Interest-Rate Option Models, by Riccardo Rebanato

Chapter 12, The Black Derman and Toy Model

Section 12.1. Introduction

Section begins with recalling Chapter 8, where much of the machinery of computational trees was developed, specifically with Ho-Lee (HL), Hull-White (HW), and Black-Derman-Toy (BDT) in mind.

Observation from Ken: *Having read Chapter 8, and before reading the bits about HL and Hull-White, it seems that the main difference between HL and BDT is that HL uses normal dynamics for the spot rate, whereas BDT uses lognormal dynamics. So HL has many more analytical results, because of the ease of working with normal dynamics. But BDT has advantages, like the impossibility of the spot rate becoming negative.*

The main points that were concentrated on in Chapter 8 about computational lattices or trees are that:

- The tree can be calibrated to the market. That is, it is possible to specify the dynamics of the spot rate inside the tree so that the zero curve implied by the tree agrees exactly with the currently observed market zero curve.
- The practice of stepping through the tree and continually taking the local discounted expected value of the instrument is validated. That is, it is shown that this numerical procedure for pricing the instrument is a valid discrete-time numerical pricing implementation, that is consistent with the absence of arbitrage.
- A forward induction method, using Green's functions and Arrow-Debreu prices, is described, as an efficient numerical scheme for carrying out the calibration to the market zero curve.

The author explains that the main reasons that the BDT model has been singled out for in-depth analysis here are:

1) Popularity among practitioners
2) Simplicity of its calibration to the market. *(Ken: More comment on this later, but as I presently understand it, this follows from BDT's assumption of lognormal dynamics, the same as for the Black-Scholes volatilities that are quoted in the market.)*
3) Straightforward analytical results, like those available for HL, are **not** available for BDT. *(Ken: Question: Singled out because analytical results not available? This seems strange. Possibly because, for other models, it's easier to deduce analytical results, but for BDT it's harder to do this, so he's going to help us out with this Chapter.)*

Important: the BDT is a one-factor model. It is constructed in such a way that it can price exactly any set of market discount bonds. Does not require specification of investors' risk preferences. *Observation from Ken:* *Interesting to note that it can be calibrated to any set of market discount bonds. For instance corporate bonds, junk bonds, etc. (Obviously it must be a homogeneous set of bonds, for instance all from the same issuer.) Is this right? Then the 'r' you will get will be some yield, not necessarily risk-free. Keep in mind for later.*

Observation that, since swap rates can be viewed as linear combinations of discount bonds, then BDT will price swap rates exactly for any vol input. HL can also do this, but BDT makes assumption of lognormal spot rate. As mentioned before, this prevents negative interest rates. Also eases calibration to market vols, since it follows market conventions. Author makes some notes about how it is now straightforward, both conceptually and computationally, to simultaneously calibrate to the yield curve and to cap or swaption vols. Stay tuned ...

Disadvantage of BDT is that numerical obstacles may loom, since the lognormal distribution is not so tractable as normal. Also, method is given as algorithm, so its implications, hidden assumptions, etc., are not so transparent. (For example, the nature of its mean-reversion.) One objective of chapter is to help display these qualities of model, including strengths and weaknesses.

Section 12.2. Analytic Characterization

Author makes the statement that the discrete-time implementation of this model was presented with the lattices of Chapter 8, so here he concentrates on the continuous-time version and its analytic properties. He writes the form of the short rate as

$$r(t) = u(t)\exp(\sigma(t)z(t)) \quad (1)$$

where:

$u(t)$ is the median of the short rate distribution at time t,
$\sigma(t)$ is the short rate volatility, and
$z(t)$ is a standard Brownian motion.

The author says this is the continuous-time equivalent of the BDT short rate process. It's a reasonable statement, since Eq. (1) is just a statement of a lognormal density. The log of $r(t)$ is a normal random variable with mean 0 and variance $\sigma^2 t$. And because there's no constant term in the exponent, then the exponent has equal probability to be positive or negative. So, $r(t)$ has equal probability to be above or below the median $u(t)$, as desired.

The calibration done in Chapter 8 is mainly aimed at finding $u(t)$, and it is the proper specification of $u(t)$ that allows market instruments to be priced properly. As usual, for HL and HW, it's possible to get an analytical formula for $u(t)$, but the lognormal dynamics prevent that here. *(Ken: Here I think u is just used to study analytical properties. That is, we don't dwell too much on talking about a specific form for u(t).)*

The author next makes the statement that from the dynamical equation (1), you can apply Ito's Lemma to $r = r(t,z(t))$, with the substitution $z(t) = [\ln r(t) - \ln u(t)] / \sigma(t)$ to get

$$dy(t) = \{c(t) + a(t)[w(t) - y(t)]\}dt + \sigma(t)dz \quad (2)$$

where:

$$y(t) = \ln r(t)$$
$$w(t) = \ln u(t)$$
$$c(t) = \partial w(t)/\partial t$$
$$a(t) = - \partial \ln(\sigma(t)) / \partial t$$

Observation from Ken: The author says this statement follows easily from Ito's Lemma, with the given substitution, but I haven't been able to make it follow. I'm just continuing with the Chapter at this point. The author uses the above relation to make various statements about the behavior of the BDT model, and if I can get these, then the chapter won't be a total loss.

An important thing to note about Equation (2) is that, if one assumes a vol that is constant with time, then $a(t) = 0$, and no mean-reversion will take place. In this case, the short rate will evolve by simple diffusion with drift $c(t)$. If the vol is instead decaying with time, then $a(t)$ will come out positive, and the desired mean-reversion will take place. An example is given with $\sigma(t) = \sigma(0)exp(-vt), v>0$.

Observation from the author: Because BDT allows for a general time-dependent short-rate vol, then BDT can recover a whole series of market (Black) cap prices. And these market cap prices will very often have a declining implied (Black) volatility. This will be studied in detail 'later on.'

(<u>Ken:</u> Don't know if 'later on' here means later on in this Chapter, or further on in the book.)

(<u>Ken:</u> Also: I don't know why the (Black) is included above, but it was in the book, so I thought I should copy it down here, just to remember exactly what he said. My guess is that the (Black) refers to the fact that in the official quotes on screens like Reuters or Bloomberg, what's quoted is a vol substituted into the Black model to get the Black price.)

Further, swap rates are also exactly recoverable by the BDT construction. And because swap rates are approximately lognormally distributed, then BDT will give a price very similar to that obtained from the Black model.

(<u>Ken:</u> I'm not sure exactly what he means here by 'swap rates are approximately lognormally distributed.' Does he mean that, by the BDT construction, the model implies an approximately lognormal distribution for the swap rates? Does he mean that in reality, swap rates evolve by an approximately lognormal process? My guess is he means the first. Because, here he's making a statement about how the BDT and Black models give about the same price. And Black assumes a lognormal distribution. So I think this statement is really about similar assumptions leading to similar prices.)

Author makes more statements about what's necessary in this model to obtain mean-reverting behavior for the short rate. Again: vol that decays with time is necessary to obtain mean-reverting behavior. So, BDT allows for the unconditional variance, $\sigma^2(t)t$ to grow unbounded with time, if that's what the fitting from today's caps implies.

However, since BDT allows for fitting of any vol shape, and since it's the vol that largely determines the price, then BDT has the advantage of being able to simultaneously price any number of caps.

It is also possible to simultaneously price a number of options to enter, at a series of times, the same maturity swap.

However, it is <u>not</u> possible to <u>simultaneously</u> price caps and swaptions. This is not a specific flaw of BDT, but rather a flaw of any one-factor model. Any one-factor model must, by construction, assume perfect correlation between rates. Note:

- A cap is portfolio of independent options (caplets), while
- a swaption is an option on a linear combination of forward rates.

So a swaption will depend on the imperfect terminal correlation of the forward rates. It was noted earlier (Chapter 4) that term decorrelation can be brought about by

- imperfect instantaneous correlation (which is not reproducable by the BDT model) and/or by
- non-flat instantaneous volatility term structure (which BDT can handle.)

So, after calibration to desired option prices, you should stop and assess what the implied evolution of the vol term structure is. This will be a unique function of the instantaneous vols of the various forward rates.

(Ken: Possibly the important thing here is that, the cap is a portfolio of independent options. Whereas a swaption is an option on a portfolio of rates. In neither case can you truthfully deal with the imperfect correlation of rates, but the two cases sweep this under the rug in different ways, so that the two taken together can't be consistent.)

One feature specific to BDT is that the vol term structure is completely determined by the future (not forward) vol of the short rate. This stems from the fact that the reversion speed, given above in Eq. (2) as

$$a(t) = - \partial \ln(\sigma(t)) / \partial t$$

is not an independent parameter, but rather is determined by $\sigma(t)$. (In the HW model, you can specify reversion speed as an independent parameter.) Implications of this statement will be discussed in more detail later on in this Chapter.

Section 12.3. Assessing the Realism of the BDT Model

It has been shown in previous sections that the BDT model can be made to fit the current market prices of caps or swaptions. In this section, the author explores how 'natural' this fitting is. In other words, a parameterization procedure for any model should consist of little more than 'fine-tuning' a fundamentally correct phenomenological model. So in this section, the author undertakes the task of assessing the overall realism of BDT.

One shortcoming of BDT is the artificial link between the future volatility of the short rate and the term structure of volatilities.

(Ken: I don't understand why this link is artificial. It seems reasonable to take the future vol of the short rate, at the time of some long maturity, and integrate it backwards in time to the present to get the vol of that maturity length. That doesn't seem artificial.)

Next a graph is given which I don't quite understand. He says it's a graph of volatilities of yields of maturities of different lengths, taken from the market on a specified date. He then says these curves are plotted for different values of the decay constant ν, taken from the expression $\sigma(t) = \sigma_0 exp(-\nu t)$, and with a reasonable value of σ_0. So, my point of confusion is, are these market yields, as he says, or are they yields for different assumptions of the decay constant ν? If they're meant to illustrate different cases of the decay constant, then they're vols taken from a model. But, he says that it's a market graph. So I don't understand what's going on here.

In the end, he says that for positive values of the decay constant ν, there is less variability for yields of longer maturities, which is in overall agreement with market observations.

Next the author makes another statement which I find confusing. He says that exact matching of the vols of yields of different maturities should not be expected, and if it does happen, then it's simply luck. I thought that was exactly the point of BDT: you input the vols for yields of varying maturities from the market, and you input the corresponding yields. Then BDT gives you the short rate as a function of time which exactly reproduce those vols and yields.

There is one more statement about how the qualitative shape of the term structure of volatilities is by and large correctly captured by BDT. But: Nothing has been said about the evolution of the term structure of volatilities.

A second test of the realism of the model, perhaps the most important test, is the hedging performance of the model. Within the framework of any approach, hedging can be approached in two ways:

- Within the model: Attempting to neutralize the exposure to the driving factor(s) of the model

- Outside the model: Obtaining price changes with respect to exogenous shocks, shocks which have virtually zero probability within the model. Examples for BDT include rigid shifts in the yield curve, or shocks to individual forwards.

The Outside the Model approach is conceptually inconsistent, because you stop and give special consideration to events considered extremely unlikely by the model. But, the Within the Model approach requires a lot of faith in the correctness of your model. (In the BDT case, for instance, this leads to one example situation where you're hedging your exposure to a 10-year yield just by making overnight deposits.)

In practice, the Outside the Model approach will always be used to some extent. However, Within-the-Model hedging tests can give very useful indications about the realism of the model.

Obtaining approximations to in-model hedging parameters is straightforward. For two assets *A* and *B*, you can obtain their relative hedge ratio through their sensitivity to the short rate:

$$(3) \qquad [A(up) - A(down)] / [B(up) - B(down)] = [(A(up) - A(down)) / (r(up) - r(down)] \;/\; [(B(up) - B(down)) / (r(up) - r(down)]$$

where the up and down refer to the single branching at the initial time node. This ratio gives a Δ-type hedge ratio, for the sensitivity of asset *A* with respect to asset *B*. A γ-type ratio – which involves second derivatives – can be approximated by bringing in asset values for the second time step – the second branching.

By employing these approximate hedge ratios, you can get the sensitivity of different bond prices to the change in the short rate. These sensitivities will usually show little dependence on the decay constant ν, but instead will show strong dependence on the shape of the yield curve instead. *(<u>Ken:</u> That is, bond price sensitivities shows weak dependence on the decay constant ν, which determines the shape of the vol term structure, but show strong dependence on the shape of the yield curve itself.)*

One of the most stringent tests you can make of the realism of the model is by doing some historical study:

- Take the approximate hedge sensitivities obtained from Equation (3)
- Take some historical time series of data for changes in the short rates
- Find the model predictions, using Eq. (3), of yield changes for the given changes in the short rate
- Compare model answers with market outcome

The author carried out such a study for market data from 1990-93, using market data for various currencies, with decay constants chosen to give a reasonable fit to market cap prices. A graph is given of the results, which shows that the correlation is prefect, by construction, for the short rate, while the correlation decreases with yield maturity.

(<u>Ken:</u> I'm not quite with him on the statement about how correlation is perfect between predicted and market yield changes for the short rate. Is he saying he can perfectly predict changes in the short rate, by fitting the model with the right parameters? I don't think that's right. For sure, the correlation between predicted vs. realized should be the highest for the short rate, but it shouldn't be perfect.)

One result of empirical studies has been the conclusion that BDT yields tend to move in far too parallel a fashion. This is supported by various graphs, which show that the changes in the yield curve at any given time can be very complex, but BDT predictions always show a very simple shape to the changes in the yield curve. The BDT predicted change is mostly flat, possibly rising or falling slightly.

In the end, BDT enjoys several positive features (good agreement with market, when vol term structure is decaying) but suffers from 2 important shortcomings:

- Inability to handle conditions where impact of second (tilt) factor could be of importance. This shortcoming is unavoidable for any one-factor model.
- Inability to specify vols of yields of different maturities independently of the future vol of the short rate. Will be discussed in later chapters how this is relevant to different options. (Example: How does it affect a long-maturity swap with principal determined by a short-maturity index?)

Notice again that nothing has been said about the evolution over time of the vol term structure. This will be dealt with in Chapter 19.

Section 12.4. Derivatives in One-Factor Models: The BDT Case

In this section the author considers a few specific derivatives, as evaluated by the BDT model. He compares some of the implications to those of the Black model. He's especially concenrned with the hedge parameters Δ and γ.

Specifically, follow BDT from their original 1990 paper, and consider the three instruments:

i. A T-maturity coupon-bearing bond B
ii. A call on this bond struck at X, which matures at time $t(opt) < T$
iii. A put on this same bond, also struck at X, also maturing at time the $t(opt) < T$

And consider the strategy of

i. Buying the bond
ii. Buying the put
iii. Selling the call

With this set-up, let the bond B pay n coupons between t_0, the present time, and $t(opt)$, the expiration time of the two options.

With this strategy, then the time-$t(opt)$ payoff of the portfolio will be the strike price X, with certainty. The total certain payoffs over the whole life of the strategy will include the n intermittent coupons C, along with the strike amount X at option expiry.

So the payoff of this portfolio is identical to that of a portfolio of discount bonds, each with face value C, maturing at the times of the intermittent coupons, along with a discount bond of face value X, maturing at the time $t(opt)$. Express this portfolio as $\{\sum Z_i\}$. (Just not to worry too much about notation. $\sum Z_i$ just represents the collection of these discount bonds, with their different expirations and face values.)

Given this set-up, then we can write:

$$P - C + B = \sum Z_i \qquad (4)$$

where P, C, and B represent the put, call, and coupon-bearing bond, respectively.

The BDT model must, by construction, correctly price any discount bond. So, the right-hand side of Eq. (4) is correctly priced. Because of this, the author draws the conclusion that put/call parity is satisfied.

Next the author considers hedge ratios. Specifically, if we take the derivative of the put/call parity equation, Eq. (4), with respect to B, the underlying, then we get

$$\partial P / \partial B \; - \; \partial C / \partial B \; + \; \partial B / \partial B \; = \; \partial(\sum Z_i) / \partial B$$

$$(5) \qquad \Delta_{put} - \Delta_{call} \; = \; 1 - \partial(\sum Z_i) / \partial B$$

The last term on the right-hand side expresses the sensitivity of the portfolio of discount bonds to the price of the coupon-paying bond *B*.

Eq. (5) is different from the corresponding equation you would get from any Black-like model. Specifically, the corresponding equation for a Black-like model comes out as

$$(6) \qquad \Delta_{put} - \Delta_{call} \; = \; 1$$

and the difference in the two equations comes from a different choice of numeraire.

(Ken: I'm taking this last statement completely on faith, since the author gives no reference to an earlier chapter, and no further explanation.)

The same type of reasoning applies to the gamma sensitivities. If we take a second derivative of Eq. (5), then we get

$$\partial^2 P / \partial B^2 \; - \; \partial^2 C / \partial B^2 \; = \; \partial^2(\sum Z_i) / \partial B^2$$

$$(7) \qquad \Gamma_{put} \; - \; \Gamma_{call} \; = \partial^2(\sum Z_i) / \partial B^2$$

Again, this differs from any Black-like model, in which you would get $\Gamma_{call} = \Gamma_{put}$. And, depending on the magnitude of the term on the right-hand side, then, for instance, Γ_{call} may come out positive or negative.

Finally, there are some observations by the author on the sizes of the magnitudes of the hedging parameters. Specifically, the author comments on the two quantities

$$(8) \qquad \partial(\sum Z_i) / \partial B \qquad \text{and} \qquad \partial^2(\sum Z_i) / \partial B^2$$

For options whose maturities are much shorter than the maturity of the underlying bond *B*, then both of these quantities should come out very small, and the associated sensitivities should look like the associated Black sensitivities.

Conversely, for options which have a much longer maturity relative to the maturity of the underlying bond *B*, then these terms should become more important.

Finally, there is a remark that these statements do not imply that the Black model is wrong, but instead they stem from then different choice of numeraire in the two cases. *(Ken: I still don't know what he means by this statement, but I'll remember the point that the Black model was not intended to be shown up here.)*

Section 12.5. Calibrating the BDT Model: Pricing FRA's, Caps, and Swaptions Using Lattice Models

This section lays out specific procedures for fitting the BDT curve to option market prices, in an efficient way. Special attention is devoted to caps, floors, and European swaptions, because of their liquidity and importance. It will be shown how the prices of these instruments can be obtained using computational trees. The specific case of the BDT model is stressed in this section, but the general principles can be easily applied to different models.

The most elementary case to consider is that of the forward rate agreement, or FRA. Reviewing the set-up of a usual, or 'plain-vanilla' FRA of tenor τ:

- Payment occurs τ years after the reset date. That is, the delay between the rest of the index rate, and the corresponding payment, is equal to the maturity of the index rate.
- The payoff occurs at time $T = t_{reset} + \tau$.
- The payoff is given by:

  $$FRA = (R - X)\tau$$

  where X is the strike, and R is the index rate at expiry of the option.
- The time-0 forward rate F spanning the period from t_i to t_{i+1} is given by the formula

  $$F = (1 / \tau) [(P_i / P_{i+1}) - 1] \quad (9)$$

- And the net present value of the FRA is given by

  $$NPV(FRA) = \tau(F - X)P_{i+1}$$

  $$= \tau F P_{i+1} - \tau X P_{i+1}$$

  $$= \tau[(P_i/P_{i+1}) - 1](1/\tau)P_{i+1} - X\tau P_{i+1}$$

  $$= P_i - P_{i+1} - X\tau P_{i+1}$$

  $$NPV(FRA) = P_i - (1 + X\tau)P_{i+1} \quad (10)$$

  where P_k is a discounting zero maturing at time t_k, time t_i is the reset time, and time t_{i+1} is the time of payment.

Eq. (10) shows that the NPV of an FRA is just a linear combination of discount bonds. In the BDT model, discount bonds are priced exactly by construction. So, then it also follows, that FRA's are priced exactly by the BDT model.

And note that the same can be said for a swap: A swap will be priced exactly, since it is just a collection of FRA's. And, since the price of the FRA is independent of the choice of volatility, then so is the swap price.

(<u>Ken:</u> This statement was a little confusing to me at first, but now it seems right. It's a collection of FRA's, all taken together, where the formula for the swap rate is used to determine the strike rate X.)

Next important distinction: When it comes to pricing options on FRA's, it's important to distinguish between portfolios of options on FRA's – caps and floors – vs. an option on a portfolio of FRA's – swaptions.

First the author considers the case of pricing a portfolio of options, for instance a portfolio of caplets, which together make up a cap. For each individual caplet, the payoff is determined by the τ-maturity rate R, reset at time t. This payoff occurs at time $T = t + \tau$.:

$$Caplet(T) = \tau\, Max[(R - X), 0] \quad (11)$$

For pricing, you need to know the density of the index rate R at the reset time t, over all the possible different states of the world. So the problem arises of how to determine the value of R at node (i, j), that is, at time i, in state j. This notation for the value of R at this node is R_{ij}.

A formula for R_{ij} can be determined in a manner similar to that for finding the forward interest rate, as shown in Eq. (9) above. To do this, consider the discount bond $P(i, j, i + \tau/\Delta t)$, which denotes the time-i, node-j value of a discount bond which pays off \$1 at time-$(i + \tau/\Delta t)$. Using this discount bond, then the formula for R_{ij} comes out as

(12) $$R_{ij} = (1 / \tau) [(1 / P(i, j, i + \tau/\Delta t)) - 1]$$

which is exactly analogous to the formula for the forward interest rate.

Now, given the form for Eq. (12), it's important to spell out exactly how the value of the discount bond $P(i, j, i + \tau/\Delta t)$ is determined. To do this, simply express it using the risk-neutral expectation:

(13) $$P(i, j, i + \tau/\Delta t) = E'[\exp(-\int_{0\ldots\tau} r(s)\, ds) \mid r = r(i, j)]$$

which is just the usual expression for the value of a discount bond.

And, Eq. (13) can be approximated, using the already-constructed computational tree of short rates. As follows:

- Consider the portion of the tree which 'fans out' from the node-(i, j) until time-$(i + \tau/\Delta t)$.
- The short rate is known at each of these nodes.
- Then substitute in for a bond that pays \$1 at each of these 'final,' time-$(i + \tau/\Delta t)$ nodes,
- and carry out the averaging/discounting procedure back to node-(i, j).

This gives the value of the discount bond $P(i, j, i + \tau/\Delta t)$ there, at node (i, j).

Finally, once each of these rates R_{ij} has been solved for in this manner, then it's possible to find all the possible payoffs of the caplet, using the payoff formula given in Eq. (11). And once all the possible payoffs are known, then you may average and discount them back through the tree to the present time, to get the present value of the caplet.

Next, the author discusses using the computational tree to price swaptions. The value at expiration of the option is given by

(14) $$Swaption = Max[SR - X, 0]B$$

where:

SR is the swap rate at the time of option expiration.
X is the strike rate of the option.
B is an annuity that pays \$1 on each of the swap payment dates.

In a previous chapter (Chapter 1) it was shown that, for a plain-vanilla swap, the swap rate SR is given by the ratio of the floating leg to the fixed leg.

(<u>Ken:</u> I haven't seen the value of the swap rate SR expressed using this sentiment, that SR is equal to the ration of the floating to fixed. See below for elaboration, given the author's definitions, etc.)

Anyway, moving on, then if we independently value the floating leg and the fixed leg of the swap, which begins at time t, and matures at time T, then we get:

(15a) $$Floating\ Leg = P(t) - P(T)$$

(15b) $B = Fixed\ Leg = X \sum_{i=0\ldots n} P(t + i\tau)\tau$

Motivation of Eqs. (15a) and (15b) is as follows:

For Eq. (15a), note that at the time of expiry of the swaption, a floating rate note (FRN) will be worth \$1, where I am assuming that the FRN pays off a face value of \$1 at its maturity. (Also note that here, I am specifically concerned with an FRN whose payoff dates line up with those of the swap, although in general any FRN will be worth \$1 when it is issued.) So, given that expiry of the swaption occurs at time-t, then the present value of this FRN is $P(t)$.

Now, in valuing the floating leg of the swap, we are only concerned with the interest payments, and not with the final payment of principle. So, in order to properly value the Floating Leg, we subtract off the present value of this principle payment, or $P(T)$. Subtracting off this amount gives the expression in Eq. (15a).

For Eq. (15b), just note that the expression is given by the fixed payments, $X\tau$, discounted back to the present time. *(<u>Ken:</u> Although the subscripts on the summation should run from i = 1…n, since the author gives the convention that the maturity of the swap occurs at time-T = t + nτ.)*

(<u>Ken:</u> Further elaboration on the statement that the swap rate is the ration of the Floating to the Fixed leg. The statement doesn't seem right to me, as given. Specifically, it would seem correct to me if the strike rate X were absent from Eq. (15b). Because then, you could take SR = Float / Fixed, and derive the relation

$$SR \sum_{i=1\ldots n} P(t + i\tau)\tau = P(t) - P(T)$$

which would just say, that at the present time, the value of the fixed payments have to be equal to the value of the floating payments. This is consistent with the usual derivation of the swap rate.

Also, the presence of the X doesn't seem right to me in Eq. (15b) because, supposedly the Equations (15) are used to determine the swap rate, irregardless of the value of the swaption. Yet Eq. (15b) is calling into play the strike rate for the swaption.)

Moving on, with the relations discussed above, then we may write the value of the swaption as:

$$Swaption = Max[SR - X, 0]B$$

$$= Max[(Float/Fixed) - X, 0]B$$

$$= Max[(Float/B) - X, 0]B$$

(16) $$Swaption = Max[Float - XB, 0]$$

With this equation, then the swaption can be interpreted as an option to exchange a floating-rate bond for a fixed-coupon bond.

A further advantage to Eq. (16) is an efficient numerical algorithm for the valuation of the swaption.

From the preceding derivations, it can be seen that cap and swaption prices can be easily computed using BDT lattice trees. The values will depend on the input volatility of the short rate. Continuing on, it will be shown that:

- Cap prices can be recovered exactly, as can swaption prices, under specific circumstances.

- Due to the intrinsic 1-factor nature of models like BDT, in general it's impossible to simultaneously match both cap and swaption prices.

For pricing a cap: The price of a cap is the sum of the prices of the individual caplets. Assume a series of market prices for caps of varying maturities, but the same tenor. Here tenor refers to the length of one time period in an individual caplet, while the maturity refers to the length of the entire cap.

So assume a series of caps, for instance with maturities of length 6 months, 12 months, 18 months, etc. For a 1-factor model which exactly fits the interest-rate term structure, and which allows for a time-varying volatility, then we have the following procedure:

- Use a trial-and-error, or numerical root-finding procedure, to find the volatility implied by the first one of these caps. The vol found this way will be the vol of the underlying variable, which in the case of BDT is the short rate. With the vol taken to be constant at this vol over the first 6-months of the tree, then the first of the caps is exactly priced.
- Move on to the second cap. Further construct the tree, using the vol from the first step over the first 6 months of the tree, and then numerically search for the vol over the second 6 months of the tree that gives the correct price for this second cap. Now both the 6-month and the 1-year cap are exactly priced, by construction.
- Continue with this procedure for the whole series of caps, with each new cap giving the vol of the short rate over one more time period.

This procedure can be carried out for as many caps as there are in the series of market caps. Because each new step gives the short rate vol over one more time period, without interfering with any of the previously found vols.

This procedure is of general validity for any 1-factor model, but it's especially easy for BDT, by the following reasoning:

- Market caps are priced using the Black formula. *(<u>Ken:</u> My understanding of this statement is that market participants use the Black formula, with whatever parameters they see fit. They then quote the price produced from Black using those parameters.)*
- The Black formula, in turn, assumes lognormal forward rates.
- The BDT model displays 'almost' lognormal forward rates. *(<u>Ken:</u> The author doesn't explain this statement here, but rather defers to Chapters 16 and 20. My guess as to what he means has to do with the difference between the short rate and forward rates. Because, BDT assumes a lognormal short rate. My guess is that a lognormal short rate implies an 'almost' lognormal curve of forward rates, in whatever sense 'almost' is defined to be.)*
- By specifying a vol $\sigma(t)$ in BDT, then you're specifying the unconditional variance of the short-rate distribution to be $\sigma^2(t)t$.
- The thing is, at option expiry, the forward rate in the Black formula becomes a spot rate, so the BDT vol $\sigma(t)$ produces almost exactly the same variance as implied by the Black model price.
- Therefore, tree calibration to market prices of caps can be carried out almost by inspection.

(<u>Ken:</u> After reading this, it seems that there are going to have to be some corrections that will be employed in any practical implementation. First, there is this difference between the lognormal rates and the 'almost' lognormal rates. Also, question: is there some kind of convexity adjustment that must be employed here? Because, it seems (intuitively) like BDT uses the vol of the future value of the short rate, while Black uses the vol of the forward rate. In Hull, Ch. 16, Sec. 5, there's a convexity adjustment that gets employed sometimes between these two. It may or may not be needed here.)