BOND AND BOND OPTION PRICING BASED ON THE CURRENT TERM STRUCTURE*

Philip H. Dybvig
Washington University in Saint Louis

September 1988
revised, March, 1996

ABSTRACT

Ho and Lee derive a term structure model that is based on the term structure at a point in time. The Ho and Lee model is shown to be a a binomial version of Vasicek's model without mean reversion, but with an added deterministic drift as a function of time. From this perspective, we derive a whole class of term structure models based on existing bond and bond option pricing models and the term structure at a point of time. Because the Ho and Lee model has an unreasonable implicit short rate process (that has a larger and larger drift over time), these alternative models are theoretically more realistic. Like the Ho and Lee model, all of these models are *ad hoc* in the sense that most shifts and twists in the yield curve are captured not by changes of state variables but rather by changing the model as a whole period by period. Exploratory data analysis indicates that a one-factor model with slight mean-reversion explains nearly all of the variability of interest rates. Therefore, only the one factor is needed for analyzing the variance, and the Ho and Lee model should fit well for ordinary options on bonds maturing in less than a year. For options on bonds maturing in more than a year, an analogous model based on the Vasicek one-factor mean-reverting model should fit well. Future research should focus on the time series properties of the common factor in bond returns and in particular its variance process.

*I am grateful for useful conversations with Jim Bodurtha, Jon Ingersoll, Chris Lamoureux, Jean Masson, and Ehud Ronn. This is a minor revision (primarily to update the statistics) of a paper written in 1988 and revised slightly in 1989.

1 Introduction

The desire to price interest rate options has placed new demands on our modeling of interest rates and the term structure. An earlier literature was content to ask about the relation between today's term structure and the expectation of future interest rates. Now, we require a complete model of co-movements of bond prices so we can compute the dynamic hedges used to price a universe of interest-sensitive securities. Ho and Lee [1986] have combined the two traditions, by building a dynamic hedging model that reflects the information contained in the current term structure. This combination comes at the expense of an *ad hoc* structure in which many types of interest rate movements are taken as shifts in the model rather than shifts in the state variables (as in a correctly specified model). This paper has two primary purposes. One primary purpose is to synthesize existing theoretical results to show how to derive a whole class of models in the spirit of Ho and Lee's model. The class allows us to take advantage of existing models for which closed-form bond and bond option pricing formulas exist. The other primary purpose is to examine theoretically and empirically the features of interest rate co-movements that have made the Ho and Lee model useful for practitioners in spite of its *ad hoc* features.

The analysis of the Ho and Lee Model in this paper has some common ground with the discussion of Ho and Lee in Heath, Jarrow, and Morton [1990,1992]. This paper's perspective on Vasicek models was previously discussed in a later section of Dybvig [1988]. The empirical results here are based on that perspective. This perspective could be viewed as consistent with leading cases of the Heath, Jarrow, and Morton [1992] approach of treating the term structure itself as the set of state variables (but without the complicated notation and exposition). However, the philosophy here is much different, with the emphasis being on having a model in which all sorts of local moves in the shape of the observed discrete points on the yield curve are possible. This need not be inconsistent with the Heath, Jarrow, and Morton model (if finitely many maturities are observed), but it is inconsistent with most implementations of the model.

2 The Ho and Lee Model

We will start with the primitive assumption that prices are given by discounted expected values, where discounting uses the rolled-over short rate. For example, in continuous time the price P_s at s of an asset paying P_t at $t > s$ is given by

$$P_s = E_s\left[P_t e^{-\int_{\tau=s}^{t} r_\tau d\tau}\right], \tag{1}$$

where r_τ is the instantaneous riskless rate at time τ and E_s is expectation conditional on information available at time s. There are two possible interpretations of this assumption. The most general is that we are working in the *risk-neutral probabilities* of Cox and Ross [1976] (referred to by Harrison and Kreps [1979] as the *equivalent martingale measure* and by Cox, Ross, and Rubinstein [1979] as the *artificial probabilities*). Provided we are concerned only about option pricing theories (and not about optimal portfolio choice), using risk-neutral probabilities is without loss of generality, although any empirical work must keep in mind that these are not the same as the actual probabilities. A more specific interpretation is that the local expectations hypothesis (Cox, Ingersoll, and Ross [1981]) holds, in which case the actual probabilities are the risk-neutral probabilities. Except as noted, it does not matter which interpretation is taken. When we refer to expectations, they will be the expectations as in (1), whichever interpretation is taken.

The discrete time version of (1) is similar, except with a sum in place of the integral. In particular, if each time interval has length Δt,

$$P_{m\Delta t} = E_{m\Delta t}\left[P_{n\Delta t} e^{-\sum_{i=m}^{n-1} r_i \Delta t}\right], \tag{2}$$

where r_τ is now the logarithmic short rate. Our analysis of the Ho and Lee [1986] model can specialize (2) to

$$P_s = E_s\left[P_t\, e^{-\sum_{\tau=s}^{t-1} r_\tau}\right], \tag{3}$$

because they take $\Delta t = 1$. (We will want to look at the general form (2) later when we want to take limits as Δt goes to 0.)

Unlike most of the literature on the term structure, the primitive in Ho and Lee's model is the vector of discount bond prices. Without repeating their motivation and derivation, here are the essential elements of their model. Let the discount bond price $D_{i,n}$ be the value at time i of receiving 1 for sure at time n (which is $n - i$ periods out). The assumed process has as parameters a risk-neutral probability π of the "upstate," the complementary risk-neutral probability $1 - \pi$ of the "downstate," and positive functions $h(j)$ and $h^*(j)$ specifying the up and down shifts of the yield curves. Specifically,

$$D_{i,n} = h(n - i)\, D_{i-1,n}/D_{i-1,i} \tag{4}$$

if the upstate occurs at time i and

$$D_{i,n} = h^*(n - i)\, D_{i-1,n}/D_{i-1,i} \tag{5}$$

if the downstate occurs at time i. There is a draw of an upstate and a downstate at each time i, and the draws are i.i.d. over time. The functions h and h^* are assumed to satisfy

$$\pi\, h(j) + (1-\pi)\, h^*(j) \equiv 1, \tag{6}$$

which ensures that π and $1-\pi$ are correctly interpreted as risk- neutral probabilities, and to satisfy

$$h(0) = h^*(0) = 1. \tag{7}$$

Additional structure is also assumed (their assumption that up-down and down-up give the same final term structure) so that h and h^* are determined up to a parameter δ to be

$$h(t) = \frac{1}{(\pi + (1-\pi)\,\delta^t)} \tag{8}$$

and

$$h^*(t) = \frac{\delta^t}{(\pi + (1-\pi)\,\delta^t)}. \tag{9}$$

Which of the upstate and downstate corresponds to an upward shift in the yield curve depends on whether δ is larger or smaller than 1.

Now we want to restate the Ho and Lee model in more familiar terms. (This is really just elaborating on the analysis in Section III of Ho and Lee's paper.) The short rate at time i is

$$r_i = -\log(D_{i,i+1}), \tag{10}$$

and the 1-period forward rate quoted at i for money borrowed or lent from n to $n+1$ is

$$f_{i,n} = -\log(D_{i,n+1}/D_{i,n}). \tag{11}$$

Note that $D_{i,i} = 1$ and therefore

$$r_i \equiv f_{i,i}. \tag{12}$$

By (4), (8), and (11), we have that

$$\begin{aligned} f_{i,n} &= -\log\left(\frac{D_{i,n+1}}{D_{i,n}}\right) \\ &= -\log\left(\frac{h(n+1-i)\, D_{i-1,n+1}/D_{i-1,i}}{h(n-i)\, D_{i-1,n}/D_{i-1,i}}\right) \\ &= f_{i-1,n} - \log\left(\frac{h(n+1-i)}{h(n-i)}\right) \\ &= f_{i-1,n} + \log\left(\frac{\pi + (1-\pi)\,\delta^{n+1-i}}{\pi + (1-\pi)\,\delta^{n-i}}\right) \end{aligned} \tag{13}$$

in the upstate. Similarly, (5), (9), and (11) imply that

$$\begin{aligned} f_{i,n} &= f_{i-1,n} - \log\left(\frac{h^*(n+1-i)}{h^*(n-i)}\right) \\ &= f_{i-1,n} + \log\left(\frac{\pi + (1-\pi)\,\delta^{n+1-i}}{\pi + (1-\pi)\,\delta^{n-i}}\right) - \log(\delta) \end{aligned} \tag{14}$$

in the downstate.

To make better sense of the formulas in (13) and (14), it is useful to separate the change in $f_{\cdot,n}$ into a constant and a mean-zero noise term. Define ε_i by

$$\varepsilon_i = \begin{cases} (1-\pi)\log(\delta) & \text{if upstate at time i} \\ -\pi\log(\delta) & \text{if downstate at time i.} \end{cases} \tag{15}$$

Then $E[\varepsilon_i] = 0$ and we can rewrite (13) and (14) as

$$f_{i,n} = f_{i-1,n} + \log\left(\frac{\pi + (1-\pi)\,\delta^{n+1-i}}{\pi + (1-\pi)\,\delta^{n-i}}\right) - (1-\pi)\log(\delta) + \varepsilon_i. \tag{16}$$

This says that $f_{i,n}$ is the sum of $f_{i-1,n}$, a constant depending on $n-i$ and the parameters π and δ (the two middle terms of (16)), and a random noise term that is the same for all maturities n and i.i.d. across time.

In (16) we have the expression for movements of the forward rates. Now we turn towards the analogous expression for movements in the short rate r. First, note that we can aggregate (16) over time to obtain the following expression for $f_{i,n}$ as it depends on $f_{0,n}$, the ε's, and the parameters π and δ. Applying (16) recursively we have that

$$\begin{aligned} f_{i,n} &= f_{0,n} + \sum_{j=1}^{i}\left(\log\left(\frac{\pi + (1-\pi)\delta^{n+1-j}}{\pi + (1-\pi)\,\delta^{n-j}}\right) - (1-\pi)\log(\delta) + \varepsilon_j\right) \\ &= f_{0,n} + \log\left(\frac{\pi + (1-\pi)\,\delta^{n}}{\pi + (1-\pi)\,\delta^{n-i}}\right) - i(1-\pi)\log(\delta) + \sum_{j=1}^{i}\varepsilon_j \end{aligned} \tag{17}$$

From (17) and (12), we can specialize (17) to an expression for r_i.

$$\begin{aligned} r_i &= f_{0,i} + \log\left(\frac{\pi + (1-\pi)\,\delta^{i}}{\pi + (1-\pi)\,\delta^{0}}\right) - i(1-\pi)\log(\delta) + \sum_{j=1}^{i}\varepsilon_j \\ &= f_{0,i} + \log\left(\pi + (1-\pi)\,\delta^{i}\right) - i(1-\pi)\log(\delta) + \sum_{j=1}^{i}\varepsilon_j \end{aligned} \tag{18}$$

Differencing (18) gives us that

$$r_i = r_{i-1} + (f_{0,i} - f_{0,i-1}) + \log\left(\frac{\pi + (1-\pi)\,\delta^{i}}{\pi + (1-\pi)\,\delta^{i-1}}\right) - (1-\pi)\log(\delta) + \varepsilon_i. \tag{19}$$

In other words, today's interest rate is the sum of yesterday's interest rate, the relevant slope of the initial yield curve, a constant depending on time and the parameters π and δ, and i.i.d. noise. Note further from (16) and (19) that the innovation to each forward rate is the same as the innovation to the spot rate, so that different yield curves at a point in time are all parallel shifts of each other.

The stochastic part of the Ho and Lee interest rate process (19) is a binomial version of the Vasicek [1977] model without mean reversion in interest rates.[1] In this degenerate model, changes in interest rates are i.i.d. normal random variables with mean zero (in the risk-neutral probabilities). (In general, the Vasicek model admits a large class of joint normally distributed interest rate processes.) This is referred to as the driftless absolute interest rate process—absolute refers to the fact that the variance of the spot rate is the same absolute number independent of the interest rate and everything else known before the realization. It is well-known that this particular model is not a useful model of the term structure (expect perhaps as an approximation at short maturities), because the model implies that $\lim_{n\uparrow\infty} f_{i,n} = -\infty$ and $\lim_{n\uparrow\infty} D_{i,n} = \infty$.[2] Intuitively, the reason is that although the mean interest rate at future dates is not unreasonable (it is equal to today's rate), its variance is too large and allows long runs of negative interest rates in some realizations, and these realizations dominate the expectation in (3) for large t.

Given that the Ho and Lee model has the same unreasonable variance process as the driftless absolute Vasicek model, how can it be consistent with reasonable initial yield curves? The answer is that the Ho and Lee model assumes an equally unreasonable mean interest rate process which ultimately is tending towards $-\infty$ at a rate linear in time. Assuming the forward rate tends to a constant at large times to maturity, the expected change in interest rates from time $i-1$ to i is

$$\lim_{i\uparrow\infty} E[r_i - r_{i-1}] = \begin{cases} -(1-\pi)\log(\delta) & \text{if } \delta < 1 \\ \pi\log(\delta) & \text{if } \delta > 1. \end{cases} \tag{20}$$

The surprising feature of (20) is that the drift in both cases is equal to minus the smaller possible value of epsilon. In other words, for large i the drift plus noise takes on two values, one of which is approximately zero and the other of which is positive and not approximately zero.[3]

[1] It may seem from (19) that the Ho and Lee model agrees essentially with the Vasicek model without mean reversion when $\delta = 1$ and the initial term structure is flat. However, when $\delta = 1$, $\varepsilon_i \equiv 0$ (by (15)).

[2] In general, suppose $r_t = r_{t-1} + \varepsilon_t$ where the ε_t's are nonconstant i.i.d. shocks with $E[\varepsilon_t] \leq 0$ under the risk-neutral probabilities. Then, letting ε be drawn from the common distribution of the ε_t's, (3) and (11) imply that $D_{i,n} = \exp(-r_i(n-i))\prod_{t=i+1}^{n} E[\exp(-\varepsilon(n-i))]$ and $f_{i,n} = r_i - \log(E[\exp(-\varepsilon(n-i))])$. Now, ε nonconstant and $E[\varepsilon] \leq 0$ imply that $\lim_{n\uparrow\infty} E[\exp(-\varepsilon(n-i))] = \infty$, and therefore $\lim_{n\uparrow\infty} f_{i,n} = -\infty$ and $\lim_{n\uparrow\infty} D_{i,n} = \infty$.

[3] An interesting corollary of (20) is that the limiting forward rate can never fall, that is, $\lim_{n\uparrow\infty} f_{i,n} \geq \lim_{n\uparrow\infty} f_{i-1,n}$ with probability one. This is a special case of a very general result proven in Dybvig, Ingersoll,

To summarize the results of this section, the 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 patch using the initial yield curve is designed to give correct pricing of discount bonds at time 0, there is every reason to believe it will give incorrect pricing of interest rate options. After all, changing means and changing variances cannot have offsetting effects simultaneously on all types of options. Therefore, we should not rely on the Ho and Lee model for pricing options except at very short maturities for which many one-factor models give similar option prices.

The next section derives a class of models in the spirit of Ho and Lee's model (in the sense of using initial term structure data), but with two important advantages. First, the models permit interest rate processes with reasonable means and variances. Second, the class of models allows us to obtain a closed form solution for the yield curve and bond option pricing using any existing model of the term structure for which closed form solutions exist.

3 Generalizing Ho and Lee

The Ho-Lee model and its successors seem to be based on two premises. First, it is important for a pricing model to fit today's term structure. Second, it is difficult to work with a multifactor model of the term structure. These two premises lead us to using a model with a single source of noise but a very flexible parameter space to fit today's term structure. In this section we start to explore these two premises. We find that it is actually easy to fit today's term structure and that it it is easy to build multifactor models. A later section shows that it is also easy to estimate multifactor models.

First we turn to the question of building multifactor models. Cox, Ingersoll, and Ross [1985] implicitly used a clever trick to generate multifactor models of the term structure from single-factor models. Using continuous-time notation (as in (1)), if $\{r_t^a\}$ and $\{r_t^b\}$ are two independent[4] single-factor term structure models with discount bond pricing given by $D_{s,t}^a = E_s[exp(-\int_{\tau=s}^t r_\tau^a d\tau)]$ and $D_{s,t}^b = E_s[exp(-\int_{\tau=s}^t r_\tau^b d\tau)]$ respectively, then the interest rate process $\{r_t^c\}$ defined by $r_t^c \equiv r_t^a + r_t^b$ has discount bond pricing given by the product $D_{s,t}^c = D_{s,t}^a D_{s,t}^b$.

Obviously, by repeating this procedure (using the derived multi-factor interest-rate processes

and Ross [forthcoming]. The result proven there (that the limiting forward rate can never fall) relies only on the existence of limiting forward rates and absence of arbitrage.

[4] Not all of the two-factor models developed by Cox, Ingersoll, and Ross use independent factor processes.

in place of the single-factor interest rate processes), one can obtain a term structure process with as many factors as one desires. As Cox, Ingersoll, and Ross pointed out, we can interpret the factors as a real rate and an inflation rate, but of course none of the analysis relies on this interpretation.

Ho and Lee [1986] can be interpreted in this context. Specifically, one of the interest rate processes is taken to be the driftless absolute version of the Vasicek model, and the other interest rate process is a deterministic function of time (a "fudge factor"). In general, the decomposition of an interest rate process into a stochastic interest rate process plus a constant function of time is ambiguous up to a constant function of time. For the Ho-Lee model, one choice is to decompose (18) as follows: the Vasicek process is

$$r_i^a = r_0 + \sum_{j=1}^{i} \varepsilon_j \tag{21}$$

(where r_0 is the observed short rate) and the deterministic process is

$$r_i^b = f_{0,i} - r_0 + \log\left(\pi + (1-\pi)\,\delta^i\right) - i(1-\pi)\log(\delta), \tag{22}$$

where r_0 is again the observed short rate and the $f_{0,i}$'s are the observed forward rates.

This observation motivates a whole class of term-structure models. Each model in the class uses a known term structure model and a deterministic model. Given the known term structure model, the deterministic model can be chosen to fit the initial yield curve. The variance assumption is as reasonable as it is in the chosen term structure model. The following Theorem shows that any bond or option pricing results available in the known term structure model can be extended to the new model.

Theorem 1 *Let $\{r_t^a\}$ be any interest rate process, let $\{r_t^b\}$ be an interest rate process that is a function of time alone, and let the interest rate process $\{r_t^c\}$ be defined by $r_t^c \equiv r_t^a + r_t^b$. Then if we let discount bond pricing for processes a and b be given by $D_{s,t}^a = E_s[exp(-\int_{\tau=s}^t r_\tau^a d\tau)]$ and $D_{s,t}^b = exp(-\int_{\tau=s}^t r_\tau^b d\tau)$ respectively, then the discount bond price for process c is given by the product $D_{s,t}^c = D_{s,t}^a D_{s,t}^b$. Furthermore, if interest rates follow the c process, consider an option that pays at t some function $f(r_t^c, D_{t,T_1}^c, \ldots, D_{t,T_n}^c)$ of the short rate r_t^c at t and various bond prices $D_{t,T_1}^c, \ldots, D_{t,T_n}^c$. At time $s < t$, this option has the same price as the option paying $D_{s,t}^b f(r_t^a + r_t^b, D_{t,T_1}^a D_{t,T_1}^b, \ldots, D_{t,T_n}^a D_{t,T_n}^b)$ at t under process a.*

PROOF By (1),

$$D_{s,t}^c = E_s\left[e^{-\int_{\tau=s}^t r_\tau^c d\tau}\right]$$

$$
\begin{aligned}
&= E_s\left[e^{-\int_{\tau=s}^{t}(r_\tau^a+r_\tau^b)d\tau}\right] \\
&= E_s\left[e^{-\int_{\tau=s}^{t} r_\tau^a d\tau}\right] e^{-\int_{\tau=s}^{t} r_\tau^b d\tau} \\
&= D_{s,t}^a D_{s,t}^b,
\end{aligned}
$$

which is the first result we are to prove.

By (1) and the first result, the price at s of receiving $f(r_t^c, D_{t,T_1}^c, \ldots, D_{t,T_n}^c)$ at t under the process c is equal to

$$
\begin{aligned}
& E_s\left[f(r_t^c, D_{t,T_1}^c, \ldots, D_{t,T_n}^c)e^{-\int_{\tau=s}^{t} r_\tau^c d\tau}\right] \\
&= E_s\left[f(r_t^a + r_t^b, D_{t,T_1}^a D_{t,T_1}^b, \ldots, D_{t,T_n}^a D_{t,T_n}^b)e^{-\int_{\tau=s}^{t}(r_\tau^a+r_\tau^b)d\tau}\right] \\
&= E_s\left[e^{-\int_{\tau=s}^{t} r_\tau^b d\tau} f(r_t^a + r_t^b, D_{t,T_1}^a D_{t,T_1}^b, \ldots, D_{t,T_n}^a D_{t,T_n}^b)e^{-\int_{\tau=s}^{t} r_\tau^a d\tau}\right] \\
&= E_s\left[D_{s,t}^b f(r_t^a + r_t^b, D_{t,T_1}^a D_{t,T_1}^b, \ldots, D_{t,T_n}^a D_{t,T_n}^b)e^{-\int_{\tau=s}^{t} r_\tau^a d\tau}\right],
\end{aligned}
$$

which is the price of the option paying $D_{s,t}^b f(r_t^a + r_b, D_{t,T_1}^a D_{t,T_1}^b, \ldots, D_{t,T_n}^a D_{t,T_n}^b)$ at t under process a, and we are done. ∎

Because the interest rate process b is known and deterministic in Theorem 1, the option pricing result says that any standard option we can always price in model a can also be priced in model c. Often, the option payoff we are pricing will depend only on one of the arguments. For example, suppose that we want to know the price at s in model c of a call option with exercise price X maturing at t on a unit discount bond maturing at T. In this case,

$$
f(r_t^c, D_{t,T_1}^c, \ldots, D_{t,T_n}^c) = max(D_{t,T}^c - X, 0)
$$

and

$$
\begin{aligned}
D_{s,t}^b f(r_t^a + r_b, D_{t,T_1}^a D_{t,T_1}^b, \ldots, D_{t,T_n}^a D_{t,T_n}^b) &= D_{s,t}^b max(D_{t,T}^a D_{t,T}^b - X, 0) \\
&= D_{s,t}^b D_{t,T}^b max(D_{t,T}^a - X/D_{t,T}^b, 0) \\
&= D_{s,T}^b max(D_{t,T}^a - X/D_{t,T}^b, 0)
\end{aligned}
$$

which is $D_{s,T}^b$ times the price at s in model a of a call option with exercise price $X/D_{t,T}^b$ maturing at t on a unit discount bond maturing at T.

In other words, Theorem 1 allows us to perform the Ho and Lee type of analysis using a perturbation (by b) of any term structure model a that is convenient and reasonable to use. Furthermore, the option pricing in the original model a extends to the new model c. This

means that we can generate the interest rate process based on the initial yield curve and still use the closed-form bond and bond option pricing models including those derived by Vasicek [1977] and Cox, Ingersoll, and Ross [1985].

In a certain sense, all of these models in this section (including the Ho and Lee model) are *ad hoc* and should be treated as temporary fixes to be used only until we are successful at building and testing more realistic multifactor models of the term structure. The unreasonable feature of these models is that in each period we change our whole model (the parameters) in response to moves in the term structure. This is like comparative statics results that assume each period's change is interpreted as a once-and-for-all change of the sort that can never happen again. This is not rational and is not reasonable. Imperfectly correlated moves in all forward rates should be a feature of the stochastic structure of the model, and ideally we should have enough flexibility from the factor structure of a to eliminate entirely the need for a "fudge factor" b. On the other hand, as a practical matter, if a model with a few factors explains almost all of the variance of interest rate movements, the Ho and Lee model or a model developed in this section will give a good approximation for pricing many bond options even if it ignores small factors.[5]

It is worth noting at this point that while we have focused on the question of fitting today's term structure of interest rates using a translation of interest rates, it is also easy to fit today's term structure of interest rate volatilities using a change of time. The change of time would suffer from the same conceptual problems we have discussed for the translation of interest rates.

The next section discusses a class of multi-factor Vasicek models that overcome the somewhat *ad hoc* nature of the Ho and Lee approach. Using that general model as a starting point, it is explored whether the Ho and Lee approach, a generalization discussed in this section, or a more general model best fits the data.

4 Multifactor Models: Empirical Exploration

In the Ho and Lee model and the extensions discussed in Section 3, we have the *ad hoc* feature that many changes in the yield curve over time are interpreted as parameter changes rather than changes in state variables. In other words, what happens again and again each period is a realization that is completely outside the model, and yet we still retain the same model. Having an assumed term structure model that is so completely inconsistent with the data every period (because the realization is outside of the support of the probability distribution) makes it hard

[5] It is an interesting intellectual question, and given widespread use of these models in practice an interesting practical question, to understand what options are badly priced by these models.

to put much faith in these models.

There are different approaches one might take to resolving this problem. One is to assume that the observed term structure is not the correct term structure, and that any deviations from our model are the result of measurement error. This interpretation is not really consistent with the Ho and Lee approach, because it says that rather than accomodate today's term structure (as in Ho and Lee's approach) we ought to use it as a guide and base our pricing on a nearby term structure conforming to the model instead.

In general, basing a pricing model primarily on today's term structure cannot be a very good approach if measurement error (for example from not knowing how to interpret the bid-ask spread, non-contemporaneous quotes, or the difficulty of inferring discount bond prices from the prices of coupon bonds in the presence of tax effects) makes it difficult to measure today's term structure.

A more palatable approach, which is the approach taken in this section, is to use a term structure model that is rich enough to admit all possible movements in the term structure. To illustrate this approach, we will use a many-factor model in the style of Vasicek [1977]. Basically, Vasicek showed that we can do bond and bond option pricing whenever the interest rate process is Gaussian (in the risk-neutral probabilities).[6] This is because joint normally distributed interest rates imply that the exponent in the valuation formula ((1), (2), or (3)) is normally distributed and therefore discount bond prices can be computed using the normal moment generating function. Also, the joint lognormality of discount bond prices and the discount factor implies that pricing European puts and calls on discount bonds involves what is basically a transformation of the Black-Scholes model. Depending on how one does this derivation, it can be viewed as arising from Merton [1973], Rubinstein [1976], or Vasicek [1977].

A simple implication of the assumption of a Gaussian interest rate process is that discount bond yields also follow a jointly Gaussian process. Predictions in a Gaussian world are linear with homoskedastic Gaussian errors: joint normality of yields across times and maturities therefore follows from the valuation formula. We use the simple discrete-time valuation formula (3), and we assume that there is a finite vector of Markov state variables. More specifically, we assume that we can take the state variables at time i to be the yields to maturity of the discount bonds maturing at times $i+1$ through $i+n$. (Or, equivalently, these yields together are an invertible function of the state variables.)

The *log discount* at s on a discount bond maturing at a later time t is given by

$$d_{s,t} \equiv -\log(D_{s,t}). \tag{23}$$

[6] This is discussed in Dybvig [1988].

We will assume that innovations in the vector $d_s \equiv (d_{s,s+1}, \ldots, d_{s,s+n})$ have a covariance matrix Σ. If we apply (3) for $t = s+1$ and price a discount bond maturing at $T \geq t$, then dividing by $D_{s,s+1}$ we have that

$$E_s[D_{s+1,T}] = \frac{D_{s,T}}{D_{s,s+1}} \tag{24}$$

$$= e^{-d_{s,T}+d_{s,s+1}}. \tag{25}$$

By joint normality of the log discounts, the value of (24) is given by the normal moment generating function as $\exp(-E_s[d_{s+1,T}] + \mathrm{var}_s[d_{s+1,T}]/2)$, where $\mathrm{var}_s[\cdot]$ indicates variance conditional on information available at s. Consequently,

$$E_s[d_{s+1,T}] = d_{s,T} - d_{s,s+1} + \frac{1}{2}\mathrm{var}_s[d_{s+1,T}]. \tag{26}$$

Because the discount factors are joint Gaussian and Markov, $\mathrm{var}_s[d_{s+1,T}]$ is a constant that depends only on $T-(s+1)$. Of course, $E_s[d_{s+1,T}]$ varies depending on the value of the state variables at t.

What we have learned in the previous paragraph is that we do not have to estimate $E_s[d_{s+1,T}]$ and $\mathrm{var}_s[d_{s+1,T}]$ separately. In fact, this is essential (in the absence of a strong assumption such as the local expectations hypothesis), because what we observe empirically is the mean and variance under actual probabilities, not under the risk-neutral probabilities. For short time periods, variances of prediction errors under the risk-neutral probabilities are the same as under the actual probabilities–in other words, the change of measure affects the local mean, not the local variance.[7] This is a generalization of the fact that the actual mean does not affect asset pricing in the Black-Scholes model. We will take this result as accurate for the actual periods we use, which seems reasonable for monthly data and perhaps less reasonable for annual data. Roughly speaking, this means that when we run regressions on the data, we will ignore the estimated regression coefficients and intercept, and that we will base our asset pricing model on the covariance matrix of the residuals.

If our data always go n periods out, the only indeterminacy will be for $E_s[d_{s+1,s+1+n}]$, which corresponds to the last bond traded at $s+1$, which is not traded at s. However, this is not an issue provided we are finding the price at s of bonds maturing at or before $s+n$, or options on those bonds. Those bonds are all priced at s: by assumption, their prices are our state variables (up to conversion from yields to prices). Furthermore, (26) implies that the distribution of their movement over time does not depend on knowing the process for $d_{s,s+n+1}$, and by (3) neither does valuation of their options. Therefore, valuation of options on assets maturing within the

[7] See Harrison and Pliska [1981]. Dybvig and Huang [1988] is an example of a more modern treatment.

horizon of observed short rates does not require us to pin down the process on $d_{s,s+n+1}$. If we want to extrapolate and price assets and options beyond this horizon, there are reasonable ways to choose $E_s[d_{s+1,T}]$ (for example by using the estimated process in actual probabilities or making some assumption about risk premia), but within this paper it will be assumed that we are concerned with pricing options on bonds maturing within our observed maturity structure.

Now, the strategy is to estimate the covariance matrix of the innovations in log discounts. Then, (26) will tell us the means under the risk-neutral probabilities. This analysis is carried out using two data sets provided by the Center for Research in Security Prices (CRSP) at the University of Chicago. One data set contains yields from the 12-month version of the Fama Treasury Bill Term Structure file. The file used here generates the term structure for 1 to 12 months out using the average-of-bid-and-ask price of T-Bills of those approximate maturities. To avoid missing data, we use maturities from 1 through 9 months over the period June, 1964 through December, 1994. The other data set is the Fama-Bliss file of derived prices for discount bonds with maturities 1 through 5 years. This file is based on an elaborate selection procedure using bonds near par that do not have any special tax treatment. The second data set has monthly prices; from them was derived an annual series for December, 1952 through December, 1994, and a monthly series for the same dates as the Fama Treasury Bill Term Structure file. The time series of one-month interest rates and their innovations (as measured in the vector auto-regression below) are shown in Figure 1.

The first step in using the data was computation of the discount factors $d_{s,t}$, which was straightforward given the form of the data. In doing so, most of the Fama-Bliss data was discarded to convert it to a non-overlapping annual series. In each data set, the innovations in the discount factors were computed as the residuals of first-order vector autoregressions. These regressions are shown in Tables 1 and 4. While individual regression coefficients are not significant in these regressions (because interest rates tend to move together), the large R^2 and F values indicate that the regressions are very significant overall.[8] Including all the lagged discount factors should tend to minimize the effect of any errors-in-variables on the residuals without a significant loss in degrees of freedom. (Some possible sources of errors-in-variables include quote errors, the handling of the bid-ask spread, yields taken from bonds not exactly one month out, tax effects, and the selection procedure in the Fama-Bliss data.)

The estimated variances and covariances of innovations in log discounts are reported in

[8] Chris Lamoureux has pointed out that adding a second lag to these regressions seems to add significantly to the explanatory power, which suggests that the particular log discounts being used may not span a full set of state variables, depending on whether the additional explanatory power is due to a factor or to a conditional risk premium. This merits additional attention.

Tables 2 and 5. To better understand these covariance matrices, each was decomposed into principal components. The principal components model takes each eigenvector (normalized to Euclidian length 1) of the covariance matrix to be an independent component of the covariance matrix, and its corresponding eigenvalue measures the amount of variance explained by the component. (Essentially, principal components analysis is like factor analysis without the idiosyncratic noise. See Gnanadesikan [1977] for a simple description of the principal components technique and a discussion of its relation to factor analysis.) The results of the principal components analysis are striking. For both monthly and annual data, there is a dominant component corresponding to co-movements throughout the yield curve. In the monthly data, the movement corresponds roughly to parallel movements in the yield curve, while the annual data shows less response in the forward rates at higher maturities (note that the forward rates are differences in the log discounts). Given that the first component explains almost the entire variability in the data, this factor will contribute the dominant variance effect in bond and bond option prices.[9]

To test whether the same factor is being measured in both data sets, a combined monthly data set is formed for the same period as the monthly regression (June, 1964 through December, 1994). The vector autoregression now regresses each maturity's log discount on all 14 log discounts from the previous month. The actual regression is similar in flavor to the other two and the actual coefficients are not reported. Table 7 reports the dominant components of the principal components decomposition of the residuals. This decomposition, which again has a totally dominant component, confirms that the dominant component is the same in both cases.

The shape of the principal component in Table 7 indicates that it may be possible to explain almost all the variation in the data using a simple one-factor model. Next to the driftless model, the simplest Vasicek model is the simple mean-reverting model for which

$$r_t = r_{t-1} + \kappa(\bar{r} - r_{t-1}) + \sigma\eta_t. \tag{27}$$

In this model,

$$r_t = \bar{r} + (1-\kappa)^t(r_0 - \bar{r}) + \sigma\sum_{\tau=1}^{t}(1-\kappa)^{t-\tau}\eta_\tau, \tag{28}$$

[9]It is possible (and equivalent theoretically) to study the covariance matrix of some linear combination of the log discounts instead of the log discounts themselves. In theory, this would change somewhat the principal components decomposition. In fact, the alternatives still yield a dominant component, with an eigenvector corresponding to comovements. Furthermore, using the log discounts is probably less sensitive to various sources of errors-in-variables. Using forward rates puts lots of weight in little bumps in the yield curve, and using yields puts too much weight on the shortest maturity (whose yield is most affected by data problems such as inaccurate quotes, the bid-ask spread, or deviation of settlement or bond payments from the nominal date.

and it is easy to show that

$$\begin{aligned} d_{s,t} - E_{s-1}[d_{s,t}] &= \sigma \sum_{\tau=s+1}^{t} (1-\kappa)^{t-(s+1)} \eta_s \\ &= \sigma \frac{1-(1-\kappa)^{t-s}}{\kappa} \eta_s. \end{aligned} \tag{29}$$

is the innovation in the log discount.

In terms of the empirics, this model would have a single component for which the eigenvector times the square root of the eigenvalue would equal

$$\frac{1-(1-\kappa)^{t-s}}{\kappa} \sigma \sqrt{\Delta}, \tag{30}$$

where $t-s$ is the maturity (1–9 months and 1–5 years in our sample) and Δ is the time interval over which we measure the change. For estimation, we use the continuous version of the model (as described for example in Dybvig [1988]), for which largest eigenvalue can be written as[10]

$$\frac{1-\exp(-\kappa(t-s))}{\kappa} \sigma \sqrt{\Delta}. \tag{31}$$

Figure 2 shows the nonlinear-least-squares fit of the dominant eigenvalue times the square root of the eigenvector to (30), and the line with the same variance but no mean reversion. These fitted equations have parameters $\kappa = 0.0126$ (standard error 0.016) and $\sigma = 0.01848$ (standard error 0.00087).[11] Apparently, this parsimonious representation has a very good fit at all maturities, and the even more parsimonious representation with $\kappa = 0$ fits very well up to a year. (Remember that this means it should do well for options on instruments maturing in less than a year.) This says that the original Ho and Lee procedure should be a good fit for options on very short assets, and the extension of Ho and Lee should be a good fit for options on longer assets.

The consistency and simplicity of the empirical results are evidence in favor of the integrity of the data sets. If the data sets contained significant errors in variables, these errors would show up as significant components in the principal components analysis. Furthermore, the nice fit in Figure 2 would be unlikely to show up.

One implication of the empirical work is that conditional on the information at a given point in time, the next period's innovations in log discounts are almost perfectly correlated. This suggests that any second factor will not be an additive second factor (such as those in

[10] Note that this is the same as (30) up to a change in parameter definitions.

[11] The estimated standard errors are based on sampling variation in the error term but not in the parameter estimates of the underlying regression, which might be unstable due to a unit root in sample.

Cox, Ingersoll, and Ross [1985] or Brennan and Schwartz [1979]). Rather, the second factor is more likely to be a variance factor which has a small effect on bond pricing (and almost none at short maturities) but perhaps a significant effect on bond option pricing.

5 Conclusion

The Ho-Lee model and its descendents (such as popular implementations of Heath, Jarrow, and Morton [1992]) are conceptually flawed because they treat all but one dimension of local shocks as complete surprises. This conceptual problem is corrected in a general model with many factors, so that all possible local movements in interest rates are possible.

Empirically, this problem may be less severe than one might expect, at least for most common options not specifically designed to exploit this weakness. The principal components analysis in this paper, although exploratory and not backed up by formal tests, indicates that there is a single factor that explains almost all variability in the term structure, which suggests that treating all but a one-dimensional space of price movements as complete surprises may do reasonably well in pricing most options (since the amount of volatility being neglected is small and is probably less important than estimation error and other sources of misspecification), although it is undoubtedly possible to identify options (e.g. on spreads) for which the error is large.

The analysis here assumes but does not imply that the dominant factor has a constant variance. The analysis of this paper suggests (as did Brown and Dybvig [1986]) that the second factor (if any) in a term structure model should be related to the variance or other distributional features of interest rates, not additive in levels of interest rates as is usually assumed. Subsequent research should integrate a study of the variance process.

$d_{s,s+i} = b_0 + \sum_{j=1}^{9} b_j d_{s-1,s-1+j} + \varepsilon_{s,s+i}$ monthly 6406–9412									
Independent	Dependent variables								
variables	$d_{s,s+1}$	$d_{s,s+2}$	$d_{s,s+3}$	$d_{s,s+4}$	$d_{s,s+5}$	$d_{s,s+6}$	$d_{s,s+7}$	$d_{s,s+8}$	$d_{s,s+9}$
intercept	0.00	0.02	0.02	0.03	0.05	0.06	0.08	0.11	0.14
$\times 100$	(-0.08)	(0.95)	(0.63)	(0.92)	(1.23)	(1.19)	(1.44)	(1.80)	(1.99)
$d_{s-1,s}$	0.03	0.05	0.29	0.59	0.88	0.98	1.18	1.36	1.49
	(0.24)	(0.24)	(0.90)	(1.40)	(1.68)	(1.61)	(1.66)	(1.68)	(1.61)
$d_{s-1,s+1}$	0.61	0.42	0.51	0.45	0.40	0.53	0.49	0.62	0.81
	(4.73)	(1.80)	(1.45)	(0.97)	(0.70)	(0.79)	(0.64)	(0.70)	(0.80)
$d_{s-1,s+2}$	-0.33	0.16	-0.29	-0.35	-0.59	-0.92	-1.12	-1.37	-1.67
	(-2.28)	(0.63)	(-0.73)	(-0.68)	(-0.92)	(-1.23)	(-1.30)	(-1.39)	(-1.49)
$d_{s-1,s+3}$	0.22	0.26	0.68	0.42	0.59	0.89	0.87	1.05	0.99
	(1.89)	(1.23)	(2.13)	(1.00)	(1.15)	(1.47)	(1.23)	(1.31)	(1.09)
$d_{s-1,s+4}$	-0.13	-0.23	-0.39	-0.10	-0.82	-1.08	-1.19	-1.40	-1.53
	(-1.36)	(-1.40)	(-1.54)	(-0.30)	(-2.00)	(-2.24)	(-2.14)	(-2.19)	(-2.11)
$d_{s-1,s+5}$	-0.11	-0.18	-0.17	-0.16	0.56	0.33	0.59	0.68	0.92
	(-1.23)	(-1.10)	(-0.69)	(-0.48)	(1.40)	(0.69)	(1.07)	(1.08)	(1.28)
$d_{s-1,s+6}$	0.04	0.02	0.03	-0.09	-0.21	-0.15	-0.68	-0.95	-1.23
	(0.42)	(0.11)	(0.11)	(-0.24)	(-0.46)	(-0.29)	(-1.12)	(-1.37)	(-1.56)
$d_{s-1,s+7}$	0.33	0.65	0.98	1.44	1.76	2.17	2.84	2.86	3.07
	(3.59)	(3.93)	(3.94)	(4.37)	(4.35)	(4.57)	(5.16)	(4.55)	(4.28)
$d_{s-1,s+8}$	-0.20	-0.39	-0.59	-0.84	-1.02	-1.09	-1.21	-0.90	-0.77
	(-3.43)	(-3.76)	(-3.71)	(-4.03)	(-3.98)	(-3.60)	(-3.46)	(-2.25)	(-1.69)
$\overline{R}^2$	0.94	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
$F(9, 357)$	596.59	754.07	765.00	782.14	810.08	840.88	843.42	836.59	816.62

Table 1: **First-order vector autoregression of monthly log discounts.** This table reports the first-order vector autoregressions of the first nine monthly log discounts for the sample period 6406–9412. The regression uses the Fed version of the Fama T-Bill data provided by CRSP. The numbers in parentheses are t-statistics; these numbers are low because the discount rates are nearly multicollinear. This is no problem for the paper because the regressions are very significant overall (see the F- statistics) and because we are interested primarily in the residuals.

$\text{cov}[\varepsilon_{s,s+1},\ldots,\varepsilon_{s,s+9}]$ monthly 6406–9412									
$\times 1e5$	$\varepsilon_{s,s+1}$	$\varepsilon_{s,s+2}$	$\varepsilon_{s,s+3}$	$\varepsilon_{s,s+4}$	$\varepsilon_{s,s+5}$	$\varepsilon_{s,s+6}$	$\varepsilon_{s,s+7}$	$\varepsilon_{s,s+8}$	$\varepsilon_{s,s+9}$
$\varepsilon_{s,s+1}$	0.03	0.05	0.07	0.09	0.11	0.12	0.14	0.16	0.18
$\varepsilon_{s,s+2}$	0.05	0.10	0.15	0.19	0.23	0.26	0.30	0.34	0.38
$\varepsilon_{s,s+3}$	0.07	0.15	0.23	0.30	0.36	0.42	0.48	0.54	0.61
$\varepsilon_{s,s+4}$	0.09	0.19	0.30	0.40	0.49	0.57	0.65	0.73	0.82
$\varepsilon_{s,s+5}$	0.11	0.23	0.36	0.49	0.61	0.70	0.81	0.92	1.04
$\varepsilon_{s,s+6}$	0.12	0.26	0.42	0.57	0.70	0.84	0.96	1.09	1.24
$\varepsilon_{s,s+7}$	0.14	0.30	0.48	0.65	0.81	0.96	1.12	1.28	1.45
$\varepsilon_{s,s+8}$	0.16	0.34	0.54	0.73	0.92	1.09	1.28	1.47	1.66
$\varepsilon_{s,s+9}$	0.18	0.38	0.61	0.82	1.04	1.24	1.45	1.66	1.90

Table 2: **Covariance matrix of residuals from the regression reported in Table 1.** Each entry is $\times 1e5$; for example the first entry 0.03 really represents a variance of 3×10^{-7}. The diagonal elements of this matrix are used in the bond and bond option pricing. Small numbers imply that the effect on bond pricing is very small at these maturities, but of course the effect on bond option pricing is significant because the numbers here represent the only source of variance.

Principal Components of $\varepsilon_{s,s+1},\ldots,\varepsilon_{s,s+9}$ monthly 6406–9412									
eigenvalues of $\text{cov}[\varepsilon_{s,s+1},\ldots,\varepsilon_{s,s+9}]$ $\times 1e5$									
	6.56	0.09	0.02	0.01	0.01	0.01	0.00	0.00	0.00
maturity	corresponding normalized eigenvectors (as columns)								
1 month	0.05	−0.24	0.47	0.17	0.40	−0.39	0.00	−0.62	0.05
2 months	0.11	−0.36	0.45	0.26	0.18	−0.01	−0.11	0.66	−0.33
3 months	0.18	−0.44	0.20	0.04	−0.21	0.41	0.13	0.02	0.71
4 months	0.24	−0.45	−0.02	−0.22	−0.41	0.21	0.14	−0.34	−0.58
5 months	0.30	−0.31	−0.26	−0.38	−0.14	−0.68	−0.19	0.21	0.20
6 months	0.36	−0.16	−0.42	−0.09	0.64	0.37	−0.33	−0.07	−0.03
7 months	0.41	0.05	−0.30	0.39	0.12	−0.16	0.73	0.07	−0.02
8 months	0.47	0.25	0.00	0.55	−0.38	−0.03	−0.50	−0.14	0.01
9 months	0.53	0.47	0.46	−0.50	0.07	0.09	0.12	0.06	0.01

Table 3: **Principal Components decomposition of the covariance matrix of residuals from the regression reported in Table 1.** The dominant eigenvector corresponds to roughly parallel shifts in the yield curve (exactly parallel shifts would give an eigenvector proportional to $(1, 2, \ldots, 9)$). The second eigenvector corresponds to changes the overall slope of the yield curve. The third eigenvector corresponds to changes of curvature in the yield curve. Subsequent eigenvectors seem hard to interpret. The first eigenvalue is dominant, at 70 times the size of the next eigenvalue. Furthermore, the third and subsequent eigenvectors are even less important because their associated eigenvalues are so small: the largest in this group is less than $1/300^{th}$ of the largest eigenvalue.

$d_{s,s+i} = b_0 + \sum_{j=1}^{9} b_j d_{s-1,s-1+j} + \varepsilon_{s,s+i}$ annual 52–94					
Independent	Dependent variables				
variables	$d_{s,s+1}$	$d_{s,s+2}$	$d_{s,s+3}$	$d_{s,s+4}$	$d_{s,s+5}$
intercept	1.24	2.26	3.64	4.51	5.79
$\times 100$	(1.90)	(1.92)	(2.28)	(2.25)	(2.40)
$d_{s-1,s}$	1.50	2.52	3.34	4.48	4.12
	(1.08)	(1.01)	(0.99)	(1.06)	(0.81)
$d_{s-1,s+1}$	1.20	2.22	3.21	3.78	5.23
	(0.86)	(0.89)	(0.95)	(0.89)	(1.02)
$d_{s-1,s+2}$	−1.66	−3.32	−5.35	−7.00	−8.56
	(−1.83)	(−2.04)	(−2.42)	(−2.53)	(−2.57)
$d_{s-1,s+3}$	−0.65	−0.93	−0.52	−0.69	−0.64
	(−0.82)	(−0.65)	(−0.27)	(−0.28)	(−0.22)
$d_{s-1,s+4}$	0.90	1.69	2.18	3.04	3.60
	(1.85)	(1.93)	(1.85)	(2.05)	(2.02)
$\overline{R}^2$	0.73	0.78	0.81	0.83	0.84
$F(5, 37)$	24.08	30.23	37.07	43.00	46.09

Table 4: **First-order vector autoregression of annual log discounts.** These results are similar to the monthly autoregression in Table 1. As in that regression, near multicollinearity makes the individual t-statistics insignificant, but each regression is significant overall as indicated by the F-statistics.

$\text{cov}[\varepsilon_{s,s+1},\ldots,\varepsilon_{s,s+9}]$ monthly 6406–9412									
$\times 1e5$	$\varepsilon_{s,s+1}$	$\varepsilon_{s,s+2}$	$\varepsilon_{s,s+3}$	$\varepsilon_{s,s+4}$	$\varepsilon_{s,s+5}$	$\varepsilon_{s,s+6}$	$\varepsilon_{s,s+7}$	$\varepsilon_{s,s+8}$	$\varepsilon_{s,s+9}$
$\varepsilon_{s,s+1}$	0.03	0.05	0.07	0.09	0.11	0.12	0.14	0.16	0.18
$\varepsilon_{s,s+2}$	0.05	0.10	0.15	0.19	0.23	0.26	0.30	0.34	0.38
$\varepsilon_{s,s+3}$	0.07	0.15	0.23	0.30	0.36	0.42	0.48	0.54	0.61
$\varepsilon_{s,s+4}$	0.09	0.19	0.30	0.40	0.49	0.57	0.65	0.73	0.82
$\varepsilon_{s,s+5}$	0.11	0.23	0.36	0.49	0.61	0.70	0.81	0.92	1.04
$\varepsilon_{s,s+6}$	0.12	0.26	0.42	0.57	0.70	0.84	0.96	1.09	1.24
$\varepsilon_{s,s+7}$	0.14	0.30	0.48	0.65	0.81	0.96	1.12	1.28	1.45
$\varepsilon_{s,s+8}$	0.16	0.34	0.54	0.73	0.92	1.09	1.28	1.47	1.66
$\varepsilon_{s,s+9}$	0.18	0.38	0.61	0.82	1.04	1.24	1.45	1.66	1.90

Table 5: **Covariance matrix of residuals from the regression reported in Table 4.** This is similar to Table 2.

Principal Components of $\varepsilon_{s,s+1},\ldots,\varepsilon_{s,s+5}$ annual 52–94					
eigenvalues of $\text{cov}[\varepsilon_{s,s+1},\ldots,\varepsilon_{s,s+5}]$ $\times 1e5$					
	628.84	5.98	0.63	0.43	0.11
maturity	corresponding normalized eigenvectors (as columns)				
1 year	0.17	−0.53	−0.48	−0.30	0.61
2 year	0.31	−0.54	−0.25	0.04	−0.74
3 year	0.42	−0.36	0.55	0.56	0.29
4 year	0.53	0.12	0.43	−0.72	−0.06
5 year	0.64	0.53	−0.47	0.28	0.06

Table 6: **Principal Components decomposition of the covariance matrix of residuals from the regression reported in Table 4.** The dominant eigenvector corresponds to co-movement of all interest rates, althought this movement no longer looks parallel as in the monthly case. This shows evidence of mean reversion in risk-neutral probabilities (exactly parallel shifts would give an eigenvector proportional to $(1,2,\ldots,5)$). As in the monthly data, the second eigenvector corresponds to changes the overall slope of the yield curve, the third eigenvector corresponds to changes of curvature in the yield curve, and subsequent eigenvectors seem hard to interpret. In this data, the largest eigenvector is even more dominant than in the monthly data: it is more than 100 times larger than the next largest eigenvalue.

Largest Principal Components of $\varepsilon_{s,s+i}'s$ combined 6406–9412						
eigenvalues of cov$[\varepsilon_{s,s+i}'s]$ $\times 1e5$						
	101.85	3.16	0.95	0.49	0.24	0.10
maturity	corresponding normalized eigenvectors (as columns)					
1 month	0.01	−0.05	−0.01	0.05	0.05	0.04
2 month	0.02	−0.10	−0.02	0.09	0.10	0.13
3 month	0.04	−0.15	−0.04	0.12	0.16	0.20
4 month	0.05	−0.19	−0.06	0.15	0.17	0.28
5 month	0.06	−0.23	−0.07	0.16	0.15	0.23
6 month	0.08	−0.26	−0.06	0.15	0.15	0.17
7 month	0.09	−0.28	−0.06	0.16	0.13	0.12
8 month	0.11	−0.31	−0.05	0.17	0.12	0.06
9 month	0.12	−0.34	−0.07	0.15	0.11	−0.05
12 month	0.16	−0.38	−0.07	0.10	0.01	−0.84
24 month	0.30	−0.34	−0.03	−0.13	−0.85	0.21
36 month	0.42	−0.17	0.17	−0.80	0.35	0.06
48 month	0.53	0.22	0.72	0.38	0.01	−0.01
60 month	0.62	0.43	−0.64	0.12	0.05	−0.01

Table 7: **Principal Components decomposition of the covariance matrix of residuals from combined monthly and annual data.** This is the decomposition (as in Table 3 or 6) of the residuals from the log discount autoregression (as in Table 1 or 4) using monthly observations of combined monthly and annual data. This table offers evidence that the dominant factor in the annual and monthly regressions are in fact the same, because essentially the same factor shows up here. In fact, the second and third factors (which are again much smaller) again can be interpreted as slope and curvature of the yield curve.

Figure 1

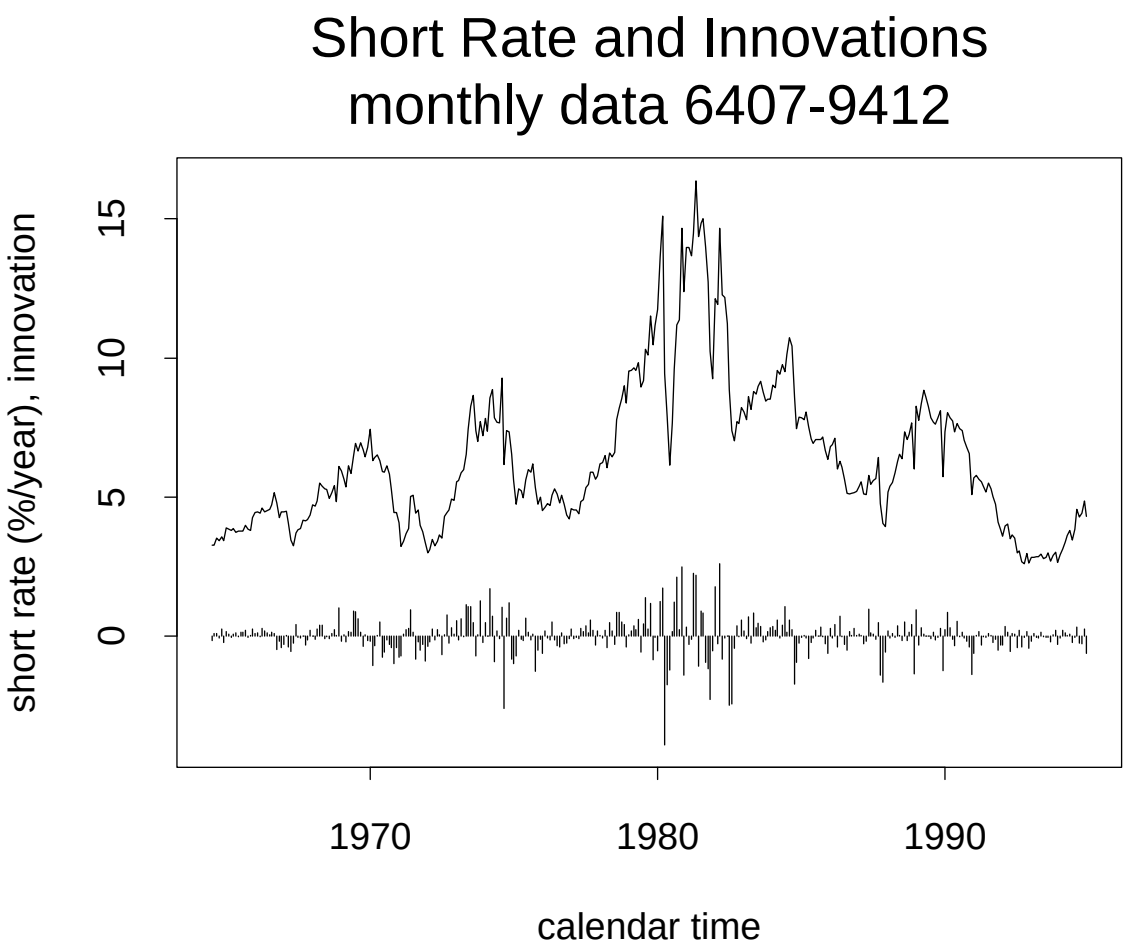
Short Rate and Innovations
monthly data 6407-9412
short rate (%/year), innovation
0
5
10
15
1970
1980
1990
calendar time

Figure 2

The points indicate the computed values of the largest eigenvector. The straight line is the fitted theoretical curve without mean reversion, and the curved line is the fitted theoretical curve with mean reversion. The estimated value of kappa corresponds to mean reversion of 12.8% per year, with sampling error of at least 1.66% per year.

References

Brennan, Michael J., and Eduardo S. Schwartz, 1979, "A Continuous Time Approach to the Pricing of Bonds" *Journal of Banking and Finance* **3**, 133–155.

Brown, Stephen and Philip Dybvig, 1986, The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates, *Journal of Finance* **41**, 617–630.

CRSP (Center for Research in Security Prices), 1995, 1994 Bond Files Guide, Graduate School of Business, University of Chicago.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1981, A Re-examination of Traditional Hypotheses about the Term Structure of Interest Rates, *Journal of Finance* **36**, 769–799.

Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, A Theory of the Term Structure of Interest Rates, *Econometrica* **53**, 385–407.

Cox, John C., and Stephen Ross, 1976, A Survey of Some New Results in Financial Option Pricing Theory, *Journal of Finance* **21**, 383–402.

Cox, John, Stephen Ross, and Mark Rubinstein, 1979, Option Pricing: A Simplified Approach, *Journal of Financial Economics* **7**, 229–263.

Dybvig, Philip H., 1988, Inefficient Dynamic Portfolio Strategies, or How to Throw Away a Million Dollars in the Stock Market, *Review of Financial Studies* **1**, 67–88.

Dybvig, Philip H., and Chi-fu Huang, 1988, Nonnegative Wealth, Absence of Arbitrage, and Feasible Consumption Streams, *Review of Financial Studies* **1**, 377–401.

Dybvig, Philip H., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, forthcoming, Long Forward and Discount Rates Can Never Fall, *Journal of Business*.

Gnanadesikan, 1977, *Methods for Multivariate Data Analysis of Multivariate Observations*, New York: Wiley.

Harrison, Michael, and David Kreps, 1979, Martingales and multiperiod securities markets, *Journal of Economic Theory* **20**, pp.381–408.

Harrison, Michael, and Stanley Pliska, 1981, Martingales and stochastic integrals in the theory of continuous trading, *Stochastic Processes and Their Applications*, **11**, pp.215–260.

Heath, David, Robert Jarrow, and Andrew Morton, 1990, Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation, *Journal of Financial and Quantitative*

Analysis **25**, 419–440.

Heath, David, Robert Jarrow, and Andrew Morton, 1992, Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation, *Econometrica* **60**, 77–105.

Ho, Thomas S. Y., and Sang-Bin Lee, 1986, Term Structure Movements and Pricing Interest Rate Contingent Claims, *Journal of Finance* **41**, 1011–1029.

Merton, Robert C., 1973, Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* **4**, 141–183.

Rubinstein, Mark, 1976, The Valuation of Uncertain Income Streams and the Pricing of Options, *Bell Journal of Economics* **7**, 407–425.

Vasicek, Oldrich, 1977, An Equilibrium Characterization of the Term Structure *Journal of Financial Economics* **5**, 177–188.