NEW YORK UNIVERSITY
STERN SCHOOL OF BUSINESS
FINANCE DEPARTMENT

Working Paper Series, 1994

Arbitrage Opportunities in Arbitrage-Free Models of Bond Pricing

David K. Backus, Silverio Foresi, and Stanley E. Zin

FD-94-5

Arbitrage Opportunities in Arbitrage-Free Models of Bond Pricing*

David K. Backus[†] Silverio Foresi[‡] Stanley E. Zin[§]

October 25, 1994

Abstract

We explore the practitioners' methodology of choosing time-dependent parameters to fit a bond model to selected asset prices, and show that it can lead to systematic mispricing of some assets. The Black-Derman-Toy model, for example, is likely to overprice call options on long bonds when interest rates exhibit mean reversion. This mispricing can be exploited, even when no other traders offer the mispriced assets. We argue more generally that time-dependent parameters cannot substitute for sound fundamentals.

JEL Classification Codes: E43, G12, G13.

Keywords: bond prices and yields, forward rates, time-dependent drift and volatility, options.

*We thank Fischer Black for extensive comments and Tony Lynch for pointing out an error in an earlier draft. Backus thanks the National Science Foundation for financial support.

[†]Stern School of Business, New York University and NBER.

[‡]Stern School of Business, New York University.

[§]Graduate School of Industrial Administration, Carnegie Mellon University and NBER.

1 Introduction

Practitioners routinely use time-dependent drift and volatility parameters to fit theoretical bond pricing models to observed asset prices. This procedure was first suggested by Ho and Lee (1986) and has subsequently been widely adopted, with notable applications by Black, Derman, and Toy (1990), Heath, Jarrow, and Morton (1992), and Hull and White (1993). A more comprehensive list of references is provided by Hull (1993, ch 15).

Time-dependent parameters give so-called arbitrage-free models a great deal of freedom to reproduce observed asset prices, and thus to overcome some of the shortcomings of more parsimoniously parameterized models. For practictioners this ability is essential. The relatively parsimonious models used by academics generally provide only rough approximations to observed bond prices and yields. The four-parameter Vasicek (1977) and Cox-Ingersoll-Ross (1985) models, for example, can be calibrated to match five points on the yield curve (the four parameters plus the short rate), but cannot generally reproduce the complete yield curve to the degree of accuracy required by bond traders. The time-dependent drift parameters of the Ho and Lee model, on the other hand, can be calibrated to match an entire yield curve exactly.

Nevertheless, we think the additional parameters of arbitrage-free models raise questions about the theory. One question concerns the sheer number of parameters. Are these extra parameters simply useful refinements of the theory, or are they "Ptolemaic epicycles" that disguise weaknesses in the theory's foundations? Ptolemy's apparent success in explaining planetary motion with the earth at the center is one of the reasons theorists ever since have been suspicious of models with large numbers of free parameters. Another question was raised by 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 asked whether these changes in parameter values through time implied that the framework itself was inappropriate.

We examine the use of time-dependent drift and volatility parameters 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 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 generate accurate prices for all state-contingent claims. In this sense, these "arbitrage-free" models allow arbitrage opportunities: a trader basing prices on, say, the Black-Derman-Toy model could be exploited by a trader who knew the true structure of the economy. The most striking examples of mispricing involve long options on long bonds and "exotic" derivatives that exhibit different sensitivity to interest rate movements than bonds.

We are sympathetic to the practitioners' suggestion that time-dependent parameters can play a useful role in fine-tuning the relatively simple models that populate the academic literature. Our comparison suggests, though, that there is also a benefit to understanding the fundamentals that drive bond prices, whether these fundamentals involve mean reversion, multiple factors, stochastic volatility, or other features. Models that misspecify the fundamentals will invariably 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), the ARMA(1,1) example of Backus and Zin (1994, Section 4). Although in some respects the model 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* $r_t = y_t^1 = f_t^0$. From these definitions we see that the maturity structure of riskfree bonds can be expressed in three equivalent ways: with prices, yields, or forward rates.

We characterize asset prices in our benchmark theory, or laboratory, with a *pricing kernel*: the stochastic process governing prices of state-contingent claims. Existence of such a process is guaranteed in any arbitrage-free environment; see Duffie (1992, ch 1) for further discussion and references to earlier work. We describe the kernel for our theoretical environment in two steps. The first step involves an abstract state variable z, whose dynamics follow

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

with $\{\epsilon_t\}$ 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

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

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. Application to bond pricing is straightforward. Since the one-period return on an $n+1$-period bond is b^n_{t+1}/b^{n+1}_t, the pricing relation gives us

$$b^{n+1}_t = E_t(m_{t+1} b^n_{t+1}). \tag{5}$$

Equation (5) allows us to compute bond prices recursively, starting with the initial condition $b^0_t = 1$.

Consider a one-period bond. From equation (5) and the initial condition $b^0_t = 1$ we see that the price is the conditional mean of the pricing kernel: $b^1_t = E_t m_{t+1}$. Since the kernel is conditionally lognormal, we need the following property of log-normal 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

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

and the short rate is

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

In other words, 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:

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

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

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

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 will 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 depends on 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 only difference from conventional applications of the Black-Scholes formula is the role of the mean reversion parameter φ in (12).

3 Parameter Values

We can get an idea of the impact of the parameters on bond prices in our benchmark theory with an informal moment matching exercise. We base parameter values on 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 for the 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

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

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 quantify differences in prices across our various 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,

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

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 our 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 of

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

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

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

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

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

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)]

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

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.

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 the time-dependent drift parameters do not adequately capture the effects of mean reversion in our benchmark theory. 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

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

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 our 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 the benchmark theory, 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-Lee price over the price generated by the benchmark theory. The benchmark prices are based on the parameter values of Section 3. The 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 our 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

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

and forward rates are

$$\begin{aligned} f_t^n &= r_t + \sum_{j=1}^{n} \alpha_{t+j} + \left[\gamma^2 - (\gamma+n)^2\right]\beta_{t+1}^2/2 \\ &\qquad + \sum_{j=1}^{n} (\gamma+n-j)^2(\beta_{t+j}^2 - \beta_{t+j+1}^2)/2 \end{aligned} \tag{23}$$

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

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

Their conditional mean and variance are

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

and

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

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:

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

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 our theory. Similar patterns of declining time-dependent volatilities are common when these methods are used in practice, Black, Derman, and 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 the time-dependent volatility parameters, the model cannot simultaneously reproduce the conditional moments of the short rate and the forward rate curve of the benchmark theory. The drift parameters that reproduce the forward rate curve, equation (10), are

$$\sum_{j=1}^{n} \alpha_{t+j} = (1-\varphi^n)(\mu - r_t) + \left[\lambda^2 - (\lambda + (1-\varphi^n)/(1-\varphi))^2\right]\sigma^2/2$$
$$- \left[\gamma^2 - (\gamma+n)^2\right]\beta_{t+1}^2/2 - \sum_{j=1}^{n}(\gamma+n-j)^2(\beta_{t+j}^2 - \beta_{t+j+1}^2)/2. \quad (29)$$

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 options on zero-coupon bonds. Once more the lognormal structure of the model means that the Black-Scholes formula, equation (11), is relevant. 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

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

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 our 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 our 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 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 our theory, decays geometrically with the time horizon τ. This feature is easily mimicked in the Black-Derman-Toy model, as we've 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 our 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 our theoretical environment, and thus understate the value of a long-term callable 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: 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,

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

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

$$\begin{aligned}\log d_t^n &= -\left(r_t + (\gamma\beta_{t+1})/2\right) + \theta \log b_t^n + (\gamma + \theta n)^2 \beta_{t+1}^2/2 - \theta\left(n\alpha_{t+1}\right) \\ &\quad -\theta \sum_{j=1}^{n} \left[(n-j)(\alpha_{t+j+1} - \alpha_{t+j}) - (\gamma + n - j)^2(\beta_{t+j+1}^2 - \beta_{t+j})^2/2\right]. \qquad (31)\end{aligned}$$

Both expressions follow from pricing relation (4).

Suppose we choose the parameters of our 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 [$\gamma = \lambda$]. Does 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, our 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 our Black-Derman-Toy analog cannot replicate the state prices of our 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 our benchmark theory implies discount factors

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

and our 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,

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

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 our Black-Derman-Toy analog cannot replicate the prices of derivative assets generated by a model with mean reversion.

6 Arbitrage and Profit Opportunities

In a world governed by our benchmark theory, the Black-Derman-Toy analog systematically misprices some assets. We illustrate how this mispricing can be exploited by another trader.

The simplest 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, 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.

When an alternative price offer is not available, we can often devise dynamic strategies against a Black-Derman-Toy trader alone. Since Black-Derman-Toy prices are consistent with a pricing kernel, they preclude riskless arbitrage. There are, nevertheless, 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 it 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 our theoretical setting and in practice, that the time-dependent parameters must be recalibrated. 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. The stochastic behavior of the drift parameters is particularly troubling, and suggests that they should be treated as additional risk factors, rather than parameters, since they bear on the uncertainty of future asset prices. All of these parameter changes are inconsistent with the logic of the model, as Dybvig (1989) and Black and Karasinski (1991) have noted, but since they improve the model's performance we build them into our trader's behavior. Our trader therefore 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

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

using the pricing kernel m for the benchmark economy. This measure is analogous to Jensen's alpha for the CAPM. For the trading strategy just described, 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 options on bonds for maturities 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 parameters of analogs of the Ho and Lee (1986) and Black-Derman-Toy (1990) models are not sufficient to reproduce prices of state-contingent claims generated by a model with mean reversion. This thought experiment does not tell us how well these models might perform in practice, but we think it suggests that extra parameters will not be able to substitute for sound fundamentals.

We focused on mean reversion because we find it an appealing feature in a bond pricing model. Although the estimated autocorrelation of 0.906 reported in Table 1 is not substantially different from one, a value of one has, in our 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 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 indicated by the properties of bond prices. Our thought experiment indicates that models without it will generally misprice some assets.

Other work indicates that the fundamentals driving bond prices extend well beyond the mean reversion of models like our benchmark. Among the many studies to make this point are Gibbons and Ramaswamy (1993), who report a striking anomaly in the one-factor Cox-Ingersoll-Ross model; Stambaugh (1988), who documents benefits of additional factors; Longstaff and Schwartz (1992), who identify a second factor with volatility in the short rate; and Das (1994), who finds evidence of fat tails in short rate innovations. However, even the best of these models provides only an 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."

This motivation for time-dependent parameters is highly persuasive but ignores, we think, some potential dangers. Just as time-dependent drift and volatility parameters cannot generally reproduce the effects of mean reversion, we would not expect them to be able to reproduce the effects of additional factors, stochastic volatility, or non-normal interest rate innovations. The mean reversion of the Black and Karasinski (1991) model, for example, is not a complete answer if fundamentals are more complex than a first-order autoregression in the logarithm of the short rate.

Indeed, practitioners seem to be aware of some of the hazards of using these arbitrage-free models. One symptom is their tendency to use such models only to price assets similar to those whose prices can be verified in the market. As Ho and Abrahamson (1990, p 332) put it: "[The parameters] should be chosen to most closely replicate the type of security on which options are to be evaluated." The suggestion is that the model is to be used only for local approximations to derivative asset prices, and is less reliable for evaluating derivatives that differ significantly from those whose prices are currently observable. Another symptom is that practitioners do not generally rely on these models to compute hedge ratios, like those implied by the sensitivity of an asset's price to changes in the short rate. We saw in Section 5, for example, that our Black-Derman-Toy analog generally overstates the sensitivity of long bond prices to changes in the short rate. A practitioner using this model to immunize an exposure to long bond prices would therefore overstate the amount of hedging required to reduce his exposure to zero. Both symptoms could reflect weaknesses in the models' theoretical foundations that practitioners have discovered through experience. Best practice, we think, is to combine current knowledge of fundamentals with enough extra parameters to make the approximation adequate for trading purposes.

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.

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 theoretical models. We would not be willing, at present, to bet our salaries on our benchmark theory. Nevertheless, we think the examples illustrate 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 complex 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 on the full range of interest-rate derivative securities. In more general settings, we would expect that a model with n arrays of parameters will be able to reproduce the term structure of prices of n classes of assets, but if the fundamentals are wrong there will always be some assets that are mispriced.

In short: fundamentals matter.

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

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

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

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

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

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

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

and

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

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:

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

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 $\eta_{t+j} = \epsilon_{t+j}$, 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

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

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

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

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.

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

$$c_t^{\tau,n} = E_t\left[M_{t,t+\tau}(b_{t+\tau}^n - k)^+\right]. \qquad \text{(A3)}$$

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

$$E\left[x_1 I(x_2 - k)\right] = \exp(\mu_1 + \sigma_1^2/2)N(d),$$

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_1x_2 for x_1):

$$E\left[(x_1x_2)I(x_2 - k)\right] = \exp(\mu_1 + \mu_2 + \sigma_1^2/2 + \sigma_2^2/2 + \sigma_{12})N(d),$$

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 $b_{t+\tau}^n$ 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

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

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

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

It has conditional mean given by the first two terms and conditional variance $2A_\tau$.

The future bond price can be represented

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

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

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

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

Backus, David, and Stanley Zin (1994), "Reverse engineering the yield curve," unpublished manuscript, New York University, March.

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*, 33-39.

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

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

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," unpublished manuscript, New York University, April.

Duffie, Darrell (1992), *Dynamic Asset Pricing Theory.* Princeton, NJ: Princeton University Press.

Dybvig, Philip (1989), "Bond and bond option pricing based on the current term structure," unpublished manuscript, Washington University, February.

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 Allen Abrahamson (1990), "Options on interest-rate sensitive securities," in S. Figlewski, W. Silber, and M. Subrahmanyam, (eds.), *Financial Options: From Theory to Practice.* Homewood IL: Business One Irwin, 314-356.

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 (1993), *Options, Futures, and Other Derivative Securities, Second Edition.* Englewood Cliffs, NJ: Prentice-Hall.

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.

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," unpublished 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.

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

Table 1

Properties of Bond Yields

Data are monthly estimates of annualized continuously-compounded zero-coupon US government bond yields computed by McCulloch and Kwon (1993). Mean is the sample mean, St Dev the standard deviation, and Auto the first autocorrelation. The sample period is January 1982 to February 1992.

Maturity	Mean	St Dev	Auto
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

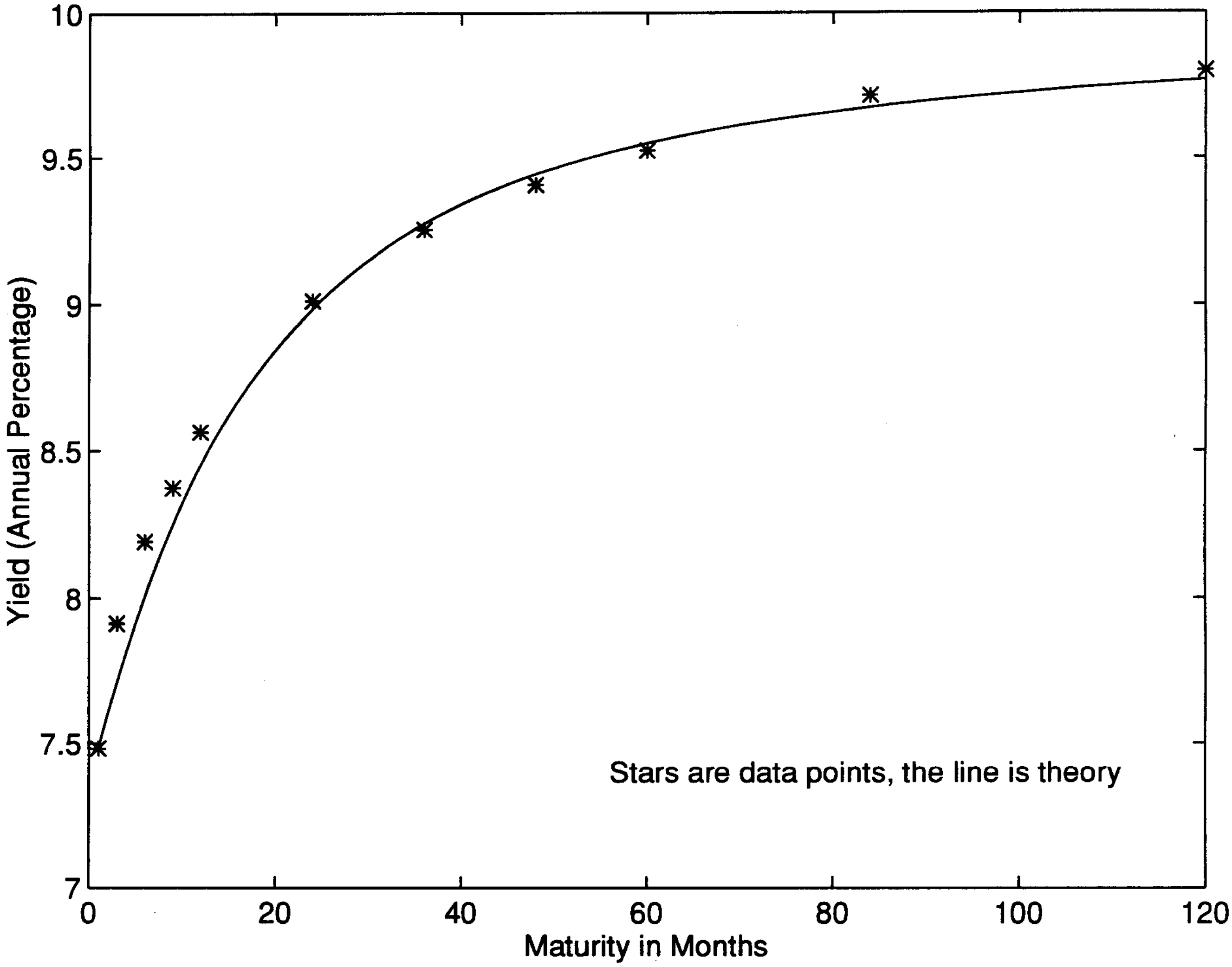


Figure 2. Two Choices of Ho–Lee Drift Parameters

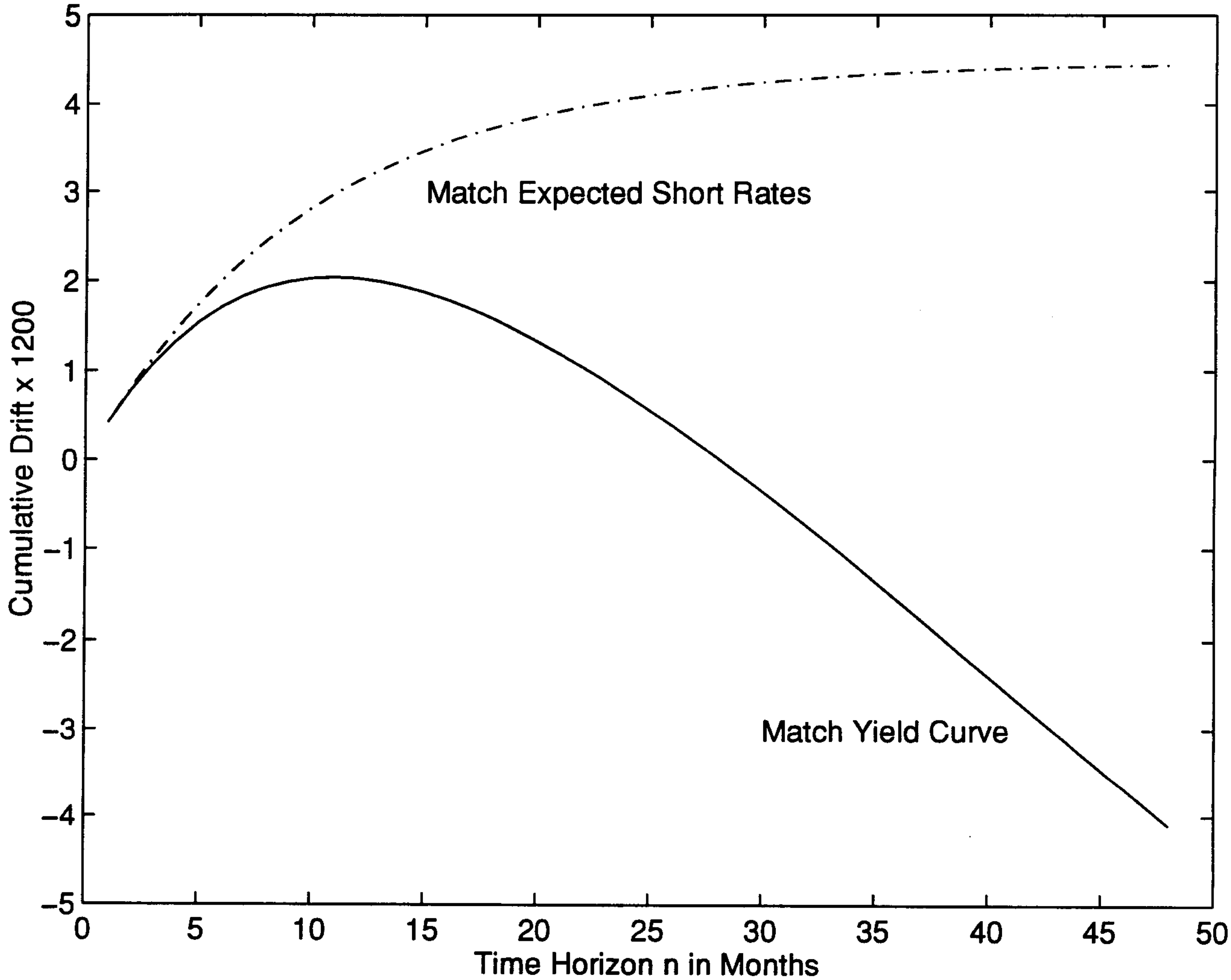


Figure 3. Ho and Lee Call Price Premium

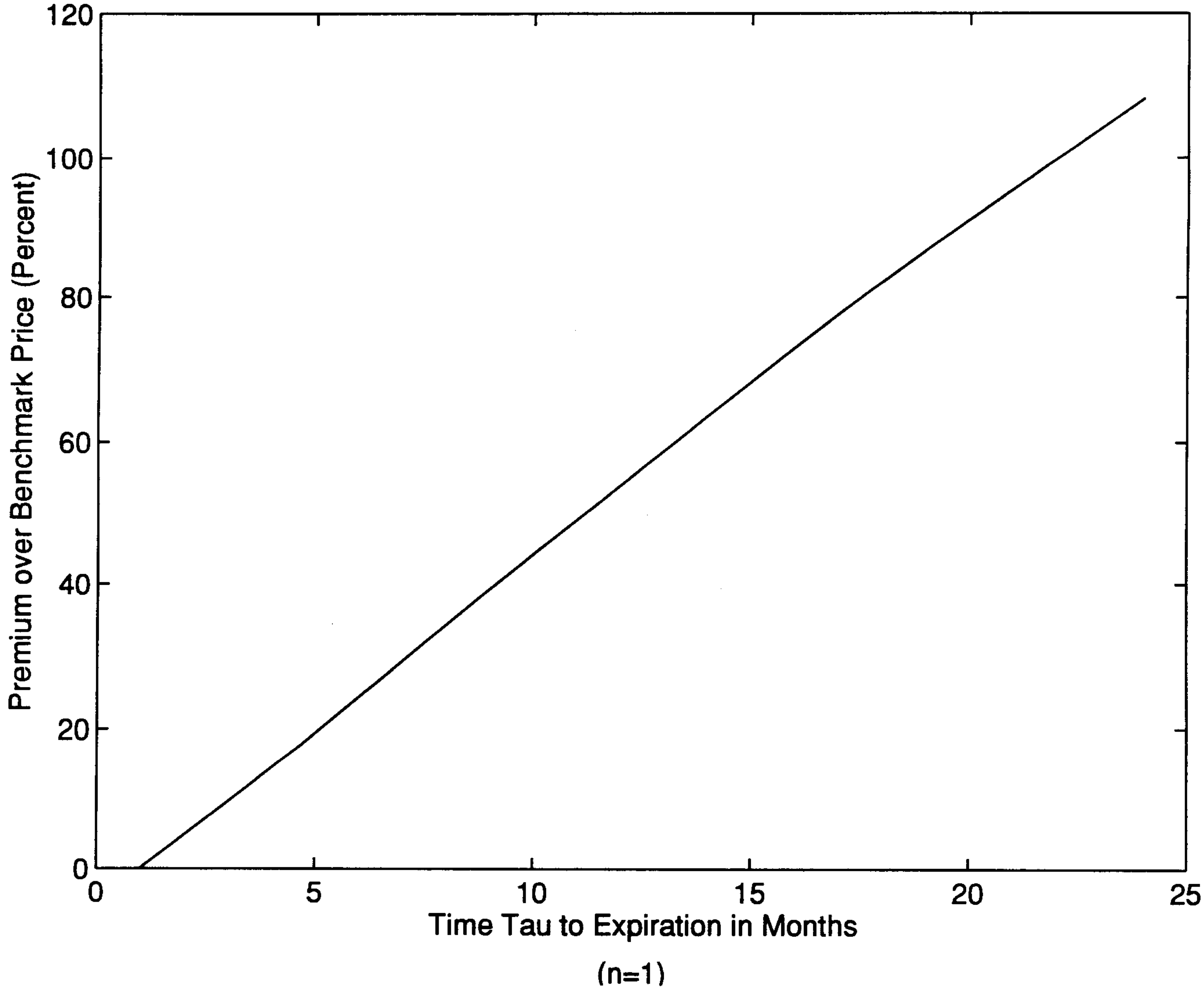


Figure 4. Black–Derman–Toy Drift Parameters

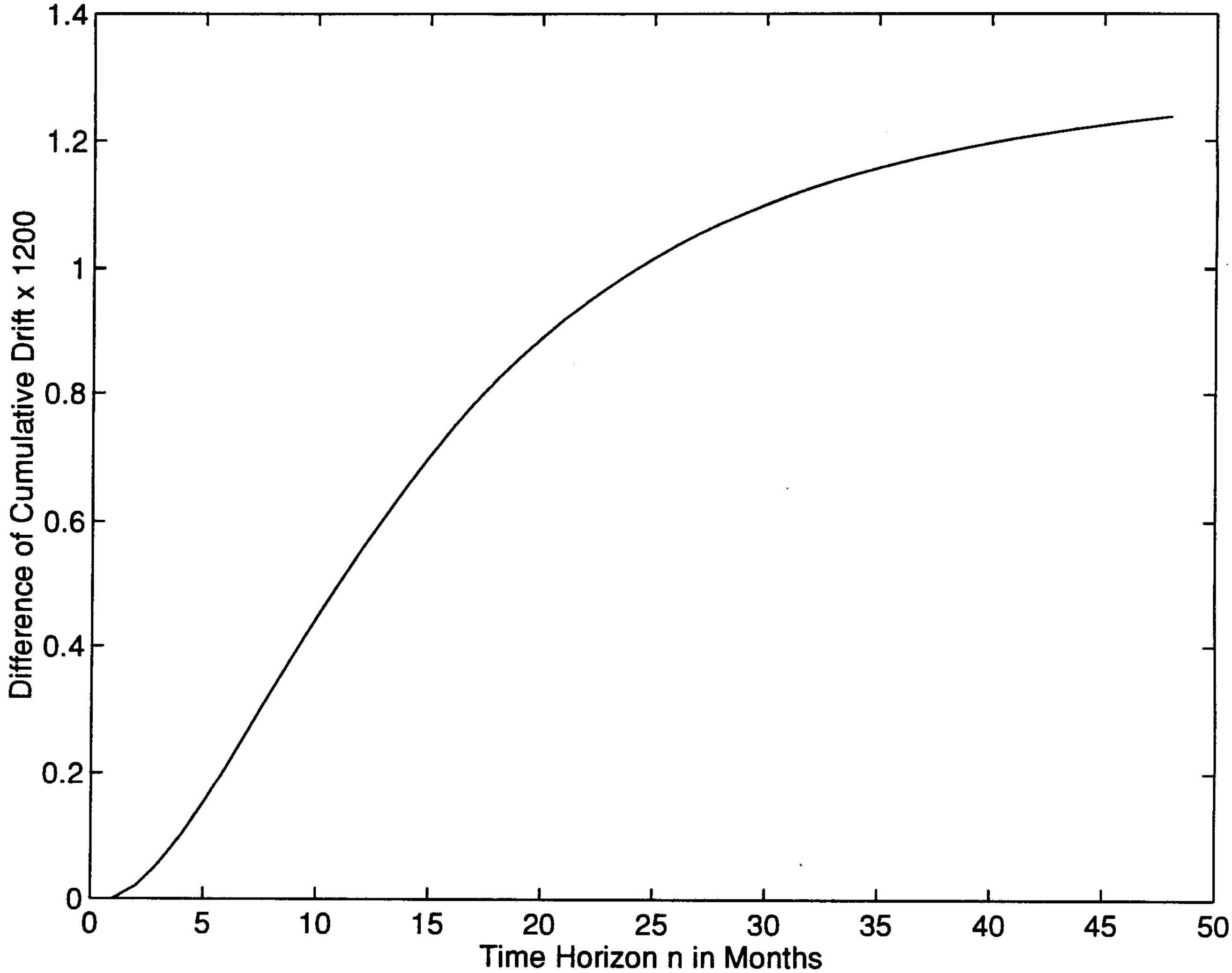


Figure 5. Black–Derman–Toy Call Price Premium

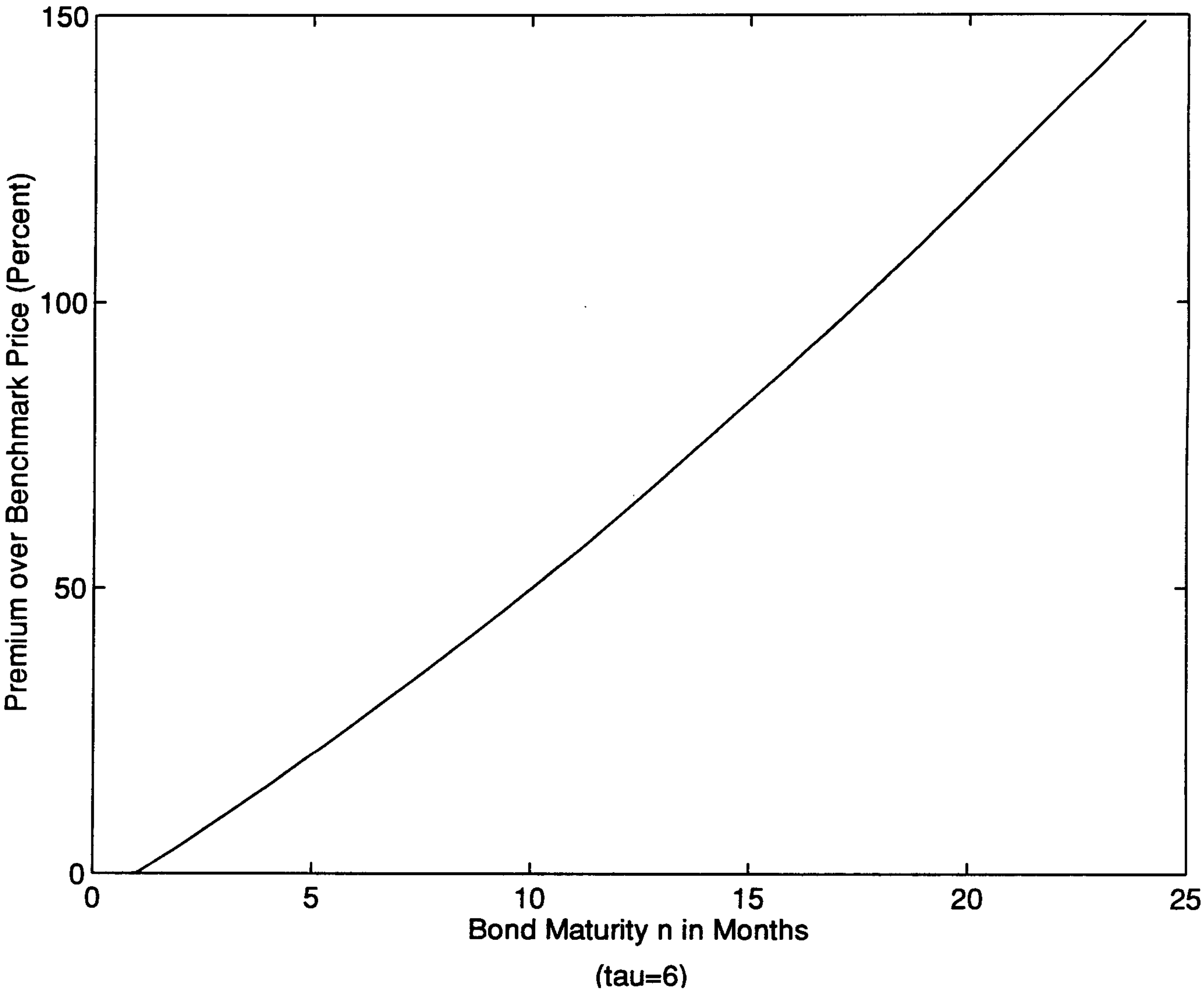
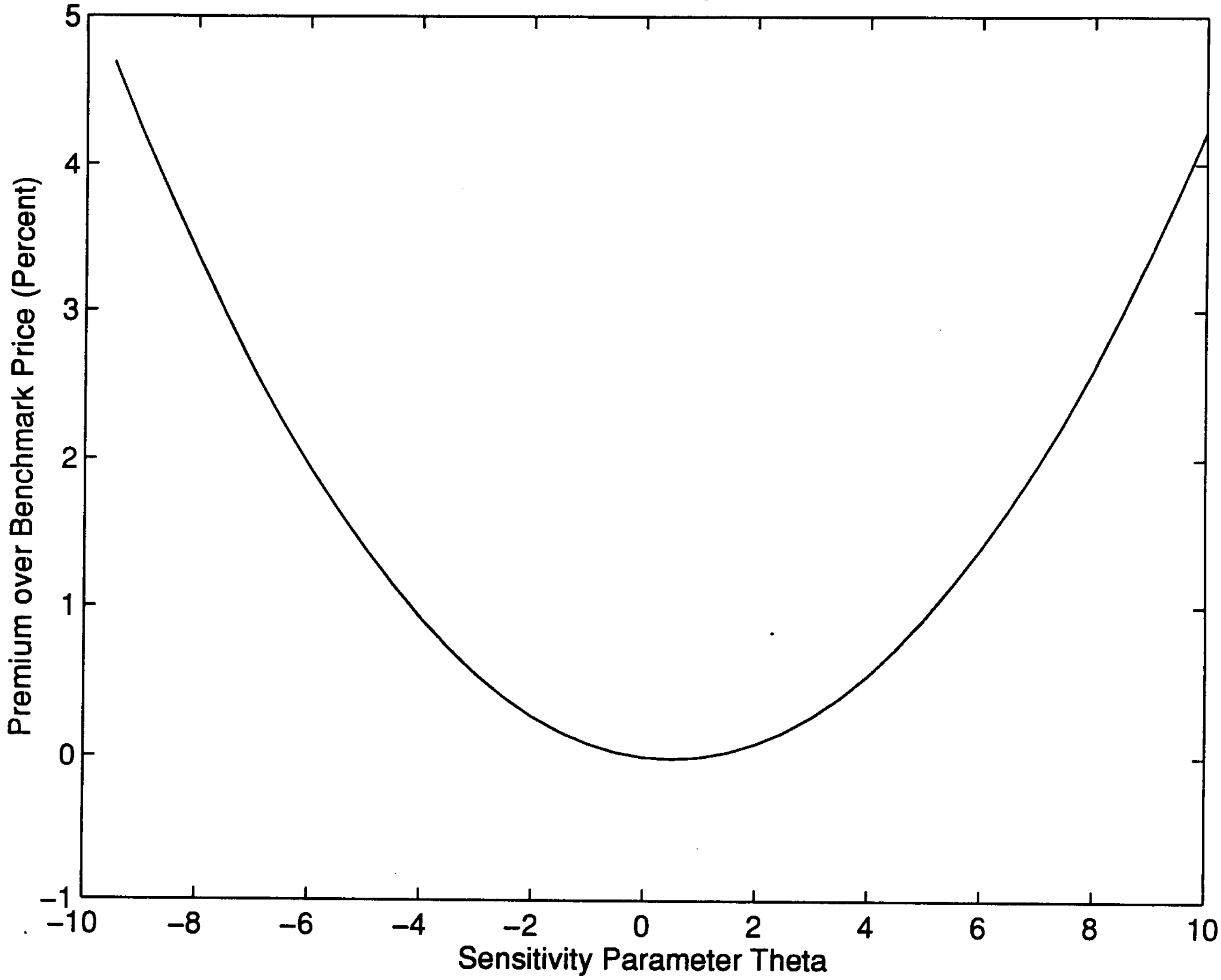


Figure 6. Black–Derman–Toy Exotic Price Premium



(n=60)

Figure 7. Initial and Revised Volatility Parameters

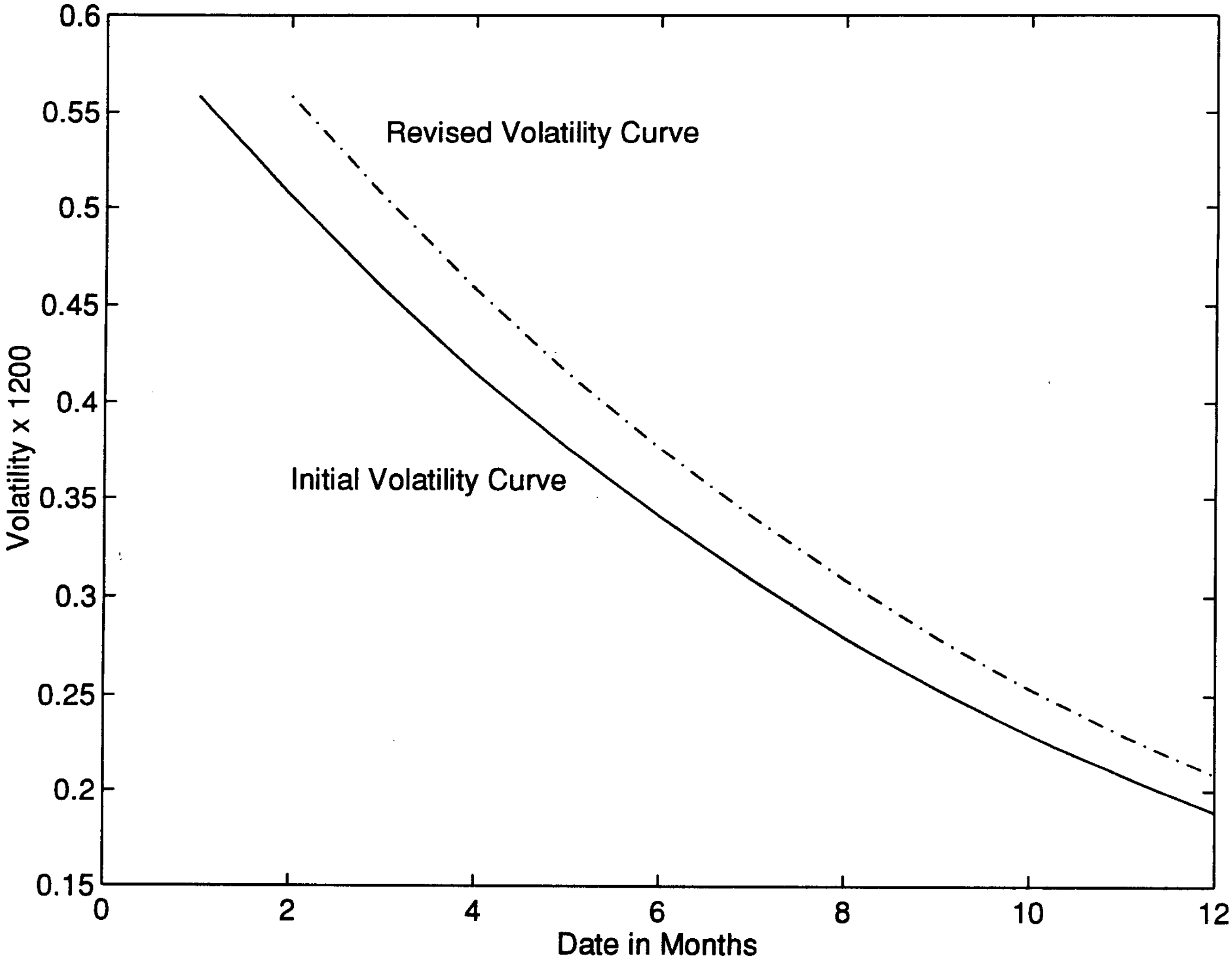
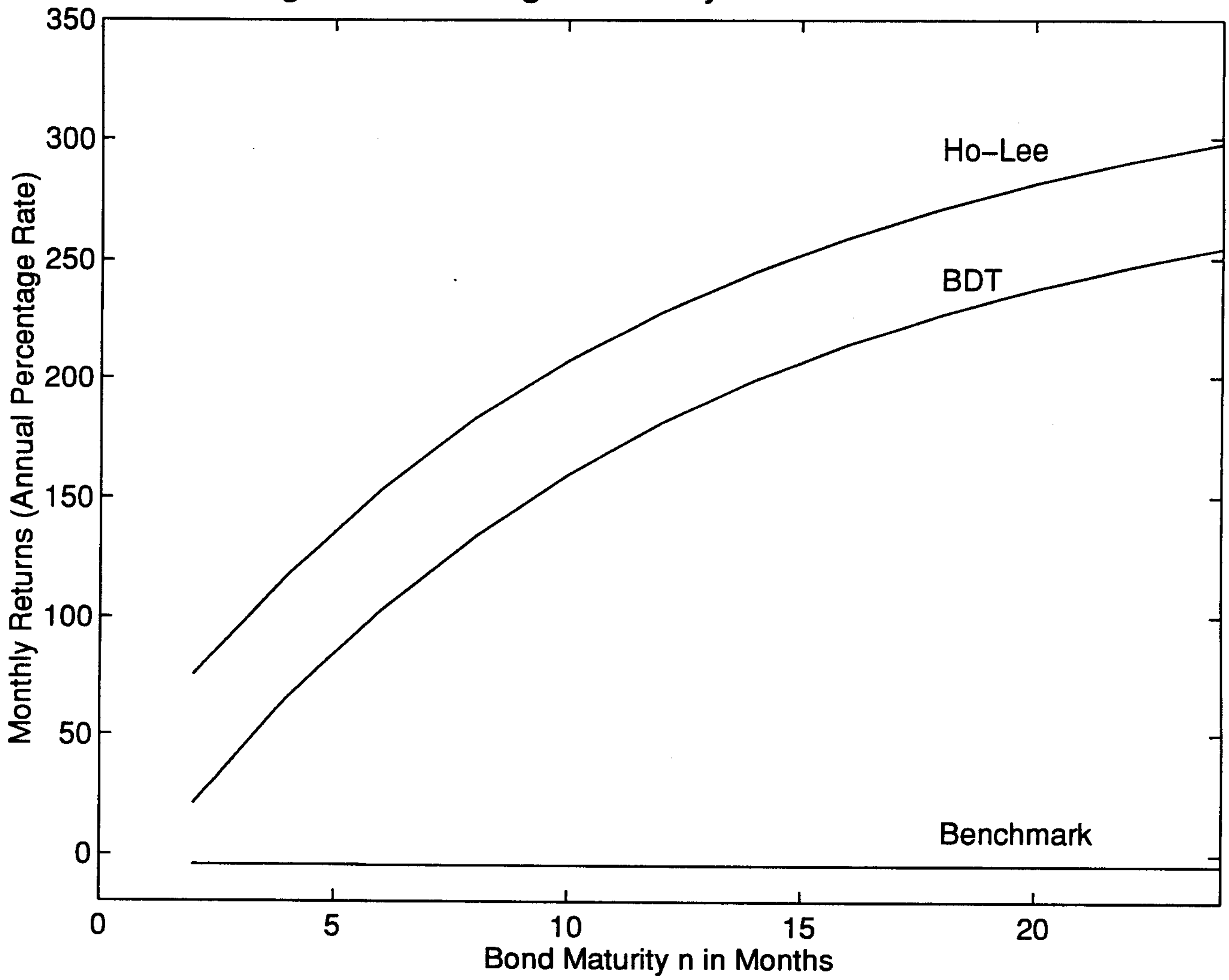


Figure 8. Average risk-adjusted excess returns



(tau=2)

Figure 9. Theoretical Mean Yield Curve with Phi = 0.99