ARBITRAGE OPPORTUNITIES IN
ARBITRAGE-FREE MODELS OF
BOND PRICING

David Backus
Silverio Foresi
Stanley Zin

ARBITRAGE OPPORTUNITIES IN
ARBITRAGE-FREE MODELS OF
BOND PRICING

ABSTRACT

Mathematical models of bond pricing are used by both academics and Wall Street practitioners, with practitioners introducing time-dependent parameters to fit "arbitrage-free" models to selected asset prices. We show, in a simple one-factor setting, that the ability of such models to reproduce a subset of security prices need not extend to state-contingent claims more generally. The popular Black-Derman-Toy model, for example, overprices call options on long bonds relative to those on short bonds when interest rates exhibit mean reversion. We argue, more generally, that the additional parameters of arbitrage-free models should be complemented by close attention to fundamentals, which might include mean reversion, multiple factors, stochastic volatility, and/or non-normal interest rate distributions.

David Backus
Stern School of Business
New York University
44 West 4th Street
New York, NY 10012-1126
and NBER

Silverio Foresi
Stern School of Business
New York University
44 West 4th Street
New York, NY 10012-1126

Stanley Zin
Graduate School of Industrial Administration
Carnegie Mellon University
Pittsburgh, PA 15213
and NBER

1 Introduction

Since Ho and Lee (1986) initiated work on "arbitrage-free" models of bond pricing, academics and practitioners have followed increasingly divergent paths. Both groups have the same objective: to extrapolate from prices of a limited range of assets the prices of a broader class of state-contingent claims. Academics study relatively parsimonious models, whose parameters are chosen to approximate "average" or "typical" behavior of interest rates and bond prices. Practitioners, on the other hand, use models with more extensive sets of time-dependent parameters, which they use to match current bond yields, and possibly other asset prices, exactly.

To practitioners, the logic of this choice is clear: the parsimonious models used by academics are inadequate for practical use. The four parameters of the Vasicek (1977) and Cox-Ingersoll-Ross (1985) models, for example, can be chosen to match five points on the yield curve (the four parameters plus the short rate), but do not reproduce the complete yield curve to the degree of accuracy required by market participants. Even complex, multi-factor models cannot generally approximate bond yields with sufficient accuracy. Instead, practitioners rely almost universally on models in the Ho and Lee (1986) tradition, including those developed by Black, Derman, and Toy (1990), Black and Karasinski (1991), Cooley, LeRoy, and Parke (1992), Heath, Jarrow, and Morton (1992), Hull and White (1990, 1993), and many others. Although analytical approaches vary across firms and even within them, the Black-Derman-Toy model is currently close to an industry standard.

Conversely, academics have sometimes expressed worry that the large parameter sets of arbitrage-free models may mask problems with their structure. A prominent example is Dybvig (1989), who noted that the changes in parameter values required by repeated use of this procedure contradicted the presumption of the theory that the parameters are deterministic functions of time. Black and Karasinski (1991, p 57) put it more colorfully: "When we value the option, we are assuming that its volatility is known and constant. But a minute later, we start using a new volatility. Similarly, we can value fixed income securities by assuming we know the one-factor short-rate process. A minute later, we start using a new process that is not consistent with the old one." Dybvig argued that these changes in parameter values through time implied that the framework itself was inappropriate.

We examine the practitioners' procedure in a relatively simple theoretical setting, a variant of Vasicek's (1977) one-factor Gaussian interest rate model that we refer to as the benchmark theory. Our thought experiment is to apply models with time-dependent drift and volatility parameters to asset prices generated by this theory. We

judge the models to be useful, in this setting, if they are able to reproduce prices of a broad range of state-contingent claims. This experiment cannot tell us how well the models do in practice, but it allows us to study the role of time-dependent parameters in an environment that can be characterized precisely. We find, in this environment, that if the world exhibits mean reversion, then the use of time-dependent parameters in a model without mean reversion can reproduce prices of a limited set of assets, but cannot reproduce the prices of general state-contingent claims. In this sense, these arbitrage-free models allow arbitrage opportunities: a trader basing prices on, say, the Black-Derman-Toy model can be exploited by a trader who knows the true structure of the economy.

A striking example of mispricing in this setting involves options on long bonds. Options of this type vary across two dimensions of time: the expiration date of the option and the maturity of the bond on which the option is written. The Black-Derman-Toy model has a one-dimensional array of volatility parameters that can be chosen to reproduce either the prices of options with any expiration date on one-period bonds or of options with common expiration dates on bonds of any maturity. But if the world exhibits mean reversion, this vector cannot reproduce the two-dimensional array of prices of bond options. If the volatility parameters are chosen to match prices of options on one-period bonds, then the model overprices options on long bonds.

While we think mean reversion is a useful feature in a model, it illustrates a more general point: that misrepresentations of fundamentals, whatever their form, cannot generally be overcome by adding arrays of time-dependent parameters. From this perspective, the role of academic research is to identify appropriate fundamentals, which might include multiple factors, stochastic volatility, and/or non-normal interest rate movements. We conjecture that arbitrage-free models with inaccuracies along any of these dimensions will misprice some assets as a result.

2 A Theoretical Benchmark

We use a one-factor Gaussian interest rate model as a laboratory in which to examine the practitioners' procedure of choosing time-dependent parameters to fit a bond pricing model to observed asset prices. The model is a close relative of Vasicek (1977). Although in some respects it is simpler than those used by practitioners, its log-linear structure is extremely useful in clarifying the roles played by its various parameters.

To fix the notation, let b_t^n be the *price* at date t of a zero-coupon bond of maturity n, the claim to one dollar at date $t + n$. By convention $b_t^0 = 1$ (one dollar today costs one dollar). Bond *yields* are

$$y_t^n = -n^{-1} \log b_t^n$$

and *forward rates* are

$$f_t^n = \log(b_t^n / b_t^{n+1}). \tag{1}$$

We label the *short rate* $r_t = y_t^1 = f_t^0$.

We characterize asset prices in our benchmark theory, or laboratory, with a *pricing kernel*: a stochastic process governing prices of state-contingent claims. Existence of such a process is guaranteed in any arbitrage-free environment. We describe the kernel for our theoretical environment in two steps. The first step involves an abstract state variable z, whose dynamics follow

$$
\begin{aligned}
z_{t+1} &= \varphi z_t + (1 - \varphi)\delta + \varepsilon_{t+1} \\
&= z_t + (1 - \varphi)(\delta - z_t) + \varepsilon_{t+1},
\end{aligned} \tag{2}
$$

with $\{\varepsilon_t\}$ distributed normally and independently with mean zero and variance σ^2. The parameter φ controls mean reversion: with $\varphi = 1$ the state follows a random walk, but with values between zero and one the conditional mean of future values of z converges to the unconditional mean δ. Step two is the pricing kernel m, which satisfies

$$-\log m_{t+1} = z_t + \lambda \varepsilon_{t+1}. \tag{3}$$

The parameter λ, which we refer to as the *price of risk*, determines the covariance between innovations to the kernel and the state and thus the risk characteristics of bonds and related assets.

Given a pricing kernel, we derive prices of assets from the pricing relation

$$1 = E_t(m_{t+1} R_{t+1}), \tag{4}$$

which holds for the gross return R_{t+1} on any traded asset. Since the one-period return on an $n + 1$-period bond is b_{t+1}^n / b_t^{n+1}, the pricing relation gives us

$$b_t^{n+1} = E_t(m_{t+1} b_{t+1}^n), \tag{5}$$

which allows us to compute bond prices recursively, starting with the initial condition $b_t^0 = 1$.

Consider a one-period bond. From equation (5) and the initial condition $b_t^0 = 1$ we see that the price is the conditional mean of the pricing kernel: $b_t^1 = E_t m_{t+1}$.

3

Since the kernel is conditionally lognormal, we need the following property of lognormal random variables: if $\log x$ is normal with mean μ and variance σ^2, then $\log E(x) = \mu + \sigma^2/2$. From equation (3) we see that $\log m_{t+1}$ has conditional mean $-z_t$ and conditional variance $(\lambda\sigma)^2$. Thus the one-period bond price satisfies

$$\log b_t^1 = -z_t + (\lambda\sigma)^2/2$$

and the short rate is

$$r_t = -\log b_t^1 = z_t - (\lambda\sigma)^2/2. \tag{6}$$

Thus the short rate r is the state z with a shift of origin. The mean short rate is $\delta - (\lambda\sigma)^2/2$, which we denote by μ in the rest of the paper.

The stochastic process for z, equation (2), implies similar behavior for the short rate:

$$r_{t+1} = r_t + (1 - \varphi)(\mu - r_t) + \varepsilon_{t+1}, \tag{7}$$

a discrete time version of Vasicek's (1977) short rate diffusion. Future values of the short rate are

$$r_{t+n} = r_t + (1 - \varphi^n)(\mu - r_t) + \sum_{j=1}^{n} \varphi^{n-j}\varepsilon_{t+j},$$

for $n \geq 1$, which yields conditional first and second moments of

$$E_t(r_{t+n}) = r_t + (1 - \varphi^n)(\mu - r_t) \tag{8}$$

and

$$Var_t(r_{t+n}) = \sigma^2 \sum_{j=1}^{n} \varphi^{2(n-j)} = \sigma^2 \left(\frac{1 - \varphi^{2n}}{1 - \varphi^2} \right). \tag{9}$$

We return to these formulas later.

Prices of long bonds follow from (5); details are provided in Appendix A.1. Their properties are conveniently summarized by forward rates, which are linear functions of the short rate:

$$f_t^n = r_t + (1 - \varphi^n)(\mu - r_t) + \left[\lambda^2 - \left(\lambda + \frac{1 - \varphi^n}{1 - \varphi} \right)^2 \right] \sigma^2/2, \tag{10}$$

for all $n \geq 0$. Given forward rates, we can compute bond prices and yields from their definitions. The right side of equation (10) has a relatively simple interpretation. If we compare it to (8), we see that the first two terms are the expected short rate n periods in the future. We refer to the last term as a *risk premium* and note that it

4

depends on three parameters: the magnitude of risk (σ), the price of risk (λ), and mean reversion (φ).

Both the Ho and Lee (1986) and Black, Derman, and Toy (1990) models are capable of reproducing an arbitrary forward rate curve (equivalently, yields or bond prices), including one generated by a theoretical model like this one. An issue we address later is whether this capability extends to more complex assets. With this in mind, consider a European call option at date t, with expiration date $t + \tau$ and strike price k, on a zero-coupon bond with maturity n at expiration. Given the lognormality of bond prices in this setting, the call price is given by the Black-Scholes (1973) formula,

$$c_t^{\tau,n} = b_t^{\tau+n} N(d_1) - k b_t^{\tau} N(d_2),\tag{11}$$

where N is the cumulative normal distribution function,

$$\begin{aligned} d_1 &= \frac{\log[b_t^{\tau+n}/(b_t^{\tau} k)] + v_{\tau,n}^2/2}{v_{\tau,n}} \\ d_2 &= d_1 - v_{\tau,n}, \end{aligned}$$

and the *option volatility* is

$$v_{\tau,n}^2 = Var_t(\log b_{t+\tau}^n) = \left(\frac{1 - \varphi^{2\tau}}{1 - \varphi^2}\right)\left(\frac{1 - \varphi^n}{1 - \varphi}\right)^2 \sigma^2.\tag{12}$$

See Appendix A.4. Jamshidian (1989) reports a similar formula for a continuous-time version of the Vasicek model. The primary difference from conventional applications of the Black-Scholes formula is the role of the mean reversion parameter φ in (12).

3   Parameter Values

We are trying, in this paper, to make a theoretical point, but we find it useful to illustrate the theory with numerical examples. The parameter values come from an informal moment matching exercise based on properties of monthly yields for US government securities computed by McCulloch and Kwon (1993). Some of the properties of these yields are reported in Table 1 for the sample period 1982-91.

We choose the parameters to approximate some of the salient features of bond yields using a time interval of one month. From equation (7) we see that μ is the unconditional mean of the short rate, so we set it equal to the sample mean of the

5

one-month yield in Table 1, 7.483/1200. (The 1200 converts an annual percentage rate to a monthly yield.) The mean reversion parameter φ is the first autocorrelation of the short rate. In Table 1 the autocorrelation is 0.906, so we set φ equal to this value. This indicates a high degree of persistence in the short rate, but less than with a random walk. The volatility parameter σ is the standard deviation of innovations to the short rate, which we estimate with the standard error of the linear regression (7). The result is $\sigma = 0.6164/1200$. Thus the values of (μ, σ, φ) are chosen to match the mean, standard deviation, and autocorrelation of the short rate. We choose the final parameter, the price of risk λ, to approximate the slope of the yield curve. Note from (10) that mean forward rates, in the theory, are

$$E\left(f_t^n\right) = \mu + \left[\lambda^2 - \left(\lambda + \frac{1 - \varphi^n}{1 - \varphi}\right)^2\right]\sigma^2/2.$$

This tells us that to produce an increasing mean forward rate curve (implying an increasing mean yield curve) we need λ to be negative. The price of risk parameter, in other words, governs the average slope of the yield curve. One way of fixing λ, then, is to select a value that makes the theoretical mean yield curve similar to the sample mean yield curve, given our chosen values for the other three parameters. An example is pictured in Figure 1, where we see that $\lambda = -750$ produces theoretical mean yields (the line in the figure) close to their sample means (the stars) for maturities between one month and ten years. With more negative values the mean yield curve is steeper, and with less negative (or positive) values the yield curve is flatter (or downward sloping).

Thus we see that all four parameters are required for the theory to imitate the dynamics of interest rates and the average slope of the yield curve. We use these benchmark values later to illustrate differences in prices across bond pricing models.

4   Ho and Lee Revisited

We turn now to the use of time-dependent parameters to fit theoretical models to observed asset prices. We apply, in turn, analogs of the models of Ho and Lee (1986) and Black, Derman, and Toy (1990) to a world governed by the benchmark theory of Section 2.

Our first example is a Gaussian analog of Ho and Lee's (1986) binomial interest rate model. The analog starts with a state equation,

$$z_{t+1} = z_t + \alpha_{t+1} + \eta_{t+1}, \tag{13}$$

6

with time-dependent parameters $\{\alpha_t\}$ and normally and independently distributed innovations $\{\eta_t\}$ with mean zero and constant variance β^2. The pricing kernel is

$$- \log m_{t+1} = z_t + \gamma \eta_{t+1}. \tag{14}$$

We use different letters for the parameters than in the benchmark theory to indicate that they may (but need not) take on different values. Our Ho and Lee analog differs from the benchmark theory in two respects. First, the short rate process does not exhibit mean reversion, which we might think of as setting $\varphi = 1$ in equation (2). Second, the state equation (13) includes time-dependent "drift" parameters $\{\alpha_t\}$.

Given equations (13,14), the pricing relation (5) implies a short rate

$$r_t = z_t - (\gamma\beta)^2/2$$

and forward rates

$$f_t^n = r_t + \sum_{j=1}^{n} \alpha_{t+j} + \left[\gamma^2 - (\gamma + n)^2\right] \beta^2/2 \tag{15}$$

for $n \geq 1$. See Appendix A.2. Note that (10) differs from (15) of the benchmark in two ways. One is the impact of the short rate on long forward rates. A unit increase in r is associated with increases in f^n of one in Ho and Lee, but $(1 - \varphi^n) < 1$ in the benchmark. The other is the risk premium, the final term in equation (15).

Despite these differences, time-dependent drift parameters allow the Ho and Lee model to reproduce some of the features of the benchmark theory. One such feature is the conditional mean of future short rates. The future short rates implied by this model are

$$r_{t+n} = r_t + \sum_{j=1}^{n} \left(\alpha_{t+j} + \eta_{t+j}\right), \tag{16}$$

which implies conditional means of

$$E_t(r_{t+n}) = r_t + \sum_{j=1}^{n} \alpha_{t+j} \tag{17}$$

for $n > 0$. If we compare this to the analogous expression for the benchmark theory, equation (8), we see that the two are equivalent if we set

$$\sum_{j=1}^{n} \alpha_{t+j} = (1 - \varphi^n)(\mu - r_t). \tag{18}$$

7

Thus the time-dependent drift parameters of the Ho and Lee model can be chosen to imitate this consequence of mean reversion in the benchmark theory.

In practice it is more common to use the drift parameters to fit the model to the current yield curve. To fit forward rates generated by the benchmark theory we need [compare (10,15)]

$$\sum_{j=1}^{n} \alpha_{t+j} = (1 - \varphi^n)(\mu - r_t) + \left[\lambda^2 - (\lambda + (1 - \varphi^n)/(1 - \varphi))^2 \right] \sigma^2/2$$
$$- \left[\gamma^2 - (\gamma + n)^2 \right] \beta^2/2. \tag{19}$$

The drift parameters implied by (18) and (19) are, in general, different. Since

$$\lim_{\varphi \to 1} \frac{1 - \varphi^n}{1 - \varphi} = n,$$

we can equate the two expressions when $\varphi = 1$ by setting $\beta = \sigma$ and $\gamma = \lambda$. But when $0 < \varphi < 1$ the two expressions cannot be reconciled. This is evident in Figure 2, where we graph the two choices of cumulative drift parameters, $\sum_{j=1}^{n} \alpha_{t+j}$, using the parameters estimated in Section 3, with $\beta = \sigma$, $\gamma = \lambda$, and $r = 3.0/1200$. The drift parameters that reproduce the conditional mean converge rapidly as the effects of mean reversion wear off. But the drift parameters that fit the current yield curve get steadily smaller as they offset the impact of maturity on the risk premium in this model. This results, for the range of maturities in the figure, in a declining term structure of expected future short rates, while the benchmark implies the reverse.

The discrepancy in the figure between the two choices of drift parameters is a concrete example of

Remark 1 *The parameters of the Ho and Lee model can be chosen to match the current yield curve, or the conditional means of future short rates, but they cannot generally do both.*

This property of the Ho and Lee model is a hint that time-dependent drift parameters do not adequately capture the effects of mean reversion in the benchmark theory. Dybvig (1989, p 5) makes a similar observation. A closer look indicates that the difficulty lies in the nonlinear interaction in the risk premium between mean reversion and the price of risk. In the benchmark theory, the risk premium on the n-period forward rate [see equation (10)] is

$$\left[\lambda^2 - \left(\lambda + \frac{1 - \varphi^n}{1 - \varphi} \right)^2 \right] \sigma^2/2.$$

8

In the Ho and Lee model, equation (15), the analogous expression is

$$\left[\gamma^2 - (\gamma + n)^2\right] \beta^2/2.$$

If we choose $\gamma = \lambda$ and $\beta = \sigma$ the two expressions are equal for $n = 1$, but they move apart as n grows. The discrepancy noted in Remark 1 is a direct consequence.

A similar comparison of conditional variances also shows signs of strain. The conditional variances of future short rates implied by the Ho and Lee model are

$$Var_t(r_{t+n}) = n\beta^2.$$

If we compare this to the analogous expression in the benchmark theory, equation (9), we see that they generally differ when $\varphi \neq 1$. With $\varphi < 1$ and $\beta = \sigma$, the conditional variances are the same for $n = 1$, but for longer time horizons they are greater in the Ho and Lee model. We summarize this discrepancy in

Remark 2 *The parameters of the Ho and Lee model cannot be chosen to reproduce the conditional variances of future short rates.*

Thus we see that additional drift parameters allow the Ho and Lee model to imitate some of the effects of mean reversion on bond yields. They cannot, however, reproduce the conditional variances of the benchmark theory. Dybvig (1989, p 5) summarized this feature of the Ho and Lee model more aggressively: "[T]he Ho and Lee model starts with an unreasonable implicit assumption about innovations in interest rates, but can obtain a sensible initial yield curve by making an unreasonable assumption about expected interest rates. Unfortunately, while this ... give[s] correct pricing of discount bonds..., there is every reason to believe that it will give incorrect pricing of interest rate options." Dybvig's intuition about options is easily verified. The Black-Scholes formula, equation (11), applies to the Ho and Lee model if we use option volatility

$$v_{\tau,n}^2 = Var_t(\log b_{t+\tau}^n) = \tau(n\beta)^2. \tag{20}$$

This formula cannot be reconciled with that of the benchmark theory, equation (12), for all combinations of τ and n unless $\varphi = 1$. If we choose the volatility parameter β to reproduce the option volatility $v_{1,1}$ of a short option on a short bond, then we overstate the volatilities of long options on long bonds. As a result, the model overvalues options with more distant expiration dates and/or longer underlying bonds.

We see in Figure 3 that call prices $c^{\tau,1}$ implied by the Ho and Lee model can be substantially higher than those generated by the benchmark theory. The figure

expresses the mispricing as a premium of the Ho and Lee price over the price generated by the benchmark theory. Benchmark prices are based on the parameter values of Section 3. Ho and Lee prices are based on drift parameters that match current bond prices, equation (19), and a volatility parameter $\beta = \sigma$ that matches the option volatility $v_{1,1}$ of a one-period option on a one-period bond. Both are evaluated at strike price $k = b_t^\tau / b_t^{\tau+n}$. For $\tau = 1$ the two models generate the same call price, but for options with expiration dates 12 months in the future the Ho and Lee price is more than 50 percent higher.

5 Black, Derman, and Toy Revisited

Black, Derman, and Toy (1990) extend the time-dependent parameters of Ho and Lee to a second dimension. They base bond pricing on a binomial process for the logarithm of the short rate in which both drift and volatility are time-dependent. We build an analog of their model that retains the linear, Gaussian structure of previous sections, but includes these two sets of time-dependent parameters. Given a two-parameter distribution like the normal, time-dependent drift and volatility can be used to match the conditional distribution of future short rates exactly and thus to mitigate the tendency of the Ho and Lee model to overprice long options. The question is whether they also allow us to reproduce the prices of other interest-rate derivative securities.

Our analog of the Black-Derman-Toy model adds time-dependent volatility to the structure of the previous section: a random variable z follows

$$z_{t+1} = z_t + \alpha_{t+1} + \eta_{t+1} \tag{21}$$

and the pricing kernel is

$$-\log m_{t+1} = z_t + \gamma \eta_{t+1}. \tag{22}$$

The new ingredient is that each η_t has time-dependent variance β_t^2. This structure differs from the benchmark theory in its absence of mean reversion (the coefficient of one on z_t in the state equation) and in its time-dependent drift (the α's) and volatility (the β's).

We approach bond pricing as before. We show in Appendix A.2 that the short rate is

$$r_t = -\log\left(E_t m_{t+1}\right) = z_t - (\gamma \beta_{t+1})^2 / 2$$

10

and forward rates are

$$f_t^n = r_t + \sum_{j=1}^{n} \alpha_{t+j} + \left[\gamma^2 - (\gamma + n)^2\right] \beta_{t+1}^2/2$$

$$+ \sum_{j=1}^{n} (\gamma + n - j)^2 (\beta_{t+j}^2 - \beta_{t+j+1}^2)/2 \qquad (23)$$

for $n \geq 1$. Equation (23) reduces to the Ho and Lee expression, equation (15), when $\beta_t = \beta$ for all t. As in the Ho and Lee model, forward rates differ from the benchmark in both the impact of short rate movements on long forward rates and the form of the risk premium.

We can use both sets of time-dependent parameters to approximate asset prices in the benchmark theory. Consider the conditional distribution of future short rates. The short rate follows

$$r_{t+1} = r_t + \gamma^2 (\beta_{t+1}^2 - \beta_{t+2}^2)/2 + \alpha_{t+1} + \eta_{t+1},$$

so future short rates are

$$r_{t+n} = r_t + \gamma^2 \left(\beta_{t+1}^2 - \beta_{t+n+1}^2\right)/2 + \sum_{j=1}^{n} (\alpha_{t+j} + \eta_{t+j}). \qquad (24)$$

Their conditional mean and variance are

$$E_t(r_{t+n}) = r_t + \gamma^2 \left(\beta_{t+1}^2 - \beta_{t+n+1}^2\right)/2 + \sum_{j=1}^{n} \alpha_{t+j} \qquad (25)$$

and

$$Var_t(r_{t+n}) = \sum_{j=1}^{n} \beta_{t+j}^2. \qquad (26)$$

This model, in contrast to Ho and Lee, is able to match the conditional variances of the benchmark theory, which we do by choosing volatility parameters that decline geometrically:

$$\beta_{t+j} = \varphi^{j-1}\sigma. \qquad (27)$$

This implies conditional variances of

$$Var_t(r_{t+n}) = \sum_{j=1}^{n} \varphi^{2(j-1)}\sigma^2 = \sigma^2 \left(\frac{1 - \varphi^{2n}}{1 - \varphi^2}\right),$$

the same as equation (9) of the theory. Similar patterns of declining time-dependent volatilities are common when these methods are used in practice, Black, Derman, and

11

Toy's numerical example included (see their Table I). To match the conditional mean of the benchmark theory, equation (8), we choose drift parameters that satisfy

$$\sum_{j=1}^{n} \alpha_{t+j} = (1 - \varphi^n)(\mu - r_t) - \gamma^2 \left(\beta_{t+1}^2 - \beta_{t+n+1}^2\right) / 2. \tag{28}$$

Thus we see, as Black, Derman, and Toy (1990, p 33) suggest, that we can fit the first two moments of the short rate with two "arrays" of parameters:

Remark 3 *The parameters of the Black-Derman-Toy model can be chosen to reproduce the conditional means and variances of future short rates.*

Given the critical role played by volatility in pricing derivative assets, this represents an essential advance beyond Ho and Lee.

Despite time-dependent volatility parameters, the model cannot simultaneously reproduce the conditional moments of the short rate and the forward rates of the benchmark theory. The drift parameters that reproduce the forward rate curve, equation (10), are

$$\sum_{j=1}^{n} \alpha_{t+j} = (1 - \varphi^n)(\mu - r_t) + \left[\lambda^2 - (\lambda + (1 - \varphi^n)/(1 - \varphi))^2\right] \sigma^2/2$$

$$- \left[\gamma^2 - (\gamma + n)^2\right] \beta_{t+1}^2/2 - \sum_{j=1}^{n} (\gamma + n - j)^2 (\beta_{t+j}^2 - \beta_{t+j+1}^2)/2. \tag{29}$$

Comparing (28) and (29) we see that the two are not equivalent, in general. We note the difference in

Remark 4 *Given volatility parameters (27) that reproduce the conditional variances of future rates, the drift parameters of the Black-Derman-Toy model can be chosen to match the current yield curve, or the conditional means of future short rates, but they cannot generally do both.*

Figure 4 plots the difference between (29) and (28). The parameter values are those of Section 3, with $\gamma = \lambda$, $\beta_{t+j} = \varphi^{j-1}\sigma$, and $r = 3.0/1200$. The differences are smaller than those of Figure 2 for the Ho and Lee model, but are nonzero nonetheless.

Remark 4 is a hint that the Black-Derman-Toy analog cannot generate accurate prices for the full range of state-contingent claims in the benchmark economy, but we

can see this more clearly by looking at specific assets. Consider European call options on zero-coupon bonds. Once more the lognormal structure of the model means that the Black-Scholes formula, equation (11), applies. If we choose drift parameters to fit the current yield curve, then any difference between call prices in the model and the benchmark theory must lie in their option volatilities. The option volatility for the Black-Derman-Toy analog is

$$v_{\tau,n}^2 = n^2 \sum_{j=1}^{\tau} \beta_{t+j}^2.$$

See Appendix A.4. If we restrict ourselves to options on one-period bonds (so that $n = 1$), we can reproduce the volatilities of our theoretical environment by choosing β's that decline geometrically with time, the same choice that replicates conditional variances of future short rates, equation (27).

However, the Black-Derman-Toy analog cannot simultaneously reproduce prices of options on bonds of longer maturities. If we use the geometrically declining parameters of equation (27), the option volatility is

$$v_{\tau,n}^2 = \left(\frac{1 - \varphi^{2\tau}}{1 - \varphi^2} \right) n^2 \sigma^2.$$

From (12) we see that the ratio of option volatilities is

$$\frac{\text{BDT Volatility}}{\text{Benchmark Volatility}} = \frac{n^2}{(1 + \varphi + \cdots + \varphi^{n-1})^2}.$$

This ratio is greater than one when $0 < \varphi < 1$ and $n > 1$, and implies that when the Black-Derman-Toy analog prices options on short bonds correctly, it will overprice options on long bonds. We see in Figure 5 that this mispricing gets worse the longer the maturity of the bond, and is greater than 150 percent for bonds with maturities of two years or more. As in Figure 3, benchmark prices are based on the parameters of Section 3, the drift parameters of the Black-Derman-Toy analog are chosen to reproduce the current yield curve, the strike price is $k = b_t^{\tau}/b_t^{\tau+n}$, and the expiration period is $\tau = 6$. Thus we have

Remark 5 *The parameters of the Black-Derman-Toy model cannot reproduce the prices of call options on bonds for all maturities and expiration dates.*

The difference in option volatilities, and hence in call prices, between the two models stems from two distinct roles played by mean reversion in determining prices

13

of long bonds. Mean reversion appears, first, in the impact of short rate innovations on future short rates. We see from (9) that the impact of innovations on future short rates, in the theory, decays geometrically with the time horizon τ. This feature is easily mimicked in the Black-Derman-Toy model, as we have seen, by using volatility parameters $\{\beta_t\}$ that decay at the same rate. The second role of mean reversion concerns the impact of short rate movements on long bond prices. In the Black-Derman-Toy model, a unit decrease in the short rate results in an n unit increase in the logarithm of the price of an n-period bond; see Appendix A.2. In the benchmark theory, the logarithm of the bond price rises by only $(1+\varphi+\cdots+\varphi^{n-1}) = (1-\varphi^n)/(1-\varphi)$; see Appendix A.1. This attenuation of the impact of short rate innovations on long bond prices is a direct consequence of mean reversion. It is not, however, reproduced by choosing geometrically declining volatility parameters and is therefore missing from call prices generated by the Black-Derman-Toy analog. Stated somewhat differently: the Black-Derman-Toy analog does not reproduce the hedge ratios of the benchmark theory.

As a practical matter, then, we might expect the Black-Derman-Toy procedure to work well in pricing options on short term instruments, including interest rate caps. For options on long bonds, however, the model overstates the option volatility and hence the call price. A common example of such an instrument is a callable bond. This procedure will generally overstate the value of the call provision to the issuer, in the theoretical environment, and thus understate the value of a callable long-term bond.

The discrepancy between option prices generated by the benchmark theory and the Black-Derman-Toy analog illustrates one difficulty of using models with time-dependent parameters: that a one-dimensional vector of time-dependent volatility parameters cannot reproduce the conditional variances of bond prices across the two dimensions of maturity and time-to-expiration. We turn now to a second example of mispricing: a class of "exotic" derivatives whose returns display different sensitivity than bonds to interest rate movements.

Consider an asset that delivers the power θ of the price of an n-period bond one period in the future. This asset has some of the flavor of derivatives with magnified sensitivity to interest rate movements made popular by Bankers Trust, yet retains the convenient log-linearity of bond prices in our framework. As with options, we compare the prices implied by the benchmark theory,

$$
\begin{aligned}
\log d_t^n = &-\left(r_t + (\lambda\sigma)^2/2\right) + \theta \log b_t^n + \left(\lambda + \theta\frac{1-\varphi^n}{1-\varphi}\right)^2 \sigma^2/2 \\
&-\theta(1-\varphi^n)(\mu - r_t),
\end{aligned} \tag{30}
$$

14

with those generated by the Black-Derman-Toy analog,

$$\log d_t^n = -(r_t + (\gamma\beta_{t+1})/2) + \theta \log b_t^n + (\gamma + \theta n)^2 \beta_{t+1}^2/2 - \theta(n\alpha_{t+1})$$

$$-\theta \sum_{j=1}^{n} \left[(n-j)(\alpha_{t+j+1} - \alpha_{t+j}) - (\gamma + n - j)^2 (\beta_{t+j+1}^2 - \beta_{t+j})^2/2 \right]. \quad (31)$$

Both follow from pricing relation (4) with the appropriate pricing kernel.

Suppose we choose the parameters of the Black-Derman-Toy analog to match the conditional variance of future short rates [equation (26)], the current yield curve [equation (29)], and the price of risk [$\gamma = \lambda$]. Can this model reproduce the prices of our exotic asset for all values of the sensitivity parameter θ? The answer is generally no. When $n = 1$ the two models generate the same price d_t^1 for all values of θ, just as we saw that the two sets of cumulative drift parameters were initially the same (see Figure 4 and the discussion following Remark 4). For longer maturities, however, the prices are generally different. Figure 6 is an example with $n = 60$ and $r = 3.00/1200$. We see that for values of θ outside the unit interval, the Black-Derman-Toy analog overprices the exotic, although the difference between models is smaller than for options on long bonds. Thus

Remark 6 *The parameters of the Black-Derman-Toy model can be chosen to match both the current yield curve and the term structure of volatility for options on one-period bonds, but they cannot generally replicate the prices of more exotic derivatives.*

These examples of mispricing illustrate a more general result: that the Black-Derman-Toy analog cannot replicate the state prices of the benchmark theory. This result is stated most clearly using *stochastic discount factors*,

$$M_{t,t+n} \equiv \prod_{j=1}^{n} m_{t+j},$$

which define n-period-ahead state prices. We show in the Appendix that the benchmark theory implies discount factors

$$-\log M_{t,t+n} = n\delta + \left(\frac{1 - \varphi^n}{1 - \varphi} \right)(r_t - \mu) + \sum_{j=1}^{n} \left(\lambda + \frac{1 - \varphi^{n-j}}{1 - \varphi} \right) \varepsilon_{t+j}$$

and the Black-Derman-Toy analog implies

$$-\log M_{t,t+n} = n\left[r_t + (\gamma\beta_{t+1})^2/2 \right] + \sum_{j=1}^{n}(n-j)\alpha_{t+j} + \sum_{j=1}^{n}(\gamma + n - j)\eta_{t+j}.$$

As long as $\varphi \neq 1$ and $\sigma \neq 0$, the second expression cannot be made equivalent to the first:

Proposition 1 *The time-dependent drift and volatility parameters of the Black-Derman-Toy model cannot be chosen to reproduce the stochastic discount factors of the benchmark theory.*

A proof is given in Appendix A.3. The difficulty lies in the innovations,

$$\left(\lambda + \frac{1 - \varphi^{n-j}}{1 - \varphi}\right) \varepsilon_{t+j} \qquad \text{and} \qquad (\gamma + n - j)\, \eta_{t+j}.$$

There is no choice of volatility parameters $\{\beta_t\}$ in Black-Derman-Toy that can replicate the nonlinear interaction of mean reversion (φ) and the price of risk (λ) in the benchmark theory.

In short, the time-dependent drift and volatility parameters of a model like the Black-Derman-Toy analog cannot replicate the prices of derivative assets generated by a model with mean reversion.

6   Arbitrage and Profit Opportunities

In the laboratory of the benchmark theory, the Black-Derman-Toy analog systematically misprices some assets. In this section we consider strategies that might be used to profit from traders using prices indicated by the Black-Derman-Toy model.

One way to exploit someone trading at Black-Derman-Toy prices is to arbitrage with someone trading at benchmark prices. If, as we have seen, a Black-Derman-Toy trader overprices options on long bonds relative to those on short rate, and a benchmark trader does not, then we can buy from the latter and sell to the former, thus making a riskfree profit. This is as clear an example of arbitrage as there is. But since such price differences are so obvious, they are unlikely to be very common.

It is less easy to exploit a Black-Derman-Toy trader when trades take place only at prices dictated by the model. Since Black-Derman-Toy prices are consistent with a pricing kernel, they preclude riskless arbitrage. We can often devise, however, dynamic strategies whose returns are large relative to their risk. One such strategy

involves call options on long bonds, which (again) the Black-Derman-Toy model generally overvalues. If we sell the option to the trader at date t, and liquidate at $t + 1$, we might expect to make a profit. This seems particularly likely in the option's final period, since the option is overvalued but the bonds on which the option is written are not.

Before we evaluate this strategy, we need to explain how the Black-Derman-Toy trader prices assets through time. A trader using the Black-Derman-Toy model generally finds, both in practice and in our theoretical setting, that the time-dependent parameters must be reset each period. In our setting, suppose the trader chooses volatility parameters at date t to fit implied volatilities from options on short bonds. As we have seen, this leads her to set $\{\beta_{t+1}, \beta_{t+2}, \beta_{t+3}, \ldots\}$ equal to $\{\sigma, \varphi\sigma, \varphi^2\sigma, \ldots\}$. If the parameters were literally time-dependent, then in the following period logic dictates that we set $\{\beta_{t+2}, \beta_{t+3}, \beta_{t+4}, \ldots\}$ equal to $\{\varphi\sigma, \varphi^2\sigma, \varphi^3\sigma, \ldots\}$. But if we calibrate once more to options on one-period bonds, we would instead use $\{\sigma, \varphi\sigma, \varphi^2\sigma, \ldots\}$. This leads to a predictable upward jump in the volatility parameters from one period to the next, as in Figure 7. Similarly, the drift parameters must be adjusted each period to retain the model's ability to reproduce the current yield curve. These changes in parameter values through time are an indication that the model is imperfectly imitating the process generating state prices, but they nevertheless improve its performance. In this sense the internal inconsistency noted by Dybvig (1989) is a symptom of underlying problems, but not a problem in its own right. Thus our trader prices call options using drift and volatility parameters that match, each period, the yield curve and the term structure of volatilities implied by options on one-period bonds, equations (29) and (27), respectively.

Now consider a strategy against such a trader of selling a τ-period option on an n-period bond and buying it back one period later. The gross one-period return from this strategy is

$$R_{t+1} = -c_{t+1}^{\tau-1,n}/c_t^{\tau,n},$$

for $\tau \geq 1$, with the convention that an option at expiration has value $c_t^{0,n} = (b_t^n - k)^+$. The profit from this strategy is not riskfree, but it can be large relative to the risk involved. In the world of the benchmark theory, the appropriate adjustment for risk is given by the pricing relation (4). We measure the mean excess risk-adjusted return by

$$a = E(m_{t+1}R_{t+1}) - 1,$$

using the pricing kernel m for the benchmark economy. This measure is analogous to Jensen's alpha for the CAPM and is the appropriate risk adjustment in this setting. For the trading strategy just described, and the parameter values of Section 3, the returns are in the neighborhood of several hundred percent per year; see Figure 8.

The returns in the figure were computed by Monte Carlo, and concern two-period (τ=2) options on bonds for maturities (n) between 1 and 24 months. The returns are greater with $\tau = 1$, but decline significantly as τ increases. For $\tau > 6$ the overpricing is difficult to exploit, since the overpricing of a five-period option the following period is almost as great.

7   Mean Reversion and Other Fundamentals

We have seen that the time-dependent drift and volatility parameters of analogs of the Ho and Lee (1986) and Black-Derman-Toy (1990) models cannot reproduce the prices of state-contingent claims generated by a model with mean reversion. This thought experiment does not tell us how well these models perform in practice, but it indicates that extra parameters are not a solution to all problems: they need to be used in the context of a structure with sound fundamentals.

We focused on mean reversion because we find it an appealing feature in a bond pricing model, yet it is missing from the most popular binomial interest rate models. The estimated autocorrelation of 0.906 reported in Table 1 is not substantially different from one, but a value of one has, in the benchmark theory, two apparently counterfactual implications for bond yields. The first is that the mean yield curve eventually declines (to minus infinity) with maturity. In fact yield curves are, on average, upward sloping, Figure 1 being a typical example. A second implication was pointed out by Gibbons and Ramaswamy (1993) and may be more telling: with φ close to one, average yield curves exhibit substantially less curvature than we see in the data. An example with $\varphi = 0.99$ is pictured in Figure 9, where we have set $\lambda = -200$ to keep the theoretical 10-year yield close to its sample mean. These two implications illustrate the added power of combining time-series and cross-section information, and suggest to us that random walk models overstate the persistence of the short-term rate of interest. For these reasons, we feel that mean reversion is suggested by the properties of bond prices.

For similar reasons, Black and Karasinski (1991), Heath, Jarrow, and Morton (1992), and Hull and White (1990, 1993) have developed ways of introducing mean-reversion into arbitrage-free models. Despite these developments, mean reversion is generally ignored by all but the most sophisticated market participants. Financial professionals do not generally reveal their methods, but they can be inferred from other sources. One source is books on fixed income modeling. Tuckman (1995) is aimed at quantitatively-oriented MBA students and is used in Salomon Brothers'

18

sales and trading training course, yet devotes only four pages to models with mean reversion. Sundaresan (1996) targets a similar audience and disregards mean reversion altogether. Perhaps a better source is Bloomberg, the leading purveyor of fixed income information services. Calculations on Bloomberg terminals of the "fair value" of interest rate derivatives are based on a log-normal analog of the Ho and Lee model, and calculations of "option-adjusted spreads" for callable bonds use the same model as the default (Bloomberg 1996), with Black, Derman, and Toy (1990) and Hull and White (1990) available as alternatives.

However, we think the central issue is not mean reversion, but whether pricing models — arbitrage-free or not — provide a good approximation to the fundamentals driving bond prices. Mean reversion, in this context, is simply an example of a fundamental that cannot be mimicked by time-dependent drift and volatility parameters. A second example is multiple factors. Work by, among others, Brennan and Schwartz (1979), Chen and Scott (1993), Duffie and Singleton (1995), Dybvig (1989), Garbade (1986), Litterman and Scheinkman (1991), and Stambaugh (1988) clearly indicates the benefits of additional factors in accounting for the evolution of bond prices through time. Longstaff and Schwartz (1992) are representative of a growing number of studies that identify a second factor with volatility, whose variation through time is apparent in the prices of options on fixed income instruments. Yet another example is the fat tails in weekly interest rate changes documented by Das (1994), who introduced jumps into the standard diffusion-based models. We would guess that models that ignore any or all of these fundamentals will, as a direct consequence, produce inaccurate state prices.

We argue, in short, that fundamentals matter, and that poor choice of fundamentals cannot generally be rectified by judicious use of time-dependent parameters.

8 Final Thoughts

We have examined the practitioners' methodology of choosing time-dependent parameters to fit an arbitrary bond pricing model to current asset prices. We showed, in a relatively simple theoretical setting, that this method can systematically misprice some assets. Like Ptolemy's geocentric model of the solar system, these models can generally be "tuned" to provide good approximations to (in our case) prices of a limited set of assets, but they may also provide extremely poor approximations for other assets. As Lochoff (1993, p 92) puts it: "Even bad models can be tuned to give good results" for simple assets.

Whether these theoretical examples of mispricing can be used to direct trading strategies in realistic settings depends on the relative magnitudes of the mispricing and the inevitable approximation errors of more parsimonious models. We would not be willing, at present, to bet our salaries on our benchmark theory. Nevertheless, we think the exercise indicates that it is important to get the fundamentals right. In our thought experiment, fundamentals were represented by the degree of mean reversion in the short rate, and we saw that a mistake in this dimension could lead to large pricing errors on some securities. The extra time-dependent parameters, in other words, are not a panacea: they allow us to reproduce a subset of asset prices, but do not guarantee accurate prices for the full range of interest-rate derivative securities. In more general settings, we expect a model with n arrays of parameters to be able to reproduce the term structure of prices of n classes of assets, but if the fundamentals are wrong there will be some assets that are mispriced.

The difficulty in practice is that even the best fundamental models provide only a rough approximation to the market prices of fixed income derivatives. That leaves us in the uncertain territory described by Black and Karasinski (1991, p 57): "[One] approach is to search for an interest rate process general enough that we can assume it is true and unchanging. ... While we may reach this goal, we don't know enough to use this approach today." Best practice, we think, is to combine current knowledge of fundamentals with enough extra parameters to make the approximation adequate for trading purposes.

A   Mathematical Appendix

We derive many of the formulas used in the text. For both the benchmark theory and our analog of the Black-Derman-Toy model, we derive the stochastic discount factors implied by the pricing kernels listed in the text, and the implied bond prices, forward rates, and prices of call options on discount bonds. The result is effectively a mathematical summary of the paper.

A.1   Benchmark Theory

We characterize bond pricing theories with stochastic discount factors,

$$
M_{t,t+n} \equiv \prod_{j=1}^{n} m_{t+j},
$$

or $-\log M_{t,t+n} \equiv -\sum_{j=1}^{n} \log m_{t+j}$. Given the pricing kernel m of our benchmark theory [equations (2) and (3)], the stochastic discount factors are

$$
-\log M_{t,t+n} \;=\; n\delta + \left(\frac{1-\varphi^{n}}{1-\varphi}\right)(z_t - \delta) + \sum_{j=1}^{n}\left(\lambda + \frac{1-\varphi^{n-j}}{1-\varphi}\right)\varepsilon_{t+j}
$$

for $n \geq 1$. Given the discount factors, we compute bond prices from $b_t^n = E_t M_{t,t+n}$ [a consequence of (5)]:

$$
-\log b_t^n \;=\; n\delta + \left(\frac{1-\varphi^{n}}{1-\varphi}\right)(z_t - \delta) - \sum_{j=1}^{n}\left(\lambda + \frac{1-\varphi^{n-j}}{1-\varphi}\right)^2 \sigma^2/2.
$$

Forward rates [see (1)] are

$$
\begin{aligned}
f_t^n &= \log b_t^n - \log b_t^{n+1} \\
&= \delta + \varphi^n(\delta - z_t) - \left(\lambda + \frac{1-\varphi^n}{1-\varphi}\right)^2 \sigma^2/2,
\end{aligned}
$$

which includes a short rate of $r_t = f_t^0 = z_t - (\lambda\sigma)^2/2$.

It is conventional to express discount factors, bond prices, and forward rates in terms of the observable short rate r rather than the abstract state variable z, which

is easily done given the linear relation between them. Since $\mu = \delta - (\lambda\sigma)^2/2$, we get discount factors

$$-\log M_{t,t+n} = n\delta + \left(\frac{1-\varphi^n}{1-\varphi}\right)(r_t - \mu) + \sum_{j=1}^{n}\left(\lambda + \frac{1-\varphi^{n-j}}{1-\varphi}\right)\varepsilon_{t+j}, \qquad \text{(A1)}$$

bond prices

$$-\log b_t^n = n\delta + \left(\frac{1-\varphi^n}{1-\varphi}\right)(r_t - \mu) - \sum_{j=1}^{n}\left(\lambda + \frac{1-\varphi^{n-j}}{1-\varphi}\right)^2 \sigma^2/2,$$

and forward rates

$$f_t^n = r_t + (1 - \varphi^n)(\mu - r_t) + \left[\lambda^2 - \left(\lambda + \frac{1-\varphi^n}{1-\varphi}\right)^2\right]\sigma^2/2,$$

as stated in equation (10).

A.2   Black, Derman, and Toy

We approach our Gaussian analog of the Black-Derman-Toy model the same way, using equations (21) and (22) to define the pricing kernel. Our analog of the Ho and Lee model is a special case with $\beta_t = \beta$ for all t. The stochastic discount factors are

$$-\log M_{t,t+n} = nz_t + \sum_{j=1}^{n}(n-j)\alpha_{t+j} + \sum_{j=1}^{n}(\gamma + n - j)\eta_{t+j}.$$

They imply bond prices of

$$-\log b_t^n = nz_t + \sum_{j=1}^{n}(n-j)\alpha_{t+j} - \sum_{j=1}^{n}(\gamma + n - j)^2\beta_{t+j}^2/2,$$

and, for $n \geq 1$, forward rates of

$$f_t^n = z_t + \sum_{j=1}^{n}\alpha_{t+j} - (\gamma + n)^2\beta_{t+1}^2/2 - \sum_{j=1}^{n}(\gamma + n - j)^2(\beta_{t+j}^2 - \beta_{t+j+1}^2)/2.$$

Since the short rate is $r_t = f_t^0 = z_t - (\gamma\beta_{t+1})^2/2$, we can rewrite these relations as

$$-\log M_{t,t+n} = n\left[r_t + (\gamma\beta_{t+1})^2/2\right] + \sum_{j=1}^{n}(n-j)\alpha_{t+j} + \sum_{j=1}^{n}(\gamma + n - j)\eta_{t+j}, \qquad \text{(A2)}$$

$$-\log b_t^n = n\left[r_t + (\gamma\beta_{t+1})^2/2\right] + \sum_{j=1}^{n}(n-j)\alpha_{t+j} - \sum_{j=1}^{n}(\gamma+n-j)^2\beta_{t+j}^2/2,$$

and

$$f_t^n = r_t + \sum_{j=1}^{n}\alpha_{t+j} + \left[\gamma^2 - (\gamma+n)^2\right]\beta_{t+1}^2/2$$
$$+ \sum_{j=1}^{n}(\gamma+n-j)^2(\beta_{t+j}^2 - \beta_{t+j+1}^2)/2,$$

as stated in equation (23). With $\beta_{t+j} = \beta$ this reduces to

$$f_t^n = r_t + \sum_{j=1}^{n}\alpha_{t+j} + \left[\gamma^2 - (\gamma+n)^2\right]\beta^2/2,$$

equation (15) of the Ho and Lee model.

A.3   Nonequivalence of Discount Factors

Our examples of assets that are mispriced by the Black-Derman-Toy model indicate that Black-Derman-Toy cannot generally reproduce the stochastic discount factors of our benchmark theory:

Proposition 1 *The parameters $\{\gamma, \alpha_t, \beta_t\}$ of the Black-Derman-Toy model cannot be chosen to reproduce the stochastic discount factors of the benchmark theory.*

The proof consists of comparing the two discount factors, (A1) for the benchmark theory and (A2) for the Black-Derman-Toy model. The deterministic terms of (A1) and (A2) are relatively simple. For, say, the first n discount factors, we can equate the conditional means of (A1) and (A2) by judicious choice of the n drift parameters $\{\alpha_{t+1}, \ldots, \alpha_{t+n}\}$. The stochastic terms, however, cannot generally be equated. We can represent the initial stochastic terms for the benchmark theory in an array like this:

	$j = 1$	$j = 2$	$j = 3$
$n = 1$	$\lambda\varepsilon_{t+1}$		
$n = 2$	$(\lambda+1)\varepsilon_{t+1}$	$\lambda\varepsilon_{t+2}$	
$n = 3$	$(\lambda+1+\varphi)\varepsilon_{t+1}$	$(\lambda+1)\varepsilon_{t+2}$	$\lambda\varepsilon_{t+3}$

For the Black-Derman-Toy model the analogous terms are

	$j = 1$	$j = 2$	$j = 3$
$n = 1$	$\gamma\eta_{t+1}$		
$n = 2$	$(\gamma + 1)\eta_{t+1}$	$\gamma\eta_{t+2}$	
$n = 3$	$(\gamma + 2)\eta_{t+1}$	$(\gamma + 1)\eta_{t+2}$	$\gamma\eta_{t+3}$

If $\sigma = 0$ the model is riskless and replicable with drift parameters alone. Alternatively, if $\varphi = 1$ we can replicate the benchmark theory by setting $\gamma = \lambda$ and $\eta_{t+j} = \varepsilon_{t+j}$, which implies $\beta_{t+j} = \sigma$ for all j. But with $\sigma \neq 0$ and $\varphi \neq 1$ it is impossible to choose the price of risk γ and volatility parameters $\{\beta_{t+1}, \beta_{t+2}, \beta_{t+3}\}$ to reproduce the benchmark theory. Suppose we try to match the terms sequentially. To match the term $(n, j) = (1, 1)$ we need

$$\gamma\eta_{t+1} = \lambda\varepsilon_{t+1},$$

which requires $\gamma\beta_{t+1} = \lambda\sigma$. Similarly, equivalence of the $(2, 1)$ terms,

$$(\gamma + 1)\eta_{t+1} = (\lambda + 1)\varepsilon_{t+1},$$

tells us (for nonzero λ) that the parameters must be $\gamma = \lambda$ and $\beta_{t+1} = \sigma$. The $(3, 1)$ term now requires

$$\gamma + 2 = \lambda + 1 + \varphi,$$

which is inconsistent with our earlier parameter choices when $\varphi \neq 1$. (When $\lambda = 0$ the same conclusion holds, but the argument starts with the $(2, 1)$ term.) Thus we see that our attempt to reproduce the discount factors of the benchmark theory with those of the Black-Derman-Toy model has failed.

It is easy to imagine similar problems with extensions of the model. We could, for example, let γ depend on time. But for similar reasons, this one-dimensional array cannot reproduce the discount factors of the benchmark model.

A.4 Bond Options

Any stream of cash flows $\{h_t\}$ can be valued by

$$E_t \sum_{j=1}^{n} M_{t,t+j} h_{t+j}$$

using the stochastic discount factors M. We use this relation to price European options on zero-coupon bonds.

Consider a European call option at date t with expiration date $t + \tau$. The option gives its owner the right to buy an n-period bond at date $t + \tau$ for the exercise price k, thus generating the cash flow

$$h_{t+\tau} = (b^n_{t+\tau} - k)^+,$$

where $x^+ \equiv \max\{0, x\}$ is the nonnegative part of x. The call price is

$$c^{\tau,n}_t = E_t \left[M_{t,t+\tau}(b^n_{t+\tau} - k)^+ \right]. \tag{A3}$$

Computing this price involves evaluating (A3) with the appropriate discount factor and bond price.

Both of our theories have lognormal discount factors and bond prices, so to evaluate (A3) we need two properties of lognormal expectations. Let us say that $\log x = (\log x_1, \log x_2)$ is bivariate normal with mean vector μ and variance matrix Σ. Formula 1 is

$$E\left[x_1 I(x_2 - k)\right] = \exp(\mu_1 + \sigma^2_1/2)N(d),$$

with

$$d = \frac{\mu_2 - \log k + \sigma_{12}}{\sigma_2},$$

where N is the standard normal distribution function and I is an indicator function that equals one if its argument is positive, zero otherwise. A similar result is stated and proved by Rubinstein (1976, Appendix). Except for the N term, this is the usual expression for the mean of a lognormal random variable. Formula 2 follows from 1 with a change of variables ($x_1 x_2$ for x_1):

$$E\left[(x_1 x_2)I(x_2 - k)\right] = \exp(\mu_1 + \mu_2 + \sigma^2_1/2 + \sigma^2_2/2 + \sigma_{12})N(d),$$

with

$$d = \frac{\mu_2 - \log k + \sigma_{12} + \sigma^2_2}{\sigma_2}.$$

Our application of these formulas to (A3) uses $M_{t,t+\tau}$ as x_1 and $b^n_{t+\tau}$ as x_2.

For the benchmark theory, we use discount factor (A1) and evaluate (A3) using the expectation formulas. To keep the notation manageable, let

$$A_n \equiv \sum_{j=1}^{n} \left(\lambda + \frac{1 - \varphi^{n-j}}{1 - \varphi} \right)^2 \sigma^2/2.$$

25

Now we look at the discount factor and future bond price. The discount factor is, from (A1),

$$
-\log M_{t,t+\tau} \;=\; \tau\delta + \left(\frac{1-\varphi^{\tau}}{1-\varphi}\right)(r_t - \mu) + \sum_{j=1}^{\tau}\left(\lambda + \frac{1-\varphi^{\tau-j}}{1-\varphi}\right)\varepsilon_{t+j}.
$$

It has conditional mean given by the first two terms and conditional variance $2A_{\tau}$. The future bond price can be represented

$$
\begin{aligned}
-\log b_{t+\tau}^{n} \;&=\; n\delta - A_n + \left(\frac{1-\varphi^{n}}{1-\varphi}\right)(r_{t+\tau} - \mu)\\[2mm]
&=\; n\delta - A_n + \left(\frac{1-\varphi^{n}}{1-\varphi}\right)\varphi^{\tau}(r_t - \mu) + \left(\frac{1-\varphi^{n}}{1-\varphi}\right)\sum_{j=1}^{\tau}\varphi^{\tau-j}\varepsilon_{t+j}\\[2mm]
&=\; -\log b_t^{\tau+n} + \log b_t^{\tau} + (A_{n+\tau} - A_{\tau} - A_n) + \left(\frac{1-\varphi^{n}}{1-\varphi}\right)\sum_{j=1}^{\tau}\varphi^{\tau-j}\varepsilon_{t+j}.
\end{aligned}
$$

An enormous amount of algebra gives us the call price

$$
c_t^{\tau,n} = b_t^{\tau+n}N(d_1) - kb_t^{\tau}N(d_2),
$$

with

$$
\begin{aligned}
d_1 \;&=\; \frac{\log[b_t^{\tau+n}/(b_t^{\tau}k)] + v_{\tau,n}^2/2}{v_{\tau,n}}\\[2mm]
d_2 \;&=\; d_1 - v_{\tau,n}
\end{aligned}
$$

and option volatility

$$
v_{\tau,n}^2 = \sigma^2\left(\frac{1-\varphi^{n}}{1-\varphi}\right)^2\sum_{j=1}^{\tau}\varphi^{2(\tau-j)} = \sigma^2\left(\frac{1-\varphi^{n}}{1-\varphi}\right)^2\left(\frac{1-\varphi^{2\tau}}{1-\varphi^2}\right),
$$

the conditional variance of the logarithm of the future bond price.

The call price for the Black-Derman-Toy model follows a similar route with discount factor (A2). If we choose the model's parameters to match bond prices, then the only difference in the call formula is the option volatility,

$$
v_{\tau,n}^2 = n^2\sum_{j=1}^{\tau}\beta_{t+j}^2.
$$

Whether this is the same as the benchmark theory depends on the choice of volatility parameters $\{\beta_t\}$.

The difficulty with the Black-Derman-Toy model with respect to pricing options in our benchmark economy is similar to that with stochastic discount factors (Proposition 1). Volatilities are defined over the two-dimensional array indexed by the length τ of the option and the maturity n of the bond on which the option is written. This array cannot be replicated by the one-dimensional vector of volatility parameters $\{\beta_t\}$.

References

Black, Fischer, Emanuel Derman, and William Toy, 1990, "A one-factor model of interest rates and its application to treasury bond options," *Financial Analysts Journal* 46, 33-39.

Black, Fischer, and Piotr Karasinski, 1991, "Bond and option pricing when short rates are lognormal," it Financial Analysts Journal 47, 52-59.

Black, Fischer, and Myron Scholes, 1973, "The pricing of options and corporate liabilities," *Journal of Political Economy* 81, 637-654.

Bloomberg, 1996, "Help for OSY," on-line help facility, May.

Brennan, Michael, and Eduardo Schwartz, 1979, "A continuous time approach to the pricing of bonds," *Journal of Banking and Finance* 3, 133-155.

Chen, Ren-Raw, and Louis Scott, 1993, "Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates," *Journal of Fixed Income* 3, 14-31.

Cooley, Thomas, Stephen LeRoy, and William Parke, 1992, "Pricing interest-sensitive claims when interest rates have stationary components," *Journal of Fixed Income* 2, 64-73.

Cox, John, Jonathan Ingersoll, and Stephen Ross, 1985, "A theory of the term structure of interest rates," *Econometrica* 53, 385-407.

Das, Sanjiv, 1994, "Discrete time bond and option pricing for jump-diffusion processes," manuscript, New York University, April.

Duffie, Darrell, and Kenneth Singleton, 1995, "An econometric model of the term structure of interest rate swap yields," manuscript, Stanford University, June.

Dybvig, Philip, 1989, "Bond and bond option pricing based on the current term structure," manuscript, Washington University, February; forthcoming in M. Dempster and S. Pliska, eds., *Mathematics of Derivative Securities*, Cambridge: Cambridge University Press.

Garbade, Kenneth, 1986, "Modes of fluctuation in bond yields," Bankers Trust, Topics in Money and Securities Markets No 20, June.

Gibbons, Michael, and Krishna Ramaswamy, 1993, "A test of the Cox, Ingersoll, Ross model of the term structure," *Review of Financial Studies* 6, 619-658.

Heath, David, Robert Jarrow, and Andrew Morton, 1992, "Bond pricing and the term structure of interest rates," *Econometrica* 60, 225-262.

Ho, Thomas S.Y., and Sang-Bin Lee, 1986, "Term structure movements and pricing interest rate contingent claims," *Journal of Finance* 41, 1011-1029.

Hull, John, and Alan White, 1990, "Pricing interest-rate-derivative securities," *Review of Financial Studies* 3, 573-592.

Hull, John, and Alan White, 1993, "One-factor interest-rate models and the valuation of interest-rate derivative securities," *Journal of Financial and Quantitative Analysis* 28, 235-254.

Jamshidian, Farshid, 1989, "An exact bond option formula," *Journal of Finance* 44, 205-209.

Litterman, Robert, and José Scheinkman, 1991, "Common factors affecting bond returns," *Journal of Fixed Income* 1, 54-61.

Lochoff, Roland, 1993, "The contingent-claims arms race," *Journal of Portfolio Management* 20, 88-92.

Longstaff, Francis, and Eduardo Schwartz, 1992, "Interest rate volatility and the term structure: A two-factor general equilibrium model," *Journal of Finance* 47, 1259-1282.

McCulloch, J. Huston, and Heon-Chul Kwon, 1993, "US term structure data," manuscript and computer diskettes, Ohio State University, March.

Rubinstein, Mark, 1976, "The valuation of uncertain income streams and the pricing of options," *Bell Journal of Economics* 7, 407-425.

Stambaugh, Robert, 1988, "The information in forward rates: Implications for models of the term structure," *Journal of Financial Economics* 21, 41-70.

Sundaresan, Suresh, 1996, *Fixed Income Markets and Their Derivatives*, Cincinnati: South-Western Publishing.

Tuckman, Bruce, 1995, *Fixed Income Securities*, New York: Wiley and Sons.

Vasicek, Oldrich, 1977, "An equilibrium characterization of the term structure," *Journal of Financial Economics* 5, 177-188.

Table 1

Properties of Government Bond Yields

Data are monthly estimates of annualized continuously-compounded zero-coupon US government bond yields computed by McCulloch and Kwon (1993). The sample period is January 1982 to February 1992.

Maturity	Mean	Standard Deviation	Autocorrelation
1 month	7.483	1.828	0.906
3 months	7.915	1.797	0.920
6 months	8.190	1.894	0.926
9 months	8.372	1.918	0.928
12 months	8.563	1.958	0.932
24 months	9.012	1.986	0.940
36 months	9.253	1.990	0.943
48 months	9.405	1.983	0.946
60 months	9.524	1.979	0.948
84 months	9.716	1.956	0.952
120 months	9.802	1.864	0.950

Figure 1. Mean Yields in Theory and Data

Stars are data points, the line is theory
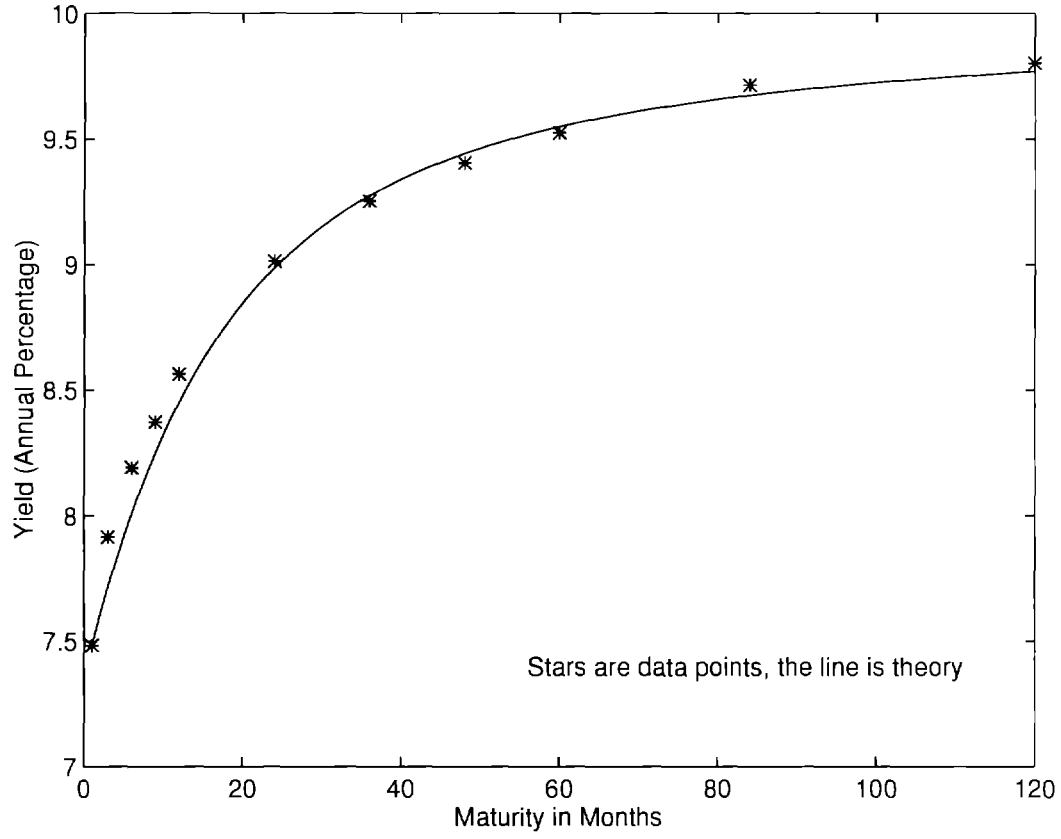
Figure 2. Two Choices of Ho–Lee Drift Parameters
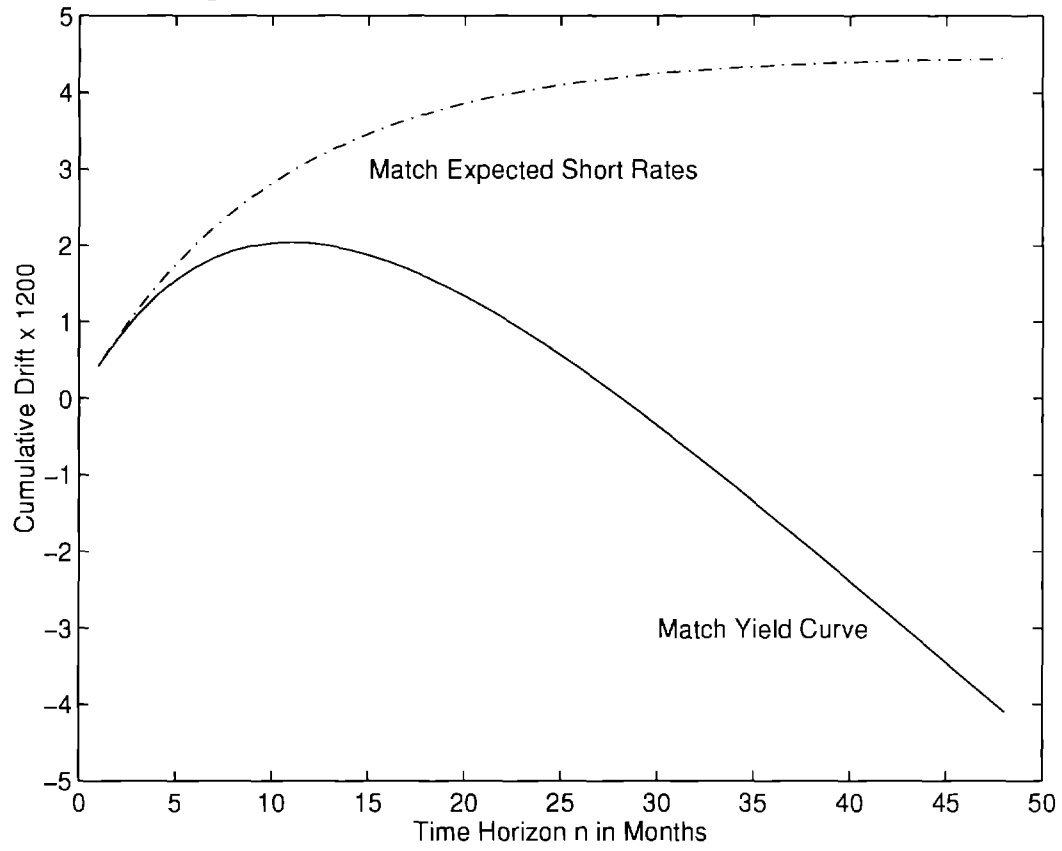
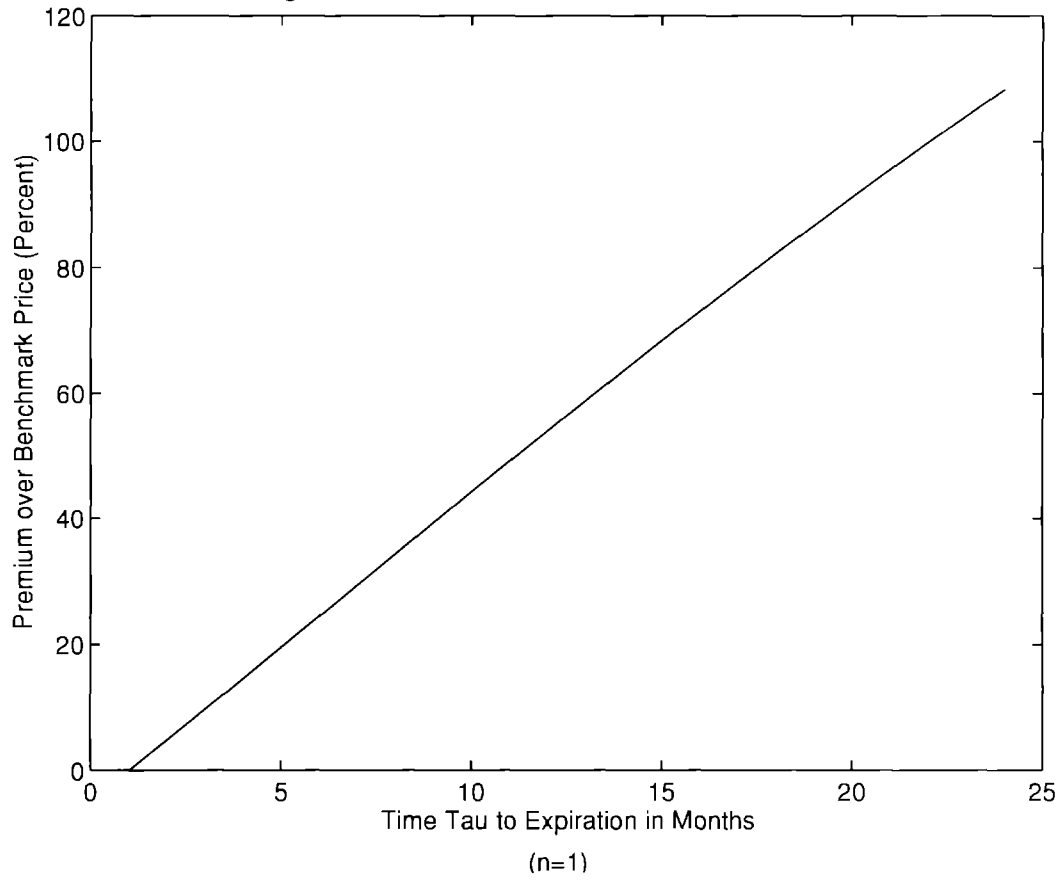Figure 3. Ho and Lee Call Price Premium
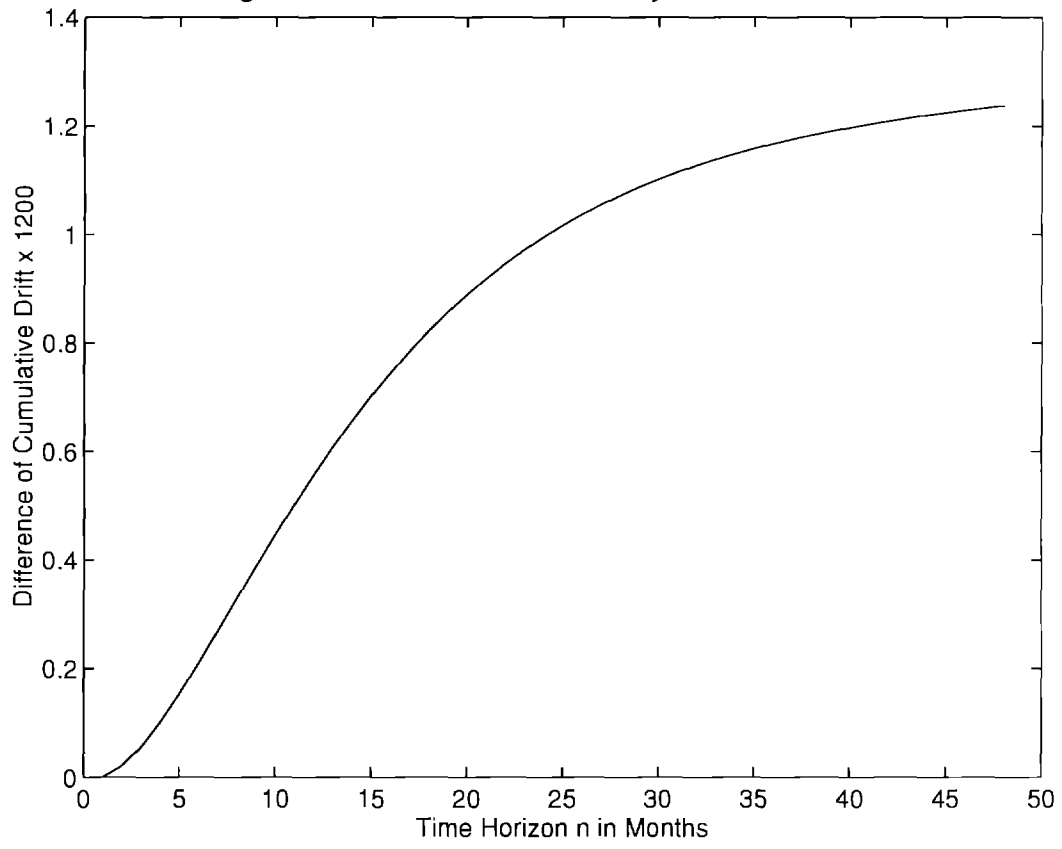
(n=1)

Figure 4. Black–Derman–Toy Drift Parameters

Figure 5. Black–Derman–Toy Call Price Premium

(tau=6)
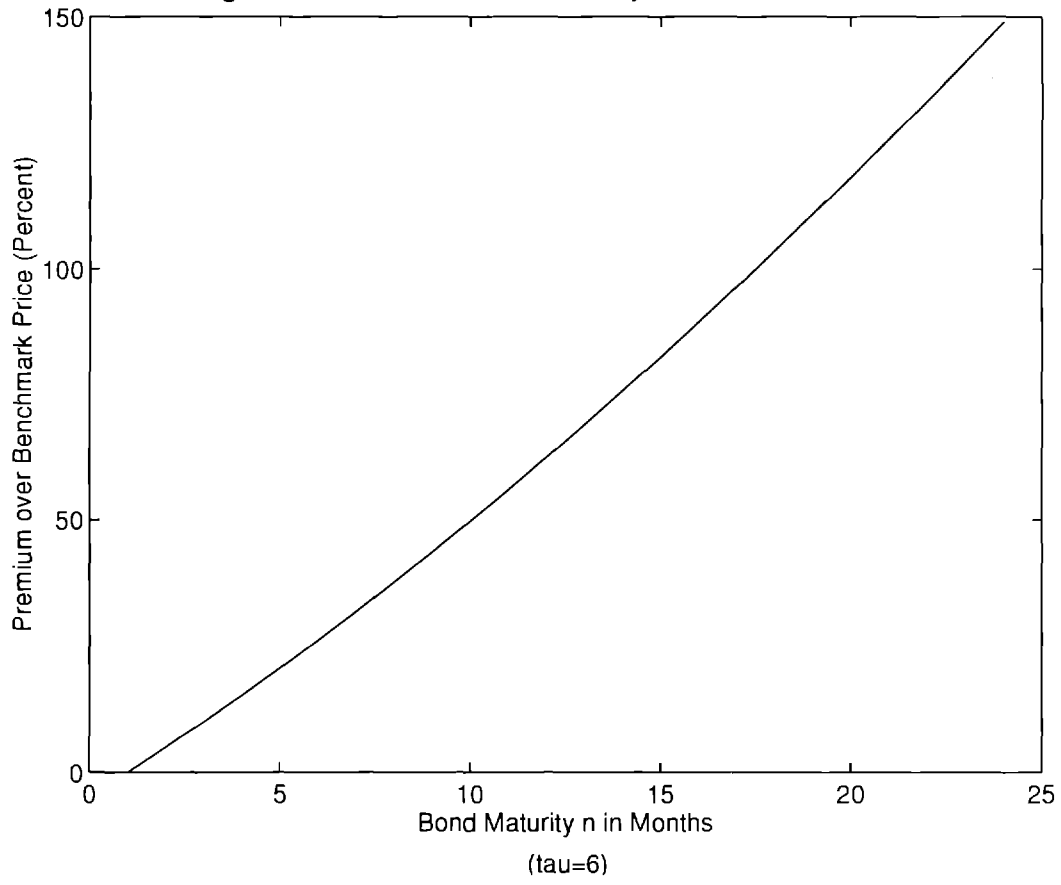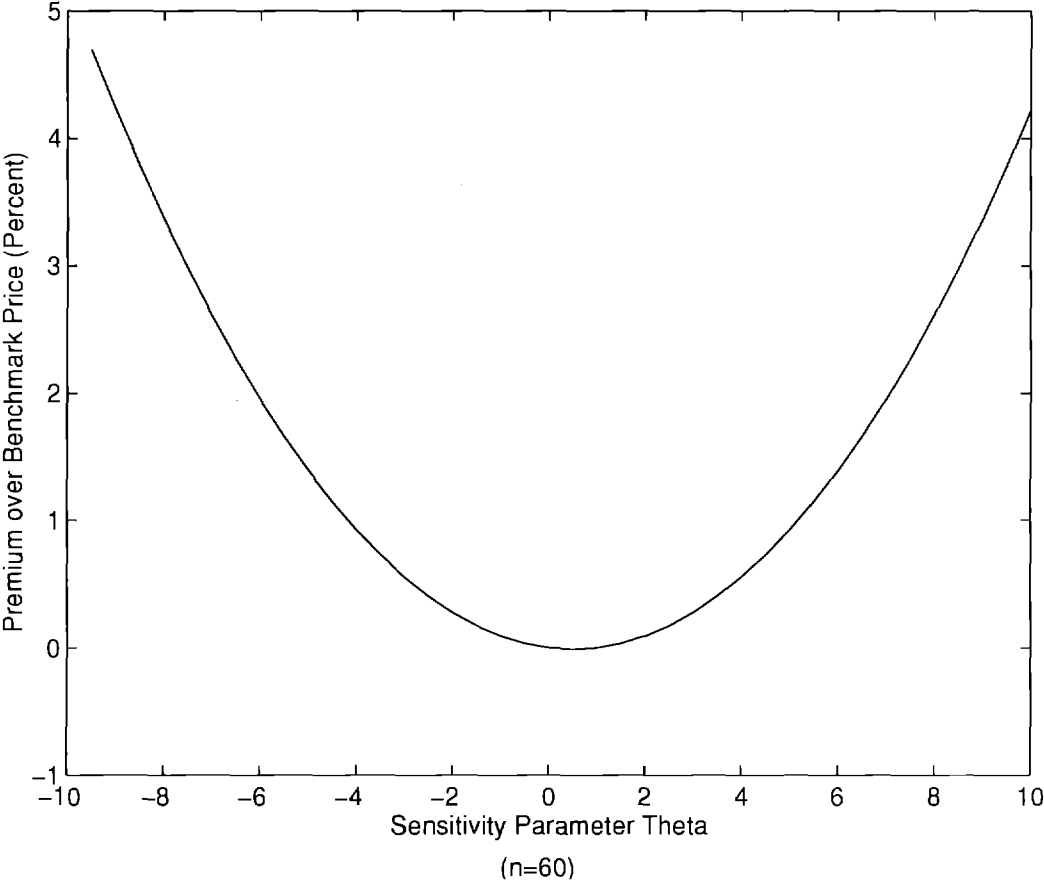
Figure 6. Black–Derman–Toy Exotic Price Premium

Premium over Benchmark Price (Percent) vs. Sensitivity Parameter Theta

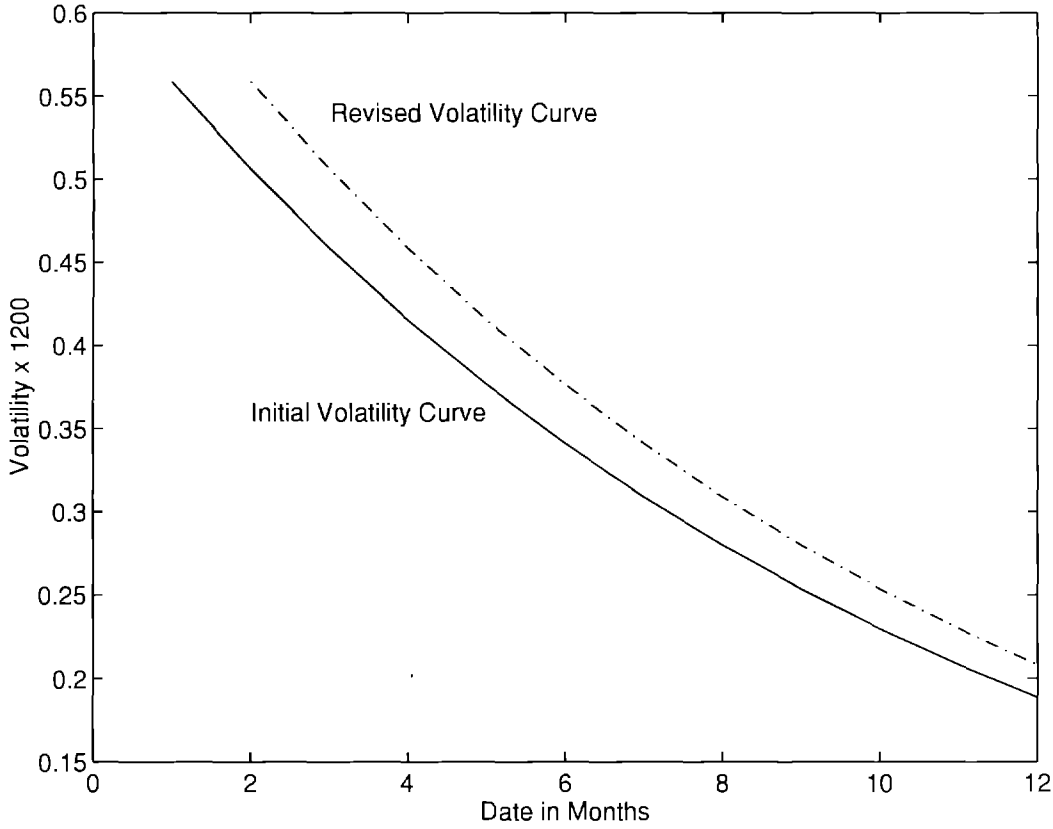(n=60)

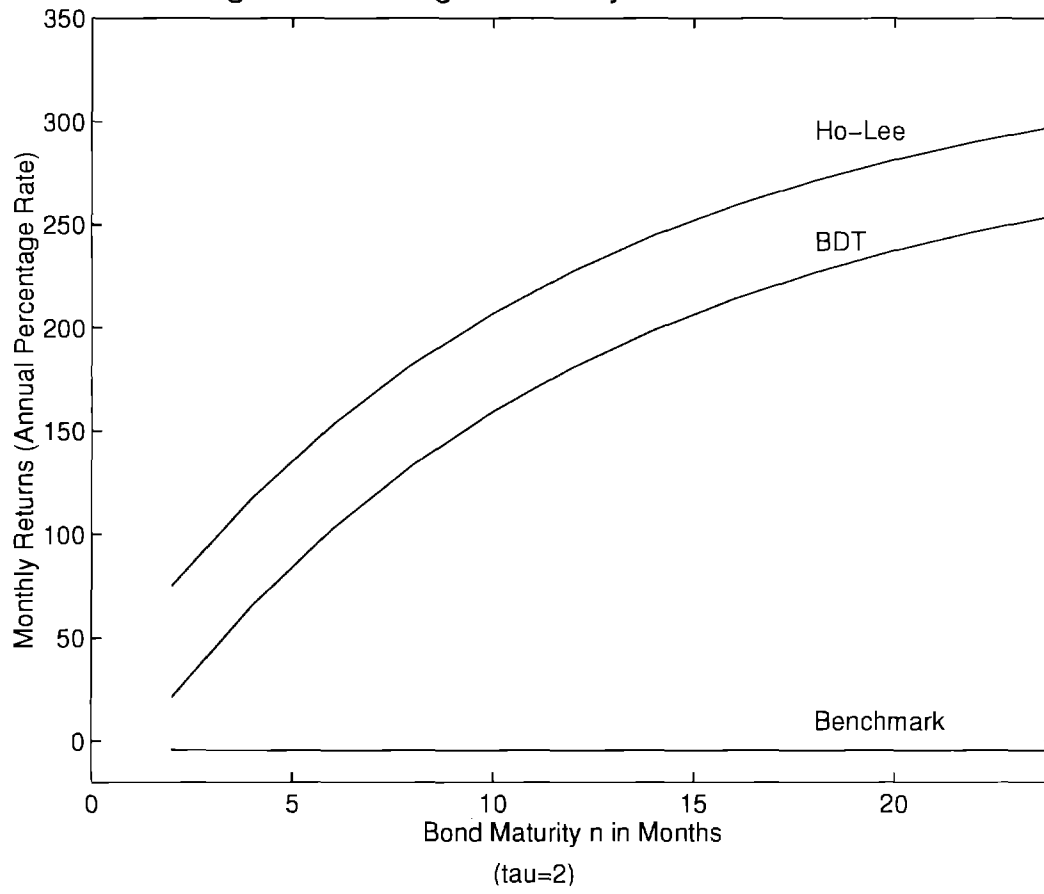Figure 7. Initial and Revised Volatility Parameters

Figure 8. Average Risk–Adjusted Excess Returns

Figure 9. Theoretical Mean Yield Curve with Phi = 0.99

Stars are data points, the line is theory