD yield curve (unless forced to by tractability considerations); and that there might well be some knockout feature based on some other exchange rate, and the complexity of pricing such an asset becomes very real.

Or at least it does if you want to use a conventional approach. But let's see how easy it becomes using the potential approach. To introduce some notation, suppose that

$$1 \text{ unit of currency } j = Y_t^{ij} \text{ units of currency } i$$

Now if (S_t^j) is a traded asset in country j , then

$$\zeta_t^j S_t^j \text{ is a martingale;}$$

also, by converting its currency- j price into currency i , it becomes a traded asset in country i , and so

$$\zeta_t^i Y_t^{ij} S_t^j \text{ is a martingale.}$$

Therefore

$$N_t^{ij} \equiv \frac{\zeta_t^i Y_t^{ij}}{\zeta_t^j}$$

is a martingale orthogonal to the space of martingales of the form $\zeta^j S^j$. Thus we can express the exchange rate Y^{ij} as

$$Y_t^{ij} = \frac{N_t^{ij} \zeta_t^j}{\zeta_t^i}.$$

In a complete market, N^{ij} must be constant, so we have the simple and appealing result that *in a complete market, the exchange rate between two countries is the ratio of the state-price densities in the two countries*. More generally, there is the possibility of some exchange-rate risk not hedgeable through other assets, represented by the martingale N^{ij} .

The beauty of the potential approach based on Markov processes is that *adding another country does* not mean adding extra sources of randomness; we simply need to build another state-price density over the *same* Markov process, which requires us only to choose a new f and α . Thus adding extra countries to the model does not need to cause problems of dimensionality (though one may well find that the treatment of the martingales N^{ij} needs to be handled cleverly so as not to lose the simplicity of the methodology.)

6 Markov chain potential models

What makes a good model for an academic is not the same as what makes a good model for a practitioner. The academic is looking for something with simple features and closed-form expressions for basic derivatives, whereas the practitioner recognises that most derivative prices and hedges will have to be computed numerically, so demands quick and accurate numerical algorithms for doing these calculations, and a decent fit to market data. It seems in general that the better a model is for one purpose the worse it is for the other!

When it comes to using the potential approach and a Markovian model, we see from (2.1) that any pricing calculation is an *integration*, and that if we are to do this numerically then we have somehow to compute a finite weighted sum over the statespace of the Markov process. Since this is so, it seems natural to work from the start with a Markov process with a finite statespace, that is, a Markov chain!

Making the assumption that the statespace \mathcal{X} is a finite set of size N has several very clear advantages:

(i) the generator of the chain is a $N \times N$ matrix Q , in terms of which the transition semigroup can be expressed as

$$(p_t(x, y))_{x, y, \in \mathcal{X}} = \exp(tQ);$$

(ii) all calculations reduce to calculations with finite matrices, and are therefore *fast*;

(iii) no splining of functions onto some finite subset of \mathcal{X} will ever be needed;

(iv) pricing of American-style options becomes an optimal-stopping problem for a finite Markov chain, which is easy to handle;

(v) we are not restricted to possibly irrelevant properties of the underlying process (such as path continuity in the case of a diffusion).

Opposed to this are two disadvantages:

(i) the size of the parameter space is $O(N^2)$, so gets quite large quite quickly;

(ii) a given model will only admit N possible values for the price of any given asset.

This latter is apparently quite restrictive, because if we were working with a 9-state chain, then the model says that only 9 possible yield curves could ever

be observed, which is simply incompatible with a casual daily observation of the interest rates reported in any decent newspaper. Our resolution of this is to interpret those prices as being in some sense a ‘market average’ of the ‘pure’ prices that would apply if we knew with certainty what state we were in. We shall explain in more detail how this may be handled in the next Section on calibration, where we address the key question, ‘*Does this work?*’

7 Calibration

The methodology outlined here is very similar to that of [8], with a couple of important variations that substantially improve the performance of the fitting. The first is to drop the restriction to symmetrizable Markov chains, used in [8] to ensure that the diagonal matrices to be computed remain real, and the second is to allow the constant α of Section 4 to become a function of the state. Thus instead of the additive functional A defined at (4.5), we shall be using

$$A_t = \int_0^t \exp\left(-\int_0^s \alpha(X_u) du\right) f(X_s) ds.$$

This change was introduced as a result of experience with the calibration presented in [8]; the goodness of fit seemed to depend quite sensitively on the (previously assumed constant) value of α , and thus allowing α to depend on state seemed a natural (and as it turned out helpful) variation to consider.

The model is parametrised by a vector⁶ θ . On day n we have a vector y_n of observations⁷. If the model were correct, the value of this observation vector y_n would be exactly equal to the model values $Y(X_n, \theta)$, but we suppose that the $\log(y_n)$ are $\log Y(X_n, \theta)$, plus some independent Gaussian noise. We adopt a Bayesian standpoint, and suppose that the initial law of X is given by $\pi = (\pi_i)_{i=1}^N$, and the initial law of θ is given by density $f_0(\theta)$; conceptually, θ is unchanging with time, even though our knowledge of it varies⁸. The notation $\mathbf{z}_n \equiv (z_0, \dots, z_n)$ serves to make formulae more compact.

Based on the assumptions above, and ignoring irrelevant constants, the likelihood

⁶For us, θ is the stacked vector of the off-diagonal entries of Q , the vector $g \equiv (\alpha - Q)^{-1}f$ and the vector α for each country involved.

⁷These observations will be market prices of certain assets.

⁸We shall soon consider what happens if we modify this assumption.

Λ_n of $(\mathbf{X}_n, \mathbf{y}_n, \theta)$ is

$$\begin{aligned}\Lambda_n &\equiv \Lambda_n(\mathbf{X}_n, \mathbf{y}_n, \theta) \\ &= f_0(\theta) \pi_{X_0} \prod_{j=1}^n p_{X_{j-1}X_j}(s_j; \theta) \exp[-b(y_j, Y(X_j; \theta))]\end{aligned}\tag{7.8}$$

where $p_{ij}(s; \theta) = P_\theta(X_s = j | X_0 = i)$, and $b(y, y') \equiv \frac{1}{2} \log(y/y') \cdot V^{-1} \log(y/y')$, where V is the covariance matrix of the Gaussian errors. We have also used the notation $s_j = t_j - t_{j-1}$ for the time between the $(j-1)$ th and j th observations. We shall be more interested in the posterior distribution of (X_n, θ) given \mathbf{y}_n , so we introduce the notation

$$L_n(x, \mathbf{y}_n, \theta) = \sum_{\mathbf{X}_n: X_n=x} \Lambda_n(\mathbf{X}_n, \mathbf{y}_n, \theta),\tag{7.9}$$

and notice that

$$L_n(x, \mathbf{y}_n, \theta) = \sum_{\xi} L_{n-1}(\xi, \mathbf{y}_{n-1}, \theta) p_{\xi x}(s_n; \theta) \exp[-b(y_n, Y(x; \theta))].\tag{7.10}$$

It is clear that for the Markov chain model in mind this expression will be far too complicated to allow exact analysis, so we make some simplifying assumptions, *specifically we assume that the likelihood L_n has the product form*

$$L_n(x, \mathbf{y}_n, \theta) = \pi_n(x, \mathbf{y}_n) l_n(\theta, \mathbf{y}_n).\tag{7.11}$$

The justification is that if we have seen so much data that we have a pretty good idea what the values of the parameters must be, then the values of θ will largely be determined by the long-run historical average behaviour of the system. On the other hand, the posterior distribution of X_n will be more influenced by recent history, because of the ergodicity of the Markov chain, and so some approximate conditional independence is reasonable; recent history tells us all we can know of X_n , distant history tells us all we can know of θ . We shall further assume that

$$l_n(\theta, \mathbf{y}_n) \propto \exp\left(-\frac{1}{2}(\theta - \hat{\theta}_n) \cdot S_n(\theta - \hat{\theta}_n)\right)\tag{7.12}$$

for some positive-definite symmetric matrix S_n . If we think that we have nearly identified the true value of θ , then such a quadratic approximation to the likelihood is quite natural.

The values $\hat{\theta}_n$, S_n , and $\pi_n(\cdot, \mathbf{y}_n)$ are computed recursively, using (7.11). Supposing that we know already $\hat{\theta}_{n-1}$, S_{n-1} , and $\pi_{n-1}(\cdot, \mathbf{y}_{n-1})$, returning to (7.10) and using

(7.11) gives

$$\begin{aligned}
L_n(x, \mathbf{y}_n, \theta) &= \sum_{\xi} \pi_{n-1}(\xi, \mathbf{y}_{n-1}) l_{n-1}(\theta, \mathbf{y}_{n-1}) p_{\xi x}(s_n; \theta) \exp[-b(y_n, Y(x; \theta))] \\
&\propto \sum_{\xi} \pi_{n-1}(\xi, \mathbf{y}_{n-1}) p_{\xi x}(s_n; \theta) \exp[-b(y_n, Y(x; \theta))] \\
&\quad \cdot \exp\left[-\frac{1}{2}(\theta - \hat{\theta}_{n-1}) \cdot S_{n-1}(\theta - \hat{\theta}_{n-1})\right] \quad (7.13)
\end{aligned}$$

Sum this expression over x , and numerically pick θ to maximise; the maximising value is our new estimate $\hat{\theta}_n$ of θ . By computing the second derivative matrix with respect to θ at $\hat{\theta}_n$ we find the value of S_n , and finally we get π_n from

$$\pi_n(x, \mathbf{y}_n) \propto \sum_{\xi} \pi_{n-1}(\xi, \mathbf{y}_{n-1}) p_{\xi x}(s_n; \hat{\theta}_n) \exp[-b(y_n, Y(x; \hat{\theta}_n))].$$

Strictly speaking, the posterior π_n for X_n should be obtained by integrating the likelihood (7.13) with respect to θ , but we approximate this by assuming that the posterior distribution for θ can be replaced by the point mass at $\hat{\theta}_n$, to avoid the need to integrate over a large number of dimensions.

As we indicated in the previous Section, this approach was modified in one vital respect; the model values $Y(x; \theta)$ were replaced by averaged values, averaging with respect to the weights

$$\sum_{\xi} \pi_{n-1}(\xi, \mathbf{y}_{n-1}) p_{\xi x}(s_n; \theta) \cdot \exp\left[-\frac{1}{2}(\theta - \hat{\theta}_{n-1}) \cdot S_{n-1}(\theta - \hat{\theta}_{n-1})\right]$$

8 Evidence from bond data

The data used here is daily yield curve data covering the period from 2nd January 1992 to 1st March 1996⁹

For each day we have values of the yield of bonds with maturity 1 month, 3 months, 6 months, 1 year, 2 years, 5 years, 7 years and 10 years. We shall use daily yield curve data for three currencies; these are sterling (GBP), the US dollar (USD) and the German Mark (DEM).

This introductory account is not the place to present an exhaustive analysis of the results of the fit. We simply report that on average the fitted values were

⁹We are grateful to Dr Simon Babbs for supplying the GBP and DEM data. The USD data was taken from the website <http://www.stls.frb.org/fred/index.html>

within 2-5 basis points¹⁰ of the actual values, with extreme bad fits of the order of 25 basis points for some bonds on some days. The number of states of the Markov chain used was not large, varying between 5 and 11. What we did find is that the calibration was good up to a point, and then the methodology we were using would ‘lose the plot’ - this would typically be after 15-20 days of running the algorithm. The graph of the mean error per fitted bond price (using $N = 9$) indexed by day is given in Figure 1.

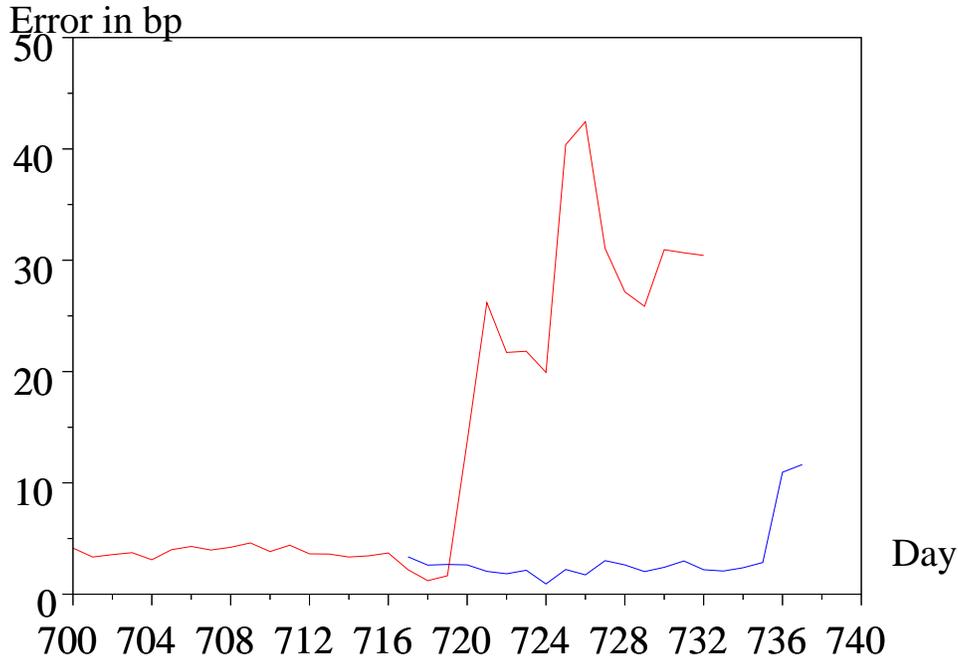


Figure 1: Mean absolute deviation of prices of 8 bonds in 3 countries

We see here two plots, one starting on day 700, which seems to stay close to the data until about day 718, after which it goes astray. Also plotted is the result of a fit begun on day 717, which tracks the data well until about day 736, when it in its turn loses the data. This instability of the fitting process is of course very undesirable; recent work on improvements of the methodology allow us to track the data with a mean error per fitted bond of 2-5 basis points over prolonged periods; we manage to hold this quality of fit over all the subsets of the data that we have tried (currently up to 150 trading days).

This is a remarkably good fit based on an extremely simple model, using only 9 possible states. At one level, it is surprising that such a simple system can do

¹⁰1 basis point = 10^{-4} .

such a good job (if we were able to fit to within 1 basis point, then we would have something that could be traded off). On the other hand, Figure 2 show why this



Figure 2: 1m LIBOR and Bank of England base rate

may not be so surprising; a plot of 1-month LIBOR along with the Bank of England's base rate shows close agreement; if 1m LIBOR is really very close to the base rate, then we should be able to do a good job modelling interest rates if we had done a good job modelling the base rate, and it is not unreasonable to consider a model for the base rate that takes only a few possible values - indeed, we expect that whatever today's base rate is, the base rate in three months from now will differ by 25, 50, 75 or 100 basis points!

9 Hedging

In conventional models, the standard way to hedge a derivative is to *delta-hedge* it. The idea here is to compute the differential of the price of the derivative with respect to the prices of the underlying instruments (so in the case of a put option, we differentiate with respect to the stock price). The differential tells us how many units of the underlying to hold to protect (to leading order) against the moves in the underlying. In the case of a complete market, this hedging methodology is exact, in the sense that if we follow it perfectly, then we will perfectly replicate the contingent claim we were trying to hedge.

If we are using a Markov chain potential model, the notion of differentiating has no meaning, nevertheless the philosophy of immunising our portfolio against possible changes will work just as well. Suppose that we have a derivative Z , and hedging instruments $z^{(1)}, z^{(2)}, \dots$. Suppose that if the state of the chain at time t is i and it jumps to j then the value of Z changes by $\Delta Z_{ij}(t)$. Then what we will do is to hold $w_r(t)$ units of asset r so that

$$\Delta Z_{ij}(t) + \sum_{r=1}^m w_r(t) \Delta z_{ij}^{(r)}(t) = 0 \quad \forall j \quad (X_t = i).$$

Thus whatever jumps of the chain occur, our hedging portfolio will be immune to them. Of course, we do not in practice claim to be able to know X_t , but this does not alter the hedging methodology; we would now make a portfolio of more hedging assets so as to ensure that

$$\Delta Z_{ij}(t) + \sum_{r=1}^M w_r(t) \Delta z_{ij}^{(r)}(t) = 0 \quad \forall i, j.$$

Following this recipe in the case of (say) a 9-state chain would entail taking a position in 72 different hedging instruments (if that many were available!) So we see that the practice of this methodology may not be quite so simple as the theory, but we can expect that the general approach will be as effective as the delta-hedging methodology is for diffusion-based models.

10 Conclusions and future directions

This brief introduction to the potential approach has shown that this modelling methodology has clear advantages:

- (i) pricing is easy;
- (ii) hedging is easy;
- (iii) handling many currencies is easy;
- (iv) calibration is perfectly feasible.

It is my belief that if there is ever to be a universal modelling methodology, it will have to look something like this. By ‘universal’ I mean a model that will account for all the different asset classes that an investment bank deals with - equity, foreign exchange, fixed income, commodities, credit risks - and the reason that a bank would like such a model is principally for what is known as *risk management*

(though more properly called risk measurement). This refers to the regulatory requirements placed on the bank to assess the riskiness of their positions, and this needs to be understood on a firm-wide basis, as well as by business unit. The potential approach to modelling really can embrace such a wide swathe of the bank's business, and even if the calculations may have to be approximate to deal with such a wide sweep, at least the approach is consistent over the whole, rather than being some patched-together pastiche of wholly different models.

So far, the potential approach has been tested only on some bond data, and the next stage of the checking has to be to try to fit other fixed-income derivatives, then extend to other asset classes, notably equities. The only obstacle here is in obtaining decent data to work with, as many of the products that are important for calibration are not traded on exchanges, so price data is hard to come by. But given such data, we can come closer to the final verdict on the potential approach: is it simply the right way to model prices, or is it a nice idea that cannot cope with the full complexity of the real data?

References

- [1] BAXTER, M., AND RENNIE, A. *Financial Calculus*, Cambridge University Press, Cambridge, 1996.
- [2] DANA, R., AND JEANBLANC, M., *Financial Markets in Continuous Time*, Springer, Berlin, 2003.
- [3] DELBAEN, F., AND SCHACHERMAYER, W., A General Version of the Fundamental Theorem of Asset Pricing, *Mathematische Annalen* **300**, 463–520, 1994.
- [4] KARATZAS, I., AND SHREVE, S. E., *Methods of Mathematical Finance*, Springer, New York, 1998.
- [5] ROGERS, L. C. G., The potential approach to the term structure of interest rates and foreign exchange rates, *Mathematical Finance* **7**, 1997, 157–176.
- [6] ROGERS, L. C. G., One for all, *RISK* **10**, 57–59, March 1997.
- [7] ROGERS, L. C. G., The origins of risk-neutral pricing and the Black-Scholes formula, in *Handbook of Risk Management and Analysis*, ed. C. O. Alexander, Wiley, Chichester, 1998.
- [8] ROGERS, L. C. G., AND YOUSAF, F. A., Markov chains and the potential approach to modelling interest rates and exchange rates, in *Mathematical Finance - Bachelier Congress 2000*, ed. H. Geman, D. Madan, S. R. Pliska, & T. Vorst, Springer, Berlin 2002.
- [9] ROGERS, L. C. G., AND ZANE, O., Fitting potential models to interest rates and foreign exchange rates, in *Vasicek and Beyond*, ed. L. P. Hughston, RISK Publications, London 1997, 327–342.
- [10] ROGERS, L. C. G., AND WILLIAMS, D., *Diffusions, Markov Processes, and Martingales, Volumes 1 and 2*, Cambridge University Press, Cambridge, 2000.