ONE FOR ALL

Chris Rogers*
University of Bath

This article contains no numerical examples, and no specific algorithms for pricing or hedging formulae; it is about general principles of modelling for financial application which are almost obvious once one thinks about the problems the right way. But the ideas presented here arose in the context of the potential approach to interest rate modelling,[1] and it does no harm to discuss how they apply to that, as we soon shall.

The foundation of all pricing recipes is the *arbitrage pricing principle* which says that if Y_T is the price of a traded asset at time T, then the price at time $t < T$ can be represented as

$$Y_t = \mathbb{E}_t[\exp(-\int_t^T r_s ds)\, Y_T], \tag{1}$$

where $(r_t)_{t\geq 0}$ is the *spot-rate process*, and $\mathbb{E}_t$ is the conditional expectation in a/the risk-neutral measure $\mathbb{P}$ given all information available at time t. All pricing formulae are applications of this general principle. So if the asset is a zero-coupon bond which matures at time T, we deduce that the price at time t of the bond is

$$P(t,T) = \mathbb{E}_t[\exp(-\int_t^T r_s ds)].$$

One approach to modelling the term structure is to model the spot rate process in the risk-neutral measure (this is the approach of Vasicek [12], Cox, Ingersoll & Ross [4], Brennan & Schwartz [2], Richard [9], ...). Once one has specified the law of r under $\mathbb{P}$, one can in principle compute prices of bonds and various interest rate derivatives, though the calculations are often quite involved. Is this approach good enough? Well, if we are only concerned with interest rate derivatives, it is; there are other frameworks within which we could describe the interest rate model (for example, the whole yield framework of Ho & Lee [7], Heath, Jarrow & Morton [6], Babbs [1],...) but these other descriptions are entirely equivalent. But if we are interested in *more* than just interest rate derivatives, what else should we be considering? The only possible answer to this question is, '*Everything!*' This is clear; look at the expression (1) for the price of the asset, where the spot rate enters explicitly.

To be more concrete, suppose we had built a good model for US interest rates, which fitted observed data well, and agreed well with prices of US interest rate derivatives. If

* Supported by EPSRC grant GR/J97281. This article has appeared in *RISK Magazine* **10**, March 1997.

[1] See 'The potential approach to the term structure of interest rates and foreign exchange rates', L C G Rogers, University of Bath preprint, 1995.

we set this up and begin trading, how long will it be before a client comes along wanting to swap floating US interest payments for fixed sterling payments? Once this happens, our model cannot cope. So we could try to make a similar model of UK interest rates, and calibrate this to UK data, and we would also have to think of some model for the exchange rates; if we do this by bringing in new sources of randomness, we are doomed. We are doomed because before long we shall have to price deals involving DEM, JPY, CHF,.... and if our response each time we meet another currency is to add further sources of randomness, then before long we may have a model with 20 Brownian motions, and pricing PDEs in this many dimensions. Quite apart from the difficulty of computation, can one be sure that when one adds another currency one may not have created arbitrage opportunities for sharp-eyed competitors?

What is the alternative? The problem we face is similar to that faced by an unsophisticated user of probability; he thinks of a single random variable in terms of its density, then he thinks of two random variables in terms of their bivariate joint density, and three random variables in terms of their trivariate joint density, and understandably finds it hard to think of an infinite sequence of random variables. But the sophisticated user knows that all random variables are mappings from the sample space to the real line, and has no difficulty in conceiving of even an uncountable family of them. The unsophisticated user comes unstuck because he is changing the underlying model every time a new random variable comes along, and this is just the problem we face if we add more randomness when we try to incorporate more currencies (or more of any sort of asset).

So our alternative will be to keep the *same* model all the time, and express the new assets *in terms of that same model.* It may be that we have to allow a limited amount of country-specific randomness, but that can easily be accommodated; it does not affect the basic principle of using a common model for all countries.

2. Generalities on interest rate modelling. To show the sort of thing which happens, let us look at interest-rate modelling in another way. Suppose that we have some reference probability $\tilde{\mathbb{P}}$ with respect to which the risk-neutral probability $\mathbb{P}$ has a density

$$\rho_t \equiv \left.\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}\right|_{\mathcal{F}_t}$$

on the collection $\mathcal{F}_t$ of events determined by time t, for every t. Now we can re-express the bond prices as

$$\begin{aligned} P(t,T) &= \mathbb{E}_t[\exp(-\int_t^T r_s ds)] \\ &= \tilde{\mathbb{E}}_t[\rho_T \exp(-\int_t^T r_s ds)]/\rho_t \\ &= \tilde{\mathbb{E}}_t[\zeta_T]/\zeta_t, \end{aligned}$$

where the process ζ is the so-called *state-price density process*, defined by

$$\zeta_t \equiv \exp(-\int_0^t r_s ds)\rho_t.$$

This final expression for the bond price corresponds to a quite different approach to the modelling of interest rates, named the *potential* approach in [10]. The first published appearance of this was in Constantinides [3]. This approach is to specify the distribution of the $\tilde{\mathbb{P}}$-supermartingale ζ, and then compute bond prices (and the prices of other interest-rate derivatives) as expectations with respect to the reference measure. If $\tilde{\mathbb{E}}[\zeta_T] \to 0$ as $T \to \infty$ (which is reasonable on economic grounds), then the supermartingale ζ is what is called a *potential* in the general theory of process, which is why the name for this approach was chosen. In some senses, this is simply a difference of viewpoint, but the payoff comes almost immediately, when we try to understand the term structure in many countries at once.

3. How FX enters. Suppose that in country i the state-price density is ζ^i, and that at time t

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

Thus if S_t^j is the price of a share in country j at time t, we know that $Y_t^{ij} S_t^j$ is the value of this expressed in country i's currency at time t. Assuming that there are no frictions, the arbitrage pricing principle says that

$$\zeta_t^i Y_t^{ij} S_t^j \text{ is a } \tilde{\mathbb{P}}\text{-martingale;}$$

but on the other hand S^j is a country-j traded asset, so we also have

$$\zeta_t^j S_t^j \text{ is a } \tilde{\mathbb{P}}\text{-martingale.}$$

Assuming complete markets, the only possible conclusion is that the two processes multiplying S^j are, up to a multiplicative constant, the *same* [2]:

$$Y_t^{ij} = Y_0^{ij} \zeta_t^j / \zeta_t^i.$$

This strikingly simple formula is apparently known,[3] though perhaps not as widely as it should be. Thus if we use the potential approach, once we have selected the state-price densities in each country, we have automatically determined the exchange rates between them in a wholly consistent manner!

4. The potential approach. Is it easy to find examples where one can model the state-price density sufficiently explicitly to permit simple calculations? Indeed it is. One place where potentials arise naturally is in Markov processes; if we have a Markov process X and take any non-negative function g, and any $\alpha > 0$, then the process

$$\zeta_t \equiv e^{-\alpha t} R_\alpha g(X_t)$$

[2] Without the completeness assumption, this may hold, but more generally one gets that $Y_t^{ij} = N_t \zeta_t^j / \zeta_t^i$ for some positive $\mathbb{P}^j$-martingale strongly orthogonal to all the $\mathbb{P}^j$-martingales of the form $\exp(-\int_0^t r_s^j ds) S_t^j$.

[3] It appears, for example, in a preprint 'The dynamics and the term structure of risk premia in foreign exchange markets', J. Saá Requejo, 1993.

is a potential, where R_α is the so-called resolvent of the Markov process,

$$R_\alpha g(x) \equiv \tilde{\mathbb{E}}[\int_0^\infty e^{-\alpha t} g(X_t) dt | X_0 = x].$$

If we use such a recipe for the state-price density, we get a simple expression for the spot rate:

$$r_t = g(X_t)/R_\alpha g(X_t),$$

and simple expressions for a range of derivative prices [4]. If one assumes a simple enough form for the Markov process and the function g, then closed-form expressions for prices can often be found; see [10] for examples.

5. Calibration & pricing. Pricing is very closely related to the calibration of models. If one has a parametric model, then the values of various market observables can be expressed as functions of the parameters θ and the underlying state X of an explanatory process (so, for the Vasicek model, this would simply be the spot rate process.) Suppose that observable i (which may be a price, or a historical volatility, for example) has model value $f^i(\theta, x)$ if the state is x (such functions would typically be quite complicated, and would be coded up as a subroutine) and that on day n the observed value is f^i_n. One thing we could do is to

$$\min_{\theta, x} L(f_n - f(\theta, x)),$$

where L is some loss function, for example, quadratic. This is not ideal, though, because it fits different θ to different days. It would be preferable from a theoretical point of view to calibrate with θ fixed. But this is going to be numerically quite heavy (minimising over θ and over $x_1, \dots, x_N$), and we don't *really* believe this anyway. We should expect and allow the parameters to move very gradually, and one way this can be achieved is to allow the parameters to diffuse slowly as Brownian motions. If we are careful, we can build a recursive procedure which not only updates our estimates of X and θ, but also tells us about the precision of our estimates; more details for the the Gaussian example above are given in the Appendix, and the full story is in Rogers & Zane [11]. The procedure (which is essentially an approximate Kalman filter) not only gives a point estimate for the unknown parameters θ and the underlying process X, but it also gives a guide to the error in those estimates. Thus we can deduce *confidence intervals* for the prices which come out of the model.

A generic calibration procedure would select certain market observables which were to be used for the calibration, and would recursively update the estimation procedure outlined above. The calibrated model would then be used to price derivatives, and if any

[4] For example, the price of a zero-coupon bond is

$$P(0, t) = e^{-\alpha t} P_t R_\alpha g(X_t)/R_\alpha g(X_0),$$

where $(P_t)_{t \geq 0}$ is the transition semigroup of the Markov process.

were found which were poorly priced by the model, then it would make sense to enlarge the calibration to use the prices of those derivatives too.

Note that there is freedom to choose a rigidity parameter ε and a matrix V_η which defines the emphasis attached to the goodnes-of-fit; see the Appendix. In the extreme case where $\varepsilon = 0$, we have a model with fixed parameters whereas if we let ε get large, then we let the parameters move freely. There is a trade-off, of course; the more constrained the parameters, the worse the fit of the model to the data is bound to be. The user is free to choose the desired level of rigidity for the model. Different weights in V_η will allow different emphasis on the closeness of fit to the observed values; for example, we may well want to insist that the most liquid derivatives must be priced with greater precision.

In many applications, a user will want an *exact* fit to the market data which was used to calibrate the model. This may in practice be accommodated by fitting a time-inhomogeneous model, but I feel it is probably best to make a running estimate of a time-homogeneous parametric model, and treat the market data as a perturbation of it. This way, one has a better chance of stability than with the usual practice of perfectly fitting the data each new day. Alternatively, and preferably, we may regard the purpose of a model to be the indication of an *interval* in which it is reasonable to trade, the actual price at which the deal is done being a matter for the trader; in that case, so long as the interval output by the model always covers the market prices used to calibrate, there is no need to do anything more to the model. Flesaker and Hughston [5] in a different approach to interest-rate modelling study examples where the state-price density (in the language of the current approach) can be expressed in the form $a(t) + \tilde{b}(t)M_t$, where a and b are positive deterministic decreasing functions, and M is a positive $\mathbb{P}$-martingale; this gives a flexible and effective approach to fitting a time-dependent model.

6. Hedging. In most diffusion models of term structure, there is enormous redundancy in hedging. For example, if one takes a complete two-factor model such as the model of Longstaff & Schwartz [8], then (in theory) *any* interest-rate derivative out to 5 years can be hedged using just two bonds of maturity greater than 5 years - and in principle *any* two bonds will do! We could even hedge using a 10-year cap in place of one of the bonds, and any such hedge is theoretically as good as any other because they all perfectly replicate the derivative! Yet in practice we have to pay transactions costs and we have imperfect estimates of the parameters of the model, so in the real world one of these theoretically-equivalent hedges may be vastly preferable to another; we shall have to return to this on another occasion....

Appendix. Let's see how the recursive updating of estimates works in the context of the Gaussian diffusion

$$dX_t = dW'_t - BX_t \, dt$$

introduced in the text. The parameters θ are the entries of B, which we shall suppose move like slow Brownian motions:

$$d\theta_t = \varepsilon dW''_t$$

for some small ε. For ease of notation, we put the parameters and the state variable into one variable, $z = (x^T, \theta^T)^T$ and express the dynamics as

$$dZ_t = \sigma dW_t + b(Z_t)dt.$$

At regular intervals of length δ, we observe $Y_n \equiv f(Z_{n\delta}) + \eta_n$, where f is some vector-valued function (typically the prices of some assets) and η is a vector of zero-mean Gaussian noises with covariance V_η. Suppose that we have a posterior distribution for $Z_{n\delta}$ at time $n\delta$ which is $N(\mu_n, V_n)$. Then defining

$$\begin{aligned}\tilde{V}_n &\equiv \delta\sigma\sigma^T + (I + \delta Db(\mu_n))V_n(I + \delta Db(\mu_n))^T \\ c_n &\equiv \mu_n + \delta b(\mu_n) \\ \Phi_{n+1}(z) &\equiv \frac{1}{2}(Y_{n+1} - f(z)) \cdot V_\eta^{-1}(Y_{n+1} - f(z)) + \frac{1}{2}(z - c_n) \cdot \tilde{V}_n^{-1}(z - c_n),\end{aligned}$$

we solve the problem

$$\min_z \Phi_{n+1}(z)$$

and take the minimising z to be μ_{n+1} and the second derivative of Φ there to be V_{n+1}^{-1}.

If the underlying diffusion were Gaussian, and f were linear, then this procedure is exactly the Kalman filter.

References.

[1] BABBS, S., A family of Itô process models for the term structure of interest rates. University of Warwick, preprint 90/24, 1990.

[2] BRENNAN, M. J. and SCHWARTZ, E. S. A continuous time approach to the pricing of bonds, *J. Banking Fin.* **3**, 1979, 133–155.

[3] CONSTANTINIDES, G., A theory of the nominal term structure of interest rates, *Rev. Fin. Studies* **5**, 1992, 531–552.

[4] COX, J. C., INGERSOLL, J. E., and ROSS, S. A., A theory of the term structure of interest rates, *Econometrica* **53**, 1985, 385–408.

[5] FLESAKER, B., and HUGHSTON, L. P., Positive interest, *RISK* **9**, 1996, 46–49.

[6] HEATH, D., JARROW, R., and MORTON, A., Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation, *Econometrica* **60**, 1992, 77–105.

[7] HO, T. S. Y. and LEE, S.-B., Term structure movements and pricing interest rate contingent claims, *J. Finance* **41**, 1986, 1011–1029.

[8] LONGSTAFF, F. A. and SCHWARTZ, E. S., Interest-rate volatility and the term structure: a two factor general equilibrium model, *J. Finance* **47**, 1991, 1259–1282.

[9] RICHARD, S. F., An arbitrage model of the term structure of interest rates, *J. Fin. Econ.* **6**, 1978, 33–57.

[10] ROGERS, L. C. G., The potential approach to the term structure of interest rates and foreign exchange rates, University of Bath preprint, 1995.

[11] ROGERS, L. C. G. and ZANE, O., Fitting potential models to interest rate and foreign exchange data, University of Bath preprint, 1996.

[12] VASICEK, O. A., An equilibrium characterization of the term structure, *J. Fin. Econ.* **5**, 1977, 177–188.