European Finance Review **1:** 105–130, 1997.

The Dynamics of Short-Term Interest Rate Volatility Reconsidered

KEES G. KOEDIJK, FRANÇOIS G. J. A. NISSEN,
PETER C. SCHOTMAN and CHRISTIAN C. P. WOLFF
Limburg Institute of Financial Economics (LIFE)[⋆]

Abstract. In this paper we present and estimate a model of short-term interest rate volatility that encompasses both the level effect of Chan, Karolyi, Longstaff and Sanders (1992) and the conditional heteroskedasticity effect of the GARCH class of models. This flexible specification allows different effects to dominate as the level of the interest rate varies. We also investigate implications for the pricing of bond options. Our findings indicate that the inclusion of a volatility effect reduces the estimate of the level effect, and has option implications that differ significantly from the Chan, Karolyi, Longstaff and Sanders (1992) model.

1. Introduction

Chan, Karolyi, Longstaff and Sanders (1992, CKLS) compare a number of widely used continuous-time models of the short-term interest rate. They estimate various models and compare the models in terms of their ability to capture the actual behavior of the short-term riskless rate. The issue of how these models compare is important because the models differ in their implications for valuing contingent claims and hedging interest rate risk.[1] The testing approach of CKLS exploits the fact that many term structure models imply dynamics for the short-term riskless rate that can be nested in one stochastic differential equation. With respect to the most successful models they conclude: “The results for the tests of the one-month Treasury Bill indicate that it is critical to model volatility correctly. The models that best describe the dynamics of the interest rates over time are those that allow the conditiorial volatility of interest rate changes to be highly dependent on the

⋆ We are grateful to the Federal Reserve Bank of St. Louis for generously providing the data. We thank Dennis Bams, Stefano Cavaglia, Nai-Fu Chen, Jean-Marie Dufour, René Garcia, Christian Gouriéroux, Ravi Jagannathan, Andrew Karolyi, Clemens Kool, Jesper Lund, Franz Palm, Pierre Perron, Richard Stapleton and participants at the American Finance Association meetings (Washington 1995), the First National Taiwan University International Conference on Money and Finance (Taipei, 1994), European Finance Association meetings (Brussels, 1994), Econometric Society European meetings (Maastricht, 1994), conference of the Money Macroeconomics and Finance Group (London, 1994) and seminars at Lancaster University, University of Montreal, University of Madrid (Carlos III) and an anonymous referee, for helpful discussion and comments on an earlier draft of this paper. We are solely responsible for all errors. Correspondence: LIFE, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands.

level of the interest rate". With regard to the parameter that measures the sensitivity of interest rate volatility to the level of the interest rate itself (γ) they report an unconstrained estimate of about 1.5.[2]

A different class of models to capture volatility dynamics in interest rates is the family of autoregressive conditional heteroskedasticity (ARCH) models, introduced by Engle (1982) and generalized (GARCH) by Bollerslev (1986). Key ingredients in these models are volatility clustering and volatility persistence. These effects are usually reliably present in estimated GARCH models of interest rate time series.[3]

In the current paper we present and estimate a model of short-term interest rate volatility, which encompasses both the level effect of CKLS and the conditional heteroskedasticity effect of the GARCH class of models. Our model – the KNSW model – exhibits a superior empirical fit relative to both pure GARCH models as well as pure single factor models as considered in CKLS. This feature results from a relatively flexible specification which allows different effects to dominate as the level of the interest rate varies.

It has long been recognized in the finance literature that the specification of volatility is one of the most important features for derivative security pricing. We investigate implications for the pricing of bond options. Specifically, we investigate the implications from different models of the short-term interest rate for the pricing of discount bond options. Our findings indicate that the inclusion of a volatility effect in the model specification, in addition to a level effect, is particularly relevant for the pricing of shorter-term options on long term bonds. The magnitude of the implied price differences is strongly dependent on the level of the interest rate.

The plan of the paper is as follows. In Section 2 we briefly review previous studies that model short-term interest rates, and introduce our new specification which nests both the level effect of CKLS and the volatility effect of the GARCH class of models. Section 3 describes our data and contains the empirical results. Section 4 considers the implications for the pricing of contingent claims. Finally, Section 5 offers some concluding remarks.

2. GARCH and Level Effects

Most of the theoretical models of the short-term interest rate which are used in finance have been developed in a continuous time setting. CKLS review a number of widely used stochastic processes that are nested within the following stochastic differential equation:

$$\mathrm{d}r_t = \kappa(\mu - r_t)\,\mathrm{d}t + \sigma r_t^{\gamma}\,\mathrm{d}W_t, \tag{1}$$

where r_t represents the short-term interest rate and W_t is a standard Brownian motion. The interest rate process is mean reverting for $\kappa > 0$. The parameter γ determines the sensitivity of the variance with respect to the level of the spot rate; we will refer to γ as the interest rate elasticity. This parameter turns out to be crucial in applications to option valuation. CKLS approximate this stochastic differential equation by the following discrete approximation:

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \sigma r_{t-1}^{\gamma} \epsilon_t, \quad \epsilon_t \sim D(0,1), \tag{2}$$

which they estimate for the one-month US Treasury bill rate. The bill yield was obtained from the Fama files within the CRSP database. The data are monthly quotations for the period 1964.06–1989.12. They used the Generalized Method of Moments to estimate the model, and report that γ is 1.5 and highly significant, which means that the conditional variance of the short-term interest rate is highly sensitive to changes in the level of the interest rate. For comparison, the Cox, Ingersoll and Ross (1985) (CIR) square root term structure model implies $\gamma = \frac{1}{2}$. In fact, the estimated elasticity is so large that stationarity of the interest rate process is not guaranteed (see Broze, Scaillet and Zakoian (1995)).

Alternatively, Longstaff and Schwartz (1993) (LS) present a two factor model for the term structure. The first factor is the short-term interest rate. The second factor is the conditional variance of changes of the short-term interest rate, which is assumed to be generated by a GARCH-class process. In their application LS estimate the following model:

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 h_{t-1}^2 + e_t, \tag{3}$$

$$h_t^2 = \beta_1 + \beta_2 e_{t-1}^2 + \beta_3 h_{t-1}^2 + \beta_4 r_{t-1}, \tag{4}$$

where $e_t = h_t \epsilon_t$ is the prediction error of the interest rate and h_t^2 is the conditional variance. The error term ϵ_t is normalized to have unit variance. This specification differs from the standard GARCH-M model by the inclusion of the lagged spot rate in the volatility equation.[4] Note that, if $\beta_2 = 0$ and $\beta_3 = 0$, the specification corresponds to a model with $\gamma = \frac{1}{2}$.[5] The level effect in the volatility proves significant, but the restriction that the interest sensitivity is equal to $\gamma = \frac{1}{2}$ might well be overly restrictive given the unrestricted estimate of γ reported in CKLS. We therefore would like to generalize the Longstaff and Schwartz specification such that it can accommodate different interest rate sensitivities in the volatility. Stated differently, we search for a specification that combines the high interest rate sensitivity of CKLS with GARCH-type volatility clustering. This motivates the specification presented below. We assume conditional volatility of the form:

$$h_t = \sigma_t r_{t-1}^{\gamma}, \tag{5}$$

which differs from CKLS by the time varying nature of σ_t, which is assumed to be generated by the GARCH(1,1) process:

$$\sigma_t^2 = \beta_1 + \beta_2(\sigma_{t-1}\epsilon_{t-1})^2 + \beta_3 \sigma_{t-1}^2. \tag{6}$$

Substituting Equation (6) in Equation (5) and solving for σ_t^2, the proposed volatility process can be rewritten as:

$$h_t^2 = \beta_1 r_{t-1}^{2\gamma} + \left(\frac{r_{t-1}}{r_{t-2}}\right)^{2\gamma} (\beta_2 e_{t-1}^2 + \beta_3 h_{t-1}^2). \tag{7}$$

We will refer to this specification as the KNSW model.[6] The CKLS model is a special case of this model; for $\beta_2 = \beta_3 = 0$. Another special case is the GARCH model, which obtains if $\gamma = 0$. A restricted version of LS obtains, i.e. with $\alpha_2 = 0$, if $\gamma = \frac{1}{2}$. An interesting feature of the above specification is the time varying persistence of shocks which depends on the interest rate level.

The unconditional distribution of the spot rate is not available in closed form. It can however be easily computed numerically by simulation for different values of the parameters. In the discrete time process negative interest rates are possible, but this is simply an artefact of the discrete time approximation of a continuous time process. The probability that the short-term interest rate will ever attain negative values is extremely small when the unconditional mean and the drift are sufficiently large relative to h_t^2.

The interest rate process of CIR is stationary. Broze, Scaillet and Zakoian (1995) show that the Euler discretization of Equation (1) is only (second-order) stationary if $\gamma < 1$. If on the other hand, $\gamma \geq 1$ the volatility at high interest rates makes it possible for interest rates to increase even further. To allow for higher variance elasticities it is necessary to introduce nonlinearities in the drift (see Aït-Sahalia (1996b)). Conley, Hansen, Luttmer and Scheinkman (1995) consider extensions of the drift in (1) by including the nonlinear terms r^{-1} and r^2. We augment the conditional mean in the discrete time model with a quadratic term:

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1}^2 + e_t. \tag{8}$$

For $\alpha_2 < 0$ the stronger mean-reverting drift is now able to pull back the interest rate to the unconditional mean from a high interest rate level in the presence of a larger variance. At the low end we assume a reflecting barrier at r = 0, which has no consequences for the estimation.

An exact continuous time limit of the discrete time process in Equation (7) is not available. One could consider the process as a simple approximation to the following process, which is closely related to the LS model:

$$\mathrm{d}r_t = \kappa(\mu - r_t)\,\mathrm{d}t + \sigma r_t^\gamma\,\mathrm{d}W_{1t}, \tag{9}$$

$$\mathrm{d}\sigma_t^2 = \theta(w - \sigma_t^2)\,\mathrm{d}t + \psi\sigma_t\,\mathrm{d}W_{2t}, \tag{10}$$

The scale parameter σ_t is not a constant as in the models considered by CKLS but follows a diffusion process as in Nelson (1990). A major difference, however, is that γ is not restricted to $\gamma = \frac{1}{2}$ in Equation (9).

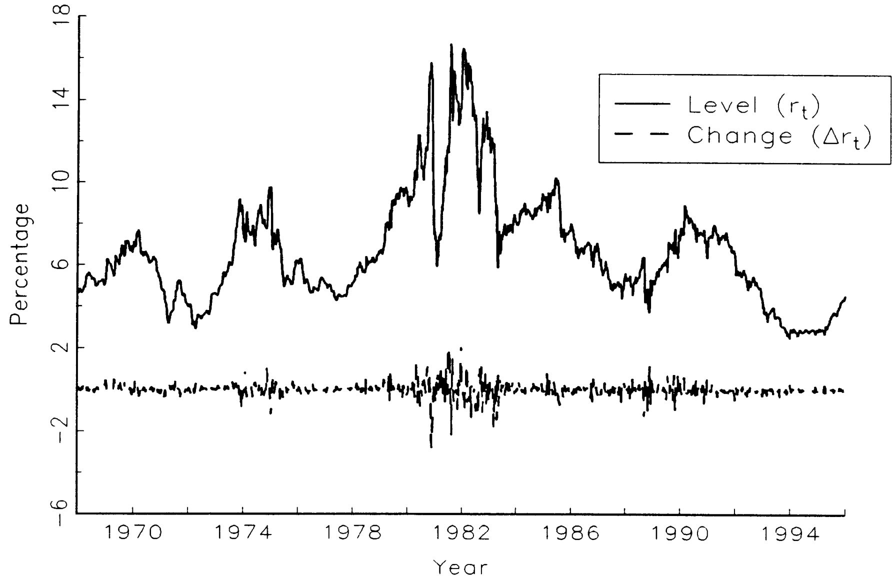

Figure 1. One-month Treasury bill.

3. Empirical Results

3.1. Data

The one-month Treasury bill rate is chosen as the short-term interest rate. Monthly and weekly yields were obtained from the Federal Reserve Bank for the period January 1968–July 1996. The monthly data are last Friday of the month observations and the weekly data are recorded at the last trading-day of the week.

The yield is expressed in annualized form. Figure 1 plots the level and the change of the one-month Treasury bill rate. The interest rate is more variable in the period subsequent to the 1979 change in Federal Reserve Bank operating procedures. Table I provides summary statistics for the two data series. The distribution of Δr_t is skewed to the left and exhibits excess kurtosis.

3.2. Parameter Estimates

In this section we provide estimation and test results for the four different models for the behavior of short-term interest rate volatility: the GARCH(1,1), CKLS, LS and KNSW. We consider the specification for the conditional mean in Equation (8) and the following volatility specification that nests the models of interest:

$$h_t^2 = \beta_0 + \beta_1 r_{t-1}^{2\gamma} + \left(\frac{r_t}{r_{t-1}}\right)^{2\gamma} (\beta_2 e_{t-1}^2 + \beta_3 h_{t-1}^2) + \beta_4 r_{t-1}. \tag{11}$$

Table I. Summary statistics

	Levels		First Differences	
	Monthly	Weekly	Monthly	Weekly
Number of observations	343	1440	342	1439
Mean	6.56	6.56	0.00	0.00
Standard deviation	2.62	2.62	0.68	0.33
Minimum	2.61	2.48	−6.10	−2.86
Maximum	15.85	16.69	3.40	2.31
Skewness	1.24	1.26	−1.73	−0.63
Excess kurtosis	1.82	2.00	22.77	14.04

Notes: Skewness is defined as m_3/s^3, with m_3 the centred third moment of the data and s the sample standard deviation. Kurtosis is defined as $(m_4/s^4) - 3$, with m_4 the centred fourth moment of the data. Units are percent per annum.

The models are estimated by the method of quasi-maximum likelihood (QML). The QML estimator is consistent and asymptotically normal for any distribution of e_t providing some regularity conditions are satisfied (see Wooldridge (1994)). The asymptotic distribution for the QML estimator $\hat{\theta}$ is then:

$$\sqrt{T}(\theta_0 - \hat{\theta}) \rightarrow N(0, \boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}), \tag{12}$$

where $\boldsymbol{A}$ denotes the information matrix, and $\boldsymbol{B}$ denotes the outer product of the gradient vector evaluated at the optimal parameter vector. The standard errors are estimated using the robust covariance matrix $\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}$.

Because of severe multicollinearity problems we are unable to estimate the general specification nesting all the specific models. It turns out that the extra constant term (β_0) of the LS model cannot be estimated if γ is a free parameter.

Panel A of Table II reports empirical results for the four models using monthly observations. The first column contains the estimates of the standard GARCH(1,1) model.[7] The GARCH parameters are highly significant even on the relatively low monthly frequency, and indicate strong persistence of variance shocks. The second and third columns report our estimates of the CKLS model in which the interest rate elasticity γ is included. The estimates of γ move from 1.40 to 2.51 when a constant term is added.[8,9] The LS model in the fourth column adds the lagged level of r_t to the GARCH specification, which appears an important improvement in terms of the likelihood function.[10] The inclusion of r_{t-1} also lowers the persistence of the variance shocks. The proposed KNSW specification is reported in the last column of Table II. It attains the highest value for the log-likelihood function. The GARCH and CKLS models are both nested within this specification, and can both be rejected at the 5% level. As in the LS model, the inclusion of the lagged interest rate lowers the persistence of the volatility shocks. Similar to the CKLS model we also find a large value for γ.

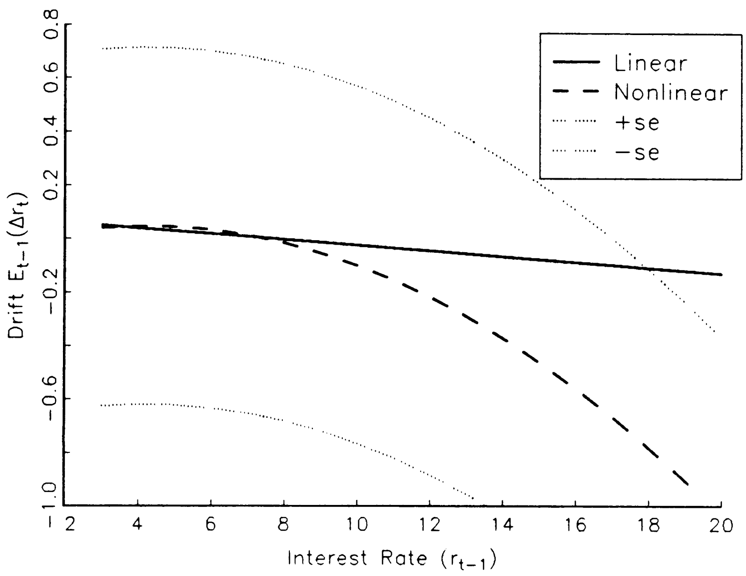

Figure 2. The figure contains the drift of the unrestricted (nonlinear) KNSW model and the (restricted) linear KNSW model. Se denotes the one-standard-error band for the nonlinear model. Table V contains the parameter estimates for the conditional mean of the linear KNSW model.

It has been illustrated by Drost and Nijman (1993) that GARCH effects are particulary dominant for high frequency data. We replicate, therefore, the estimates of Panel A using weekly observations. The parameter estimates for the weekly data are reported in Panel B. The higher frequency of the observation shows up in the parameters of the GARCH(1,1) model, where the point estimates β_2 and β_3 even add up to 1.03, although we can never reject the hypothesis that they add up to a number smaller than one. The interest rate sensitivity γ of the CKLS model is identical to the estimate from monthly data. Inclusion of the nonlinearity term α_2 has no influence on the estimated values of the parameters in the volatility specification. But the negative point estimates ensure stationarity even if $\gamma > 1$.

Figure 2 plots the lagged interest rate level against the drift, $\mathbf{E}_{t-1}[\Delta r_t]$, for the nonlinear KNSW model as well as for the linear KNSW model with $\alpha_2 = 0$. For moderate interest rate levels there is very slight mean reversion, however at interest rates higher than 15% the drift sharply decreases. Note however that the standard error of the drift term is quite large. Aït-Sahalia (1996a,b), Andersen and Lund (1996b), Conley, Hansen, Luttmer and Scheinkman (1995), Pfann, Schotman and Tschernig (1996), Stanton (1995) and Tauchen (1996) report similar nonlinearities in the dynamics of the short-term interest rate.

Table II. Unrestricted parameter estimates

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1}^2 + e_t$$

$$h_t^2 = \beta_0 + \beta_1 r_{t-1}^{2\gamma} + \left(\frac{r_t}{r_{t-1}}\right)^{2\gamma} (\beta_2 e_{t-1}^2 + \beta_3 h_{t-1}^2) + \beta_4 r_{t-1}.$$

	GARCH	CKLS1	CKLS2	LS	KNSW
A: Monthly:					
$\alpha_0 \times 10$	−0.25	0.01	0.45	−0.26	−0.30
	(0.15)	(0.07)	(0.73)	(0.09)	(0.30)
$\alpha_1 \times 10$	0.37	0.20	0.00	0.34	0.36
	(0.53)	(0.71)	(0.01)	(0.25)	(0.81)
$\alpha_2 \times 10^2$	−0.58	−0.26	−0.10	−0.42	−0.44
	(0.78)	(0.88)	(0.20)	(0.34)	(0.94)
$\beta_0 \times 10$	–	–	0.70	–	–
	–	–	(3.49)	–	–
$\beta_1 \times 10^2$	0.77	0.13	0.001	–	0.02
	(0.96)	(2.18)	(0.53)	–	(1.15)
β_2	0.26	–	–	0.25	0.18
	(3.89)	–	–	(2.31)	(3.08)
β_3	0.75	–	–	0.70	0.74
	(12.03)	–	–	(5.11)	(8.90)
$\beta_4 \times 10^2$	–	–	–	0.31	–
	–	–	–	(1.70)	–
γ	–	1.40	2.51	–	1.24
	–	(10.13)	(5.25)	–	(4.90)
Loglik	−214	−217	−209	−208	−198
B: Weekly:					
$\alpha_0 \times 10$	0.12	0.10	0.11	0.07	0.11
	(0.26)	(0.26)	(0.26)	(0.12)	(0.32)
$\alpha_1 \times 10$	0.01	0.01	0.01	0.03	0.01
	(0.04)	(0.28)	(0.18)	(0.13)	(0.09)
$\alpha_2 \times 10^2$	−0.01	−0.01	−0.01	−0.07	−0.03
	(0.26)	(0.62)	(0.48)	(0.30)	(0.25)
$\beta_0 \times 10$	–	–	0.10	–	–
	–	–	(1.83)	–	–
$\beta_1 \times 10^2$	0.01	0.08	0.01	–	0.01
	(1.62)	(2.82)	(0.78)	–	(1.81)
β_2	0.22	–	–	0.23	0.31
	(2.96)	–	–	(3.70)	(3.40)
β_3	0.81	–	–	0.77	0.60
	(14.05)	–	–	(13.79)	(3.81)
$\beta_4 \times 10^2$	–	–	–	0.04	–
	–	–	–	(2.29)	–

Table II. Continued

	GARCH	CKLS1	CKLS2	LS	KNSW
B: Weekly:					
γ	–	1.21	1.67	–	1.31
	–	(13.59)	(6.03)	–	(6.33)
Loglik	83	-76	-71	94	101

Notes: Loglik denotes the log-likelihood value. Robust t-values are given in parentheses. CKLS, LS and KNSW denote Chan, Karolyi, Longstaff and Sanders (1992), Longstaff and Schwartz (1992) and Koedijk, Nissen, Schotman and Wolff respectively.

3.3. Diagnostic Tests

In order to investigate the adequacy of the conditional variance model we employ a series of Lagrange multiplier (LM) tests as suggested by Bollerslev, Engle and Nelson (1994). With the LM tests we search for directions in which the model could be improved. Let z_t be a vector of explanatory variables that we like to test for inclusion in the volatility equation:

$$h_t^2 = f(x_t, \theta) + \delta z_t, \tag{13}$$

where $f(x_t, \theta)$ is the volatility specification under the null hypothesis. Given the fat-tailedness of the data the conventional LM tests, described in Engle (1984) are no longer applicable. However, from Wooldridge (1994) and Bollerslev and Wooldridge (1992) a robust LM test may be computed from a simple set of auxiliary regressions. First, run the regression from

$$\frac{\partial \ln h_t^2}{\partial \theta} \quad \text{on} \quad \frac{\partial \ln h_t^2}{\partial \delta} = z_t / \hat{h}_t^2$$

both evaluated at the QML estimates under the null hypothesis. Next, calculate the score that is orthogonal to the scores under the null hypothesis as $s_{\delta t} = (\hat{e}_t^2/\hat{h}_t^2 - 1)\hat{\nu}_{\delta t}$ where $\hat{e}_t$, the prediction errors, and $\hat{h}_t^2$ are evaluated at the QML estimates under the null hypothesis and $\hat{\nu}_{\delta t}$ are the residuals of the first regression. An asymptotically valid LM statistic is then calculated as the TR^2 from a regression of a vector of ones on $s_{\delta t}$ with T the number of observations and R^2 the uncentered multiple correlation coefficient. The test statistic is asymptotically distributed as $\chi^2(k)$, with k the number of elements in z_t. The directions of misspecification that we consider are:

1. Additional level effects: $z_t = r_{t-1}$.
2. Outliers between 79:10 and 81:12: $z_t = D_{1t}$ a dummy variable that takes the value one in the period 79:10–81:12 and is zero elsewhere.

3. A permanent variance shift after 79:10: $z_t = D_{2t}$, a dummy variable that takes the value one after 79:10, and zero elsewhere. This test can be interpreted as a test for the stability of the model over subperiods.
4. Sign bias (see Engle and Ng (1993)): $z_t = S_t^+$, a dummy variable that takes the value one if $e_{t-1} > 0$ and zero elsewhere.
5. Size bias (see Engle and Ng (1993)): $z_{1t} = S_t^+ e_{t-1}$, $z_{2t} = S_t^+ e_{t-1}^2$, $z_{3t} = S_t^- e_{t-1}$ and $z_{4t} = S_t^- e_{t-1}^2$ where $S_t^- = 1 - S_t^+$.

All these tests indicate directions in which to search for improved specifications. Some of the tests have been proposed with specific alternatives in mind. The last two tests were suggested by Engle and Ng (1993) as powerful diagnostics for possible asymmetries in the conditional variance. The negative size bias for example would suggest the leverage effect of Nelson's (1991) Exponential-GARCH model. Tests 2 and 3 would point at instability of the parameter estimates.

Table III shows the diagnostics for the different volatility specifications. The diagnostics for the monthly and weekly data show that the pure GARCH model fails on the level test and the 79–82 dummy. The CKLS model has severe problems on the ARCH test. The KNSW and LS models both capture the level effect as well as the GARCH effects. The difference between the LS model and the KNSW specification is in the interest rate elasticity of volatility. This elasticity is restricted to $\gamma = \frac{1}{2}$ in the LS model, while it is estimated as 1.24 in the monthly KNSW model.

The diagnostics indicate that none of the models can cope with the 79–82 high volatility episode. To see the impact of this period on the parameter estimates of the volatility equation, we re-estimate the KNSW model where we exclude the high volatility period. The estimate of γ now drops to 0.88 for the monthly frequency. This result shows that the high sensitivity of volatility with respect to the level can partially be explained by this period.[11]

Normality is strongly rejected for all models. While this implies that we must be cautious in interpreting distributional implications of the models, it does not invalidate the parameter estimates, since QML is robust to departures from normality. Following Bollerslev (1987) we also considered the standardized t-distribution. The estimated degrees of freedom parameter ranges from three to five, thereby reflecting substantial fat-tailedness. However, the adjustment for fat-tailedness has no significant impact on the other parameters.[12]

3.4. Unconditional Distribution

The parameter estimates have implications for the unconditional moments of the interest rate. The unconditional moments also provide a test of the specification of the model, when the implied moments are compared to the sample moments of the interest rate level. For the unconditional distributions we would also have to take into account possible nonlinearities in the conditional mean (see Section 2).

Table III. Diagnostics

	GARCH	CKLS2	LS	KNSW
A: Monthly				
Skewness	0.10	-0.03	0.01	-0.08
Excess Kurtosis	2.13***	0.89***	2.05***	1.08***
Jarque-Bera	65.27***	11.54***	60.05***	17.15***
LM(level)	11.21***	0.78	0.01	0.10
LM(level^2)	13.94***	0.44	6.85***	0.85
LM(ARCH(4))	0.56	42.20***	2.00	1.54
LM(79:10)	4.23**	3.11**	4.99**	1.19
LM(79:10–81:12)	7.97***	9.69***	17.23***	5.36**
LM(sign)	0.06	7.22***	2.78*	2.19
LM(size)	0.83	30.74***	2.89	7.90*
B: Weekly				
Skewness	-0.10*	0.25	-0.13**	-0.14**
Excess Kurtosis	4.79***	7.13***	4.92***	4.40***
Jarque-Bera	1379.73***	3062.00***	1465.03***	1168.17***
LM(level)	11.27***	1.41	1.99	0.22
LM(level^2)	14.09***	0.24	8.85***	2.12
LM(ARCH(16))	6.75	176.97***	5.58	6.54
LM(79:10)	0.01	10.69***	0.17	0.67
LM(79:10–81:12)	4.85**	54.79***	9.87***	5.77**
LM(sign)	0.01	0.01	0.79	0.22
LM(size)	9.71**	143.74***	4.45	9.13**

Notes: Skewness is defined as m_3/s^3, with m_3 the centred third moment of the data and s the sample standard deviation. Kurtosis is defined as $m_4/s^4 - 3$, with m_4 the centred fourth moment of the data. Both are computed on the scaled residuals $\hat{e}_t/\hat{h}_t$. LM(level), LM(ARCH(p)), LM(79:10), LM(79:10–81:12), LM(sign), denote LM test statistics for the squared interest rate level, ARCH effects with p lags, a shift dummy after October 1979, a shift dummy for 79:10–81:12, a sign effect, and a combined size effect respectively. The size effect combines: a positive size effect, a negative size effect, a positive size square effect and a negative size square effect. * (**) [***] denote rejection at the 10% (5%) [1%] level.

To focus on the implications of the volatility dynamics, however, we compare the interest models under the restriction that they have a linear conditional mean.

In Section 3 we discussed that this linear specification defines a nonstationary process if $\gamma > 1$. For the KNSW and the CKLS model we, therefore, impose the restriction $\gamma = 1$ to obtain stationary distributions. Furthermore we consider the KNSW model with $\gamma = \frac{1}{2}$. Finally we investigate the GARCH model. For the GARCH specification it follows from panel B of Table II that the sum of the parameter estimates β_2 and β_3 is larger than unity. This restriction implies that the interest rate is not a covariance stationary process. We therefore impose

Table IV. Period 79–82 excluded

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-1}^2 + e_t,$$

$$h_t^2 = \beta_1 r_{t-1}^{2\gamma} + \left(\frac{r_t}{r_{t-1}}\right)^{2\gamma} (\beta_2 e_{t-1}^2 + \beta_3 h_{t-1}^2).$$

	$\alpha_0 \times 10$	$\alpha_1 \times 10$	$\alpha_2 \times 10^2$	$\beta_1 \times 10^2$	β_2	β_3	γ
Monthly	0.20	0.13	−0.21	0.18	0.16	0.56	0.88
	(0.20)	(0.32)	(0.50)	(0.14)	(0.47)	(0.30)	(2.06)

Notes: Loglik denotes the log-likelihood value. Robust t-values, are given in parentheses.

Table V. Parameter estimates of restricted models

$$r_t - r_{t-1} = \alpha_0 + \alpha_1 r_{t-1} + e_t,$$

$$h_t^2 = \beta_1 r_{t-1}^{2\gamma} + \left(\frac{r_t}{r_{t-1}}\right)^{2\gamma} (\beta_2 e_{t-1}^2 + \beta_3 h_{t-1}^2).$$

	GARCH	CKLS	KNSW1	KNSW2
α_0	0.03	0.03	0.02	0.02
	(1.62)	(1.77)	(1.55)	(1.63)
α_1	−0.01	−0.01	−0.01	−0.01
	(1.27)	(1.31)	(0.95)	(1.12)
$\beta_1 \times 10^2$	0.01	0.17	0.01	0.03
	(2.05)	(2.05)	(2.59)	(2.14)
β_2	0.17	–	0.26	0.23
	(3.43)	–	(4.62)	(3.36)
β_3	0.82	–	0.68	0.76
	(15.97)	–	(9.37)	(11.99)
γ	0	1	1	0.5
Loglik	83	−92	100	94

Notes: Loglik denotes the log-likelihood value. Robust t-values, conditional on γ are given in parentheses. The GARCH model is restricted to $\beta_2 + \beta_3 = 0.99$; the CKLS model has the restriction $\gamma = 1$; the two KNSW models have been estimated under the restriction $\gamma = 1$ and $\gamma = \frac{1}{2}$ respectively. No restrictions on β_2 or β_3 were needed in the KNSW models.

the restriction $\beta_2 + \beta_3 < 1$. Table V contains weekly parameter estimates of the restricted models.

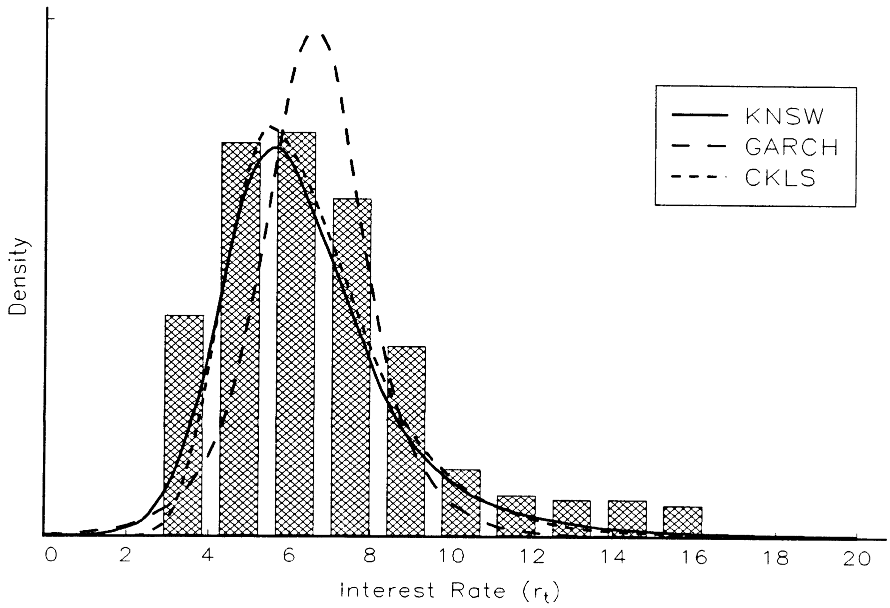

Figure 3. Unconditional distributions. The shaded area is the sample histogram at the weekly frequency.

Appendix A describes an algorithm to compute the implied unconditional distributions for the GARCH model, the CKLS model and the KNSW model for the parameter estimates of the restricted models in Table V. Figure 3 shows the unconditional distribution of the weekly data together with a histogram of the actual distribution of the one-month spot rate. The implied distribution of the KNSW model captures much of the skewness of the actual data. The skewness of the CKLS model is small and the GARCH model is symmetric. Note that we obtain this reasonable good fit despite the nonnormality of the errors. It appears that the rejection of the normality test is mainly due to a few large outliers in the 79–82 period, which have little impact on the unconditional distribution.

3.5. News Impact Curves

An insightful way to graphically illustrate the differences between the various volatility specifications is the news impact curve introduced in Engle and Ng (1993). The news impact curve shows the effect of the last shock, e_t, on the conditional volatility h_t^2. Writing r_t explicitly as a function of e_t and r_{t-1}, the general volatility specification Equation (7) becomes a function of e_t given values of the other state variables (r_{t-1}, h_{t-1}):

$$h_t^2 = \tilde{f}(e_t \mid r_{t-1} = r, h_{t-1}^2 = h^2). \tag{14}$$

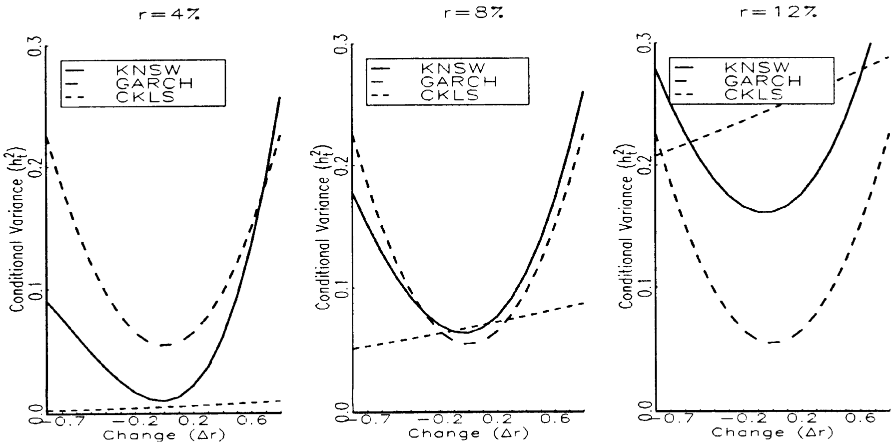

Figure 4. News impact curves.

Since the mean reversion is negligible over a one period horizon, we approximate $r_t \approx r_{t-1} + e_t$ and thus obtain h_t^2 as a function of e_t and the other state variables. We write the function as:

$$\tilde{f}(e \mid r, h^2) = \beta_1(r+e)^{2\gamma} + \left(1 + \frac{e}{r}\right)^{2\gamma} (\beta_2 e^2 + \beta_2 h^2). \tag{15}$$

Equation (15) is quadratic for the GARCH specification ($\gamma = 0$), but can be very asymmetric for the KNSW specification. For example, in the special case $\gamma = \frac{1}{2}$ the news impact curve is a cubic polynomial in e. The shape is also very different for different levels of the interest rate. A negative shock has two effects on the volatility. The first, the GARCH type volatility clustering, increases the volatility; the second effect is the decrease of the level and decreases the volatility. Eventually, for very large negative shocks, the level effect dominates.

The news impact curve of the KNSW model depends on the interest rate level, the last period's innovation and the last period's conditional variance. The news impact curve of the GARCH model depends on both last period's innovation and last period's conditional variance. The news impact curve of the CKLS model only depends on last period's interest rate level.

We will construct the news impact curves at different levels: the low level (r = 4%), the moderate level (r = 8%), and the high level (r = 12%).[13] Figure 4 contains the news impact curves at three interest rate levels based on the parameter estimates in Table V. In the first panel a negative shock does not have large impact on volatility for the KNSW model, while it increases volatility for the GARCH model. The curve for the low level clearly displays the asymmetry of the KNSW model. At the intermediate level the GARCH and KNSW models are very close with respect to upward shocks.

Again, for a negative shock the level effect and the GARCH effect almost cancel in the KNSW model so that volatility is not affected by downward shocks in the interest rate, while the CKLS model remains very asymmetric with a negative shock lowering volatility. The figures show the flexibility of the KNSW model: GARCH effects dominate at ‘normal’ levels while the asymmetry implied by the level effect is very strong at low levels. The KNSW model thus combines the features of the CKLS and GARCH models, shifting smoothly from one to the other as the level of interest rate varies. At high levels the three models diverge most in their volatility estimates. The CKLS model, which has the highest interest rate elasticity implies the highest conditional variance. The GARCH model which has no level dependence in the volatility does not show any big increase after a positive or negative shock. The figure is the same as the previous panels. The KNSW model is less asymmetric at high levels. Figure 5 combines the news impact curves of the previous figures. The surface gives the news impact curves of the KNSW model.

4. Bond Option Implications

The economic differences between the various volatility specifications can best be illustrated by considering the valuation of bond options under each of the models of the conditional heteroskedasticity. The models will have different option implications, because they imply different conditional densities for future spot rates. Figure 3 already highlighted the different unconditional densities of the short-term interest rate. These densities have similar unconditional first and second moments, but very different higher order moments. The higher the value of γ, the more right-tail skewness is introduced. The stronger the ARCH effects, the higher the fourth moments of the conditional densities. These distributional properties carry over to option prices. Options that pay off if interest rates are high, will be more valuable, ceteris paribus, the fatter the right hand tail. In this section we consider the differences in short horizon predictive densities.

We consider option implications, because option values will be much more sensitive to the distributional assumption than prices of long term bonds, which depend predominantly on conditional first and second moments. We will concentrate on options on long term bonds, but with a short expiration period. Over longer expiration horizons the option value will depend on both the volatility dynamics as well as the degree of mean reversion of the short-term interest rate. It will then be impossible to identify the sources of the differences between the various models.

Since the volatility specifications have been developed in discrete time, we will also develop the option implications in discrete time. Let $P_0^{(n)}$ be the price of an n-period discount bond at the time $t = 0$. A European call option on this bond with strike price K and expiration date $t = m$ is defined as the risk neutral expected present value of the payoff at the expiration date:

$$C_0(m, n, K) = \mathbf{E}_0[X_m[P_m^{(n-m)} - K]^+], \tag{16}$$

where $_X m = \exp(-\sum_{i=0}^{m-1} r_i)$ is the discount factor, and the operator $[Z]^+$ is defined as max$(Z, 0)$.

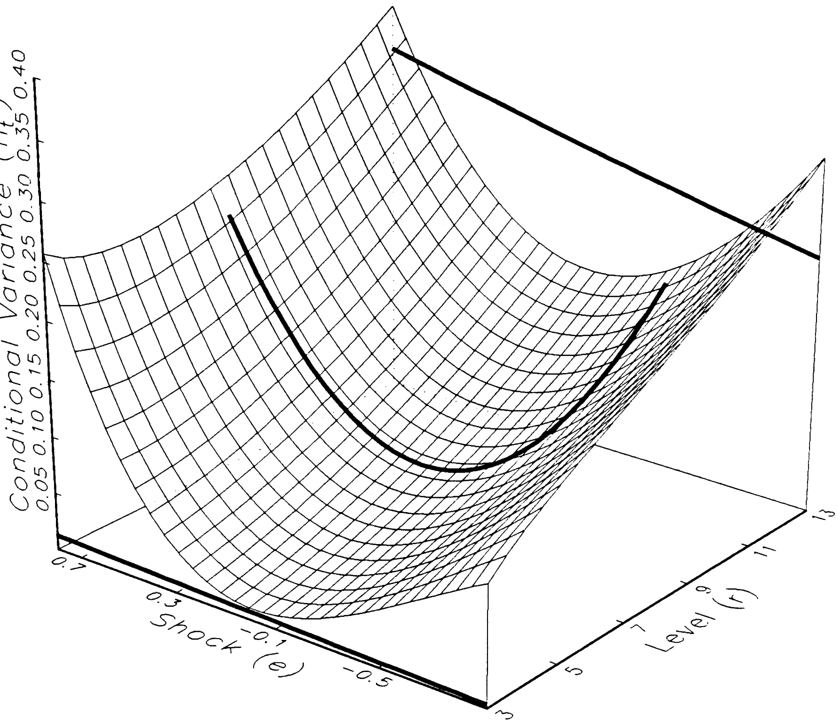

Figure 5. News Impact Surface. The surface are the news impact curves of the KNSW model. The parabola is the news impact curve of the GARCH model. The other two curves are the news impact curves of the CKLS model at the low level (4%) and at the high level (12%).

4.1. THE PRICE OF RISK

The expectation in Equation (16) has to be taken with respect to the risk neutral probability measure associated with the interest rate process, and not the actual interest rate process that we have estimated from the time series data. We therefore need an assumption about the price of risk. We assume that the risk adjustment takes the form of a change of the conditional mean of the interest rate process to:

$$\mathbf{E}_t[r_{t+1}] = \mu + \rho(r_t - \mu) + \lambda h_{t+1}, \tag{17}$$

where λ is a constant parameter, representing the price of risk. If $\gamma = 1$ the volatility specification is $h_{t+1} = \sigma_{t+1} r_t$ and Equation (17) can now be written as:

$$\mathbf{E}_t[r_{t+1}] = \tilde{\mu}_t + \tilde{\rho}_t(r_t - \tilde{\mu}_t). \tag{18}$$

Equation (17) defines the risk neutral conditional mean parameters: $\tilde{\rho}_t = \rho + \lambda\sigma_{t+1}$, and $\tilde{\mu}_t = (1-\rho)/(1-\tilde{\rho}_t)\mu$. Since $\lambda > 0$, the adjustment has the effect of increasing the mean of the interest rate ($\tilde{\mu}_t > \mu$), and reducing the amount of mean reversion ($\tilde{\rho}_t > \rho$) thereby increasing the unconditional variance.

The CKLS model belongs to the class of one factor models, with the spot rate as the only state variable. The KNSW model is a two factor model with volatility as a second factor. For comparability across models we assume that the price of volatility risk is equal to zero, so that we use a single risk price λ.

4.2. THE TERM STRUCTURE

The main obstacle in computing the option values is that the payoff at the expiration date $t = m$ depends on the value of a discount bond with maturity $(n - m)$, which is itself a function of all state variables, and not available in closed form. The value of a bond with some maturity k is a function of all state variables. In the KNSW model there are three state variables: the level of the short-term rate r, the volatility h, and the latest shock e:

$$P_t^{(k)} = p(r_t, h_t, e_t). \tag{19}$$

In simulating the payoff of the option we must be able to compute the bond at any value that the state variables can attain at time m. Since the function $p(\cdot)$ is not known analytically it must be computed numerically. Using a naive Monte Carlo method this would require a simulation for each particular combination of state variables. This naive simulation procedure is illustrated in Figure 6. To obtain the distribution of the state variables at time m, (r_m, e_m, h_m), conditional on the starting values, (r_0, e_0, h_0), we have to sample L_m paths of length m. To obtain the payoff at time $t = m$, we have to sample L_n paths of length $(n - m)$ to compute the bond price $P_m^{(n-m)}$ at each value (r_m, e_m, h_m).

In practice, one could define a three dimensional grid for the state variables, compute the bond price only at the grid points and use interpolation for points in between. This is still computationally very costly. We therefore opted for a different approach based on the same sampling idea as for the computation of the unconditional density described in Appendix A. Figure 7 shows how this simulation reduces the number of sample paths. Again we first have to obtain the distribution of the state variables, (r_m, e_m, h_m), at the expiration date of the option.

Let r_s, $(s = 1, \ldots, N)$ be a single long realization from the risk neutral interest rate process. At each period s we store the state variables r_s, h_s and e_s, and also the quantities

$$I_s^{(k)} = \exp\left(-\sum_{i=0}^{k-1} r_{s+i}\right), \tag{20}$$

for different values of k. We approximate the bond price function by a polynomial correction to the expectations hypothesis:

$$P^{(k)} = \beta \exp\left(-k\tilde{\mu} - \frac{1-\tilde{\rho}^k}{1-\tilde{\rho}}(r - \tilde{\mu})\right)$$

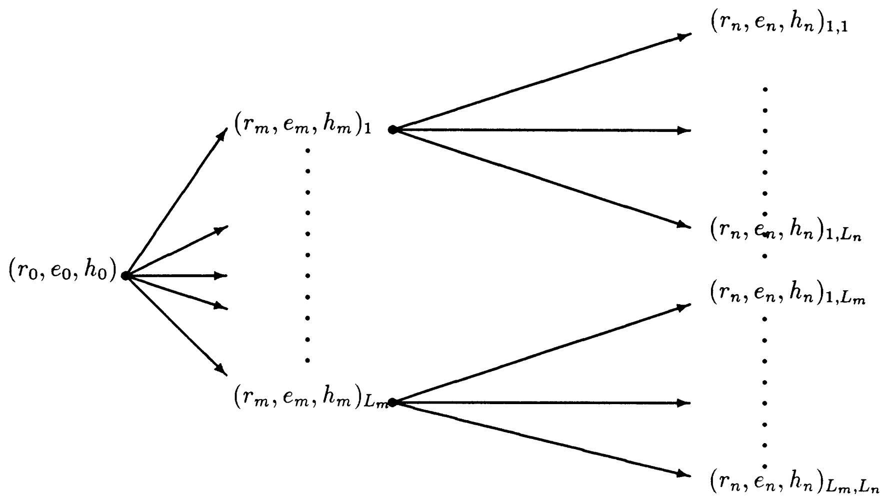

Figure 6. Native simulation. m, n, L_m, L_n denote the maturity of the option, the maturity of the bond, the number of combinations (r_m, e_m, h_m) and the number of combinations (r_n, e_n, h_n) respectively. L_n simulations are used to compute $P_m^{(n-m)}$ at each of L_m possible states at time m.

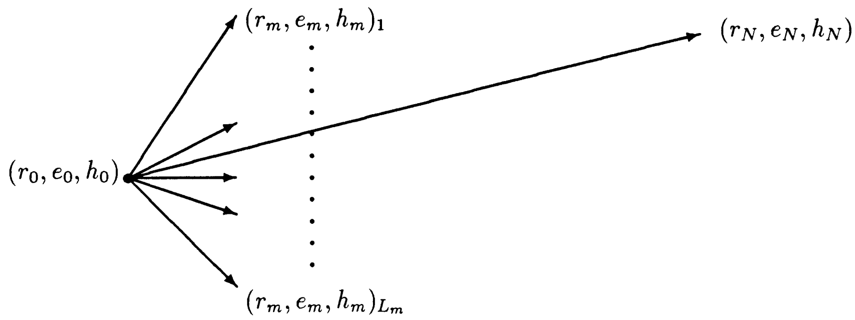

Figure 7. Efficient simulation. m, n, L_m, N denote the maturity of the option, the maturity of the bond, the number of combinations (r_m, e_m, h_m) and the length of a single long realization. The bond price is computed as a time invariant function $p(r, e, h)$.

$$+a_0 + \sum_{i=1}^{3} a_i x_i + \sum_{i=1}^{3} \sum_{j=1}^{i} a_{ij} x_i x_j + \sum_{i=1}^{3} \sum_{j=1}^{i} \sum_{\ell=1}^{j} a_{ij\ell} x_i x_j x_\ell, \qquad (21)$$

where x_i, ($i = 1, \ldots, 3$), are the state variables (r, e, h), and β, a_i, a_{ij} and $a_{ij\ell}$ are parameters. The leading term in Equation (21) is an approximation to the bond price according to the linearized expectations hypothesis under the risk neutral conditional mean in Equation (17), which would set the price at time t equal to $\exp(-\mathbf{E}_t[\sum_{i=0}^{k-1} r_{t+i}])$. The constant parameters $\tilde{\rho}$ and $\tilde{\mu}$ are fixed at the time series average of $\tilde{\rho}_t$ and $\tilde{\mu}_t$ from the simulation. The other parameters are estimated by OLS from the linear regression:

$$
\begin{aligned}
I_s^{(k)} &= \beta \exp\left(-k\tilde{\mu} - \frac{1-\tilde{\rho}^k}{1-\tilde{\rho}}(r_s - \tilde{\mu})\right) \\
&\quad + a_0 + \sum_{i=1}^{3} a_i x_{is} + \sum_{i=1}^{3}\sum_{j=1}^{i} a_{ij} x_{is} x_{js} \\
&\quad + \sum_{i=1}^{3}\sum_{j=1}^{i}\sum_{\ell=1}^{j} a_{ij\ell} x_{is} x_{js} x_{\ell s} + u_s. \qquad (22)
\end{aligned}
$$

If the simulated sample size $N \to \infty$ the regression function converges to an approximation of the true bond price function.[14] We now use the regression function to calculate the bond price for each combination of (r_m, e_m, h_m).

In the discrete time process negative interest rates can occur during the simulations. We handle negative interest rates by introducing a reflecting barrier at $r = 0$. This means that negative draws of r are rejected (see Black (1995) for a motivation introducing such a reflecting barrier).

In the application we consider four week options on a ten year discount bond, i.e. $m = 4$ weeks, and $n = 520$ weeks. For the different specifications the risk price λ was calibrated such that the estimated ten year discount bond price implies an average yield equal to the average ten year yield observed in the data. Using CRSP data for the period 1970–1995 the ten year discount yield is 8.5%. For example, for the model with $\gamma = 1$ this gives $\gamma = 0.05$. The average autocorrelation parameter of the risk neutral process then comes out at $\tilde{\rho} = \rho + \lambda\tilde{\sigma} = 0.990 + 0.05 \times 0.052 = 0.993$.[15] Figure 8 shows the unconditional density of the short rate and the risk neutral density of the short rate. The mean of the risk neutral density is about 8.5% and the standard deviation is larger than that of its empirical counterpart. We checked whether the approximation in Equation (21) gives admissible bond prices $0 < P^{(k)}(r, e, h) < 1$ for $k = m, n, n-m$. This appears true at all points realized in the simulation.

4.3. Option Simulation

Although all three specifications are univariate time series models, the conditional distribution of next period's spot rate for the KNSW model depends on three state variables: the spot rate, the innovation to the spot rate and the conditional variance. For a full comparison of the different models, we must compare the implications at different levels of all the state variables. In order to keep things manageable we present our results in a two-way table, distinguishing three different levels of the spot rate (r_0 = 4%, 8%, and 12%), and three different shocks.

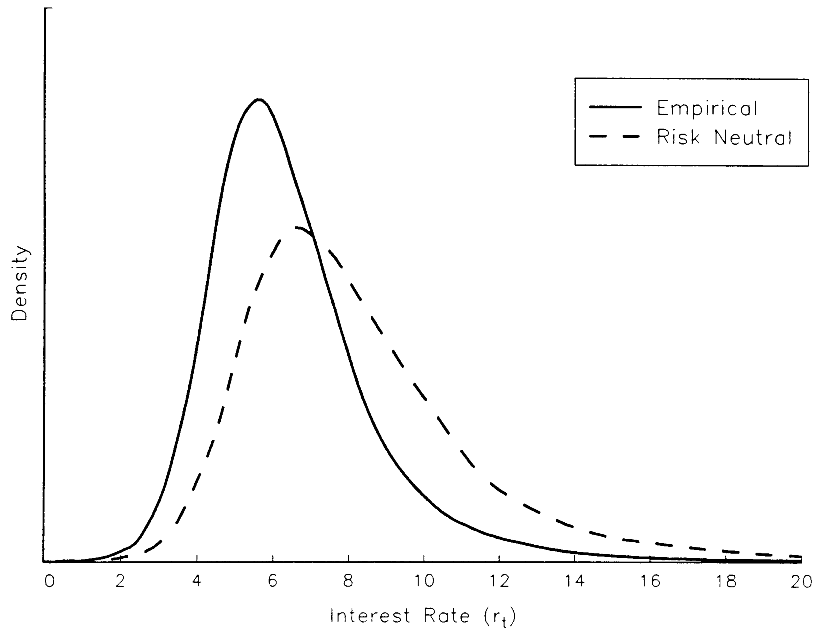

Figure 8. Risk neutral unconditional distribution. The risk neutral density assumes that the price of risk $\lambda = 0.05$.

The size of the typical shock at these different interest rates is calculated from the simulated data as:

$$|e_0| = \frac{\sum_{s=1}^{N} w_s |e_s|}{\sum_{s=1}^{N} w_s}, \tag{23}$$

with w_s a weight function based on a simple normal kernel:

$$w_s = \exp\left(-\frac{1}{2\Delta V_r^2}(r_s - r_0)^2\right), \tag{24}$$

where Δ is a bandwidth parameter and V_r^2 is the unconditional variance of the simulated interest rate.[16] Typical shocks we consider for each initial interest rate r_0 are $-e_0$, 0, and e_0. At each of these initial conditions we set the initial conditions of the third state variable h^2 at its conditional expectation given the other state variables:

$$h_0^2 = \mathbf{E}[h^2 \mid r_0 = r, e_0 = e]. \tag{25}$$

The conditional expectation is computed from the same simulated as used for the construction of the yield curve:

$$\hat{h}_0^2 = \frac{\sum_{s=1}^{N} w_s h_s^2}{\sum_{s=1}^{N} w_s}, \tag{26}$$

where w_s is obtained from the multivariate kernel:

$$w_s = \exp\left(-\frac{1}{2\Delta}(x_s - x_0)'\Omega^{-1}(x_s - x_0)\right), \tag{27}$$

where x_s is redefined as the subvector of state variables (r_s, e_s) and Ω is the unconditional covariance matrix of x_s from the simulation of length N.

The future bond price $P_m^{(n-m)}$ in Equation (16) is replaced by its approximate functional form $p(r_m, e_m, h_m)$ in Equation (21). With this approximation we use a conventional Monte Carlo simulation to estimate the option value for various initial conditions $(r_0, e_0, \hat{h}_0^2)$ and strike prices K. The option prices in Table VI are computed for at-the-money options. The strike price for an at-the-money option is defined as the initial forward price:

$$K = P_0^{(n)}/P_0^{(m)}, \tag{28}$$

where the bond prices are consistent with the implied term structure, i.e.:

$$P_0^{(k)} = \frac{\sum_{s=1}^{N} w_s I_s^{(k)}}{\sum_{s=1}^{N} w_s}, \tag{29}$$

for $k = m, n$ and where w_s is the same as in Equation (27). Each sample path for the spot rate depends on *m* drawings for the random variables ϵ_s, $(s = 1, \ldots, m)$. Using the sequence ϵ_s a sample path for r_{t+s} $(s = 0, \ldots, m)$ is constructed based on the parameter estimates. The option value $C_0(m, n, K)$ is estimated by averaging over L_m simulated paths. For the tables below we have set N = 100,000 and L_m = 5000.

4.4. Results

In Table VI we report option values as a percentage of the underlying long-term discount bond, i.e. $C_0(m, n, K)/P_0^{(n)}$. The table consists of seven rows and three columns. The columns refer to the three different levels of the spot rate; the rows represent the three different values for last period's shock to the spot rate. The upper panel and the middle panel report option values according to the KNSW model, with different restrictions on the parameters. The option values in the lower panel, only depend on the level of the spot rate. Appendix B describes how standard errors are computed.

From the table we draw several conclusions. The standard errors of the option values of the KNSW model are small, meaning that small differences with other models will lead to statistically significant differences in option valuation. Parameter uncertainty is not a big issue here.

Both dimensions, level and shock, are important for valuing options in the KNSW model. The effect of a shock is to increase volatility which will lead to

Table VI. Percentage option value

	4%	8%	12%
KNSW1	$\alpha_2 = 0, \gamma = 1$		
Zero shock	0.42	0.47	0.60
	(0.01)	(0.01)	(0.02)
Positive shock	0.59	0.63	0.79
	(0.01)	(0.02)	(0.03)
Negative shock	0.56	0.61	1.03
	(0.01)	(0.02)	(0.04)
KNSW2	$\alpha_2 = 0, \gamma = \frac{1}{2}$		
Zero shock	0.40	0.40	0.42
	(0.01)	(0.01)	(0.01)
Positive shock	0.53	0.49	0.55
	(0.01)	(0.01)	(0.01)
Negative shock	0.52	0.50	0.76
	(0.01)	(0.01)	(0.02)
CKLS	$\alpha_2 = 0, \beta_2 = \beta_3 = 0$		
	0.42	0.55	0.70
	(0.01)	(0.01)	(0.02)

Notes: Results pertain to an at-the-money European call option on a ten year discount bond with four weeks to expiration. Option prices are based on the weekly parameter estimates. Standard errors reflecting parameter uncertainty are in parentheses. The size of the shocks at the 4%, 8% and 12% interest rate level are (−0.2%, 0%, 0.2%), (−0.3%, 0%, 0.3%) and (−0.6%, 0% 0.6%) respectively.

higher option values. At low and moderate interest rates the effect of a positive shock on the option value is larger than the effect of a negative shock. At high interest rates, however, the effect of a negative shock is higher.

For all specifications differences along a column in the table depend on the value of γ. At low and moderate values of r the option values when $\gamma = 1$ are close to the option values when $\gamma = \frac{1}{2}$. At the high level the implied option value is positively related to the interest rate elasticity of the model.

5. Conclusions

In this paper we presented and estimated a model for the short-term interest rate volatility, that encompasses both the level effect in the CKLS model and the conditional heteroskedasticity effect of the GARCH class of models. The flexible specification of the conditional variance equation allows different effects to dominate as the level of the interest rate varies. The different models were estimated for monthly as well as for weekly data. We find that both GARCH effects and level effect are important determinants of interest volatility.

The parameter that measures the sensitivity of interest rate volatility with respect to the interest rate level, γ, is highly significant. The important empirical difference of the KNSW model and the CKLS type specifications is the smaller estimate of γ. For the estimation of the interest rate sensitivity in the variance specification one cannot ignore the strong GARCH effects in monthly and weekly data. Ignoring GARCH creates an omitted variables problem for the estimate of the level effect in the volatility.

The most precise estimates of the volatility specification are obtained at the weekly frequency. The estimated value of γ ranges from 1.40 for the CKLS model for the monthly frequency to 1.21 for the CKLS model for the weekly frequency. The parameter estimate of γ is not significantly different from unity for the CKLS and KNSW models.

As the volatility of the short-term interest rate is one of the determinants for the pricing of interest rate contingent claims, we investigate the implications of the dynamics of short-term interest rate volatility for the pricing of discount bond options. Our results suggest that the inclusion of a GARCH effect in addition to a level effect in the model specification is relevant for the pricing of short-term discount bond options. This result is related to the lower estimated value of γ when volatility effects are included. We show that at interest rate levels of 12% a change in the value of γ results in a large change of the relative option value.

Notes

[1] The research of CKLS has generated a lot of discussion in the finance literature. Recent contributions to the debate on the interest rate volatility dynamics include Aït-Sahalia (1996a,b), Andersen and Lund (1996a,b), Brenner, Harjes and Kroner (1995), Conley, Hansen, Luttmer and Scheinkman (1995), Tauchen (1996) and Torous and Ball (1995).

[2] Sensitivity is defined as $\partial \ln h / \partial \ln r$, where r is the interest rate level and h is the standard deviation.

[3] Bollerslev, Chou and Kroner (1992) provide a survey of empirical studies in this vein.

[4] Bomhoff and Schotman (1988) estimate the same model for monthly data for Germany, Japan and the United States and find that the level effect improves the specification of the volatility equation. The GARCH-M effect turns out to be insignificant.

[5] The constant term β_1 also enters the conditional variance equation under exact aggregation of the continuous time process of CKLS. See also Section 4 below.

[6] Brenner, Harjes and Kroner (1995) propose a different approach. In Equation (6) they use the unscaled prediction error e_t instead of the scaled prediction error $\sigma_t \epsilon_t$. The stationarity conditions of their specification are hard to establish.

[7] The initial condition for h_0 is the unconditional variance.

[8] For example, in the CIR model ($\gamma = \frac{1}{2}$) the exactly aggregated conditional volatility takes the form $h_t^2 = \beta_0 + \beta_4 r_{t-1}$, with $\beta_0 > 0$ and $\beta_4 \neq 1$ (see DeMunnik and Schotman (1994)). Pagan, Hall and Martin (1994) focus on temporal aggregation problems of the CKLS model.

[9] The estimate of γ is lower than in CKLS. We have not been able to exactly replicate their results due to some differences between their data and ours. For the overlapping sample period we have some different data points in 1987.

[10] In order to ensure that estimates are comparable across models we set the GARCH-M parameter

in LS to zero. If this parameter is included it is not significantly different from zero. When β_1 is a free parameter it attains a negative value. Therefore we also impose $\beta_1 = 0$.

[11] These results also hold for the CKLS model. Aït-Sahalia (1996a,b) shows that the value of the γ parameter is a nonlinear function of the interest rate level. The value of γ decreases for high interest rate levels.

[12] The tables for the Student-t results are not included in this paper.

[13] The conditional variance is specified as $h^2 = \left(\frac{1}{T}\sum_{t=1}^{T} e_t^2/r_{t-1}^{2\gamma}\right) r^{2\gamma}$.

[14] A consistent estimator of the true, instead of an approximate, implied bond price can be obtained through a nonparametric kernel method using the output of the simulation. But it would be computationally expensive to run the kernel estimator for every iteration in the subsequent Monte Carlo simulation of the option price.

[15] The parameter estimates ρ and $\bar{\sigma}$ are based on the results in Table V. The autocorrelation parameter is calculated as: $\rho = 1 - \alpha_1 = 1 - 0.010 = 0.990$ and the average volatility is calculated as

$$\bar{\sigma} = \left(\frac{\beta_1}{1 - \beta_2 - \beta_3}\right)^{1/2} = \left(\frac{0.14 \times 10^{-2}}{1 - 0.23 - 0.76}\right)^{1/2} = 0.052.$$

[16] The bandwidth parameter Δ is chosen as $\Delta = 1.06 V_r N^{-0.2}$ (see Silverman (1986)).

Appendix A: Marginal Distribution of Spot Rate

To compute the unconditional density we follow Geweke (1994). Let $f(x \mid y)$ denote the conditional density of r_t given r_{t-1}. The unconditional density of the spot rate is the solution, if it exists, to the integral equation:

$$g(x) = \int_0^\infty f(x \mid y) g(y)\, \mathrm{d}y. \tag{30}$$

Since the spot rates are highly correlated, a Monte Carlo simulation is very inefficient for obtaining the unconditional density. The accuracy can be improved by drawing a sequence of conditional densities and averaging the densities. Define a grid of points ($i = 1, \ldots, N$) at which to estimate the density $g(y)$. Now draw a time series sample of spot rates r_t ($t = 1, \ldots, T$), and for each r_t evaluate the conditional density $f(y_i \mid r_t)$ at each of the grid points. The unconditional density at a point y_i is finally estimated as:

$$g(y_i) = \frac{1}{T}\sum_{t=1}^{T} f(y_i \mid r_t) \quad (i = 1, \ldots, N). \tag{31}$$

Whenever a negative value of r_t is drawn, the draw is rejected and the value of r_t is set to zero, this way truncating the distribution to positive interest rates. The number of Monte Carlo draws was set to 1×10^6 and the number of negative interest rate drawings was zero for the weekly data. The starting value for the Monte Carlo runs was set to the sample mean.

Appendix B: Computation of Standard Errors

The standard errors of the implied option values reflect the uncertainty about the parameter estimates but are conditional on an exogenously given initial term structure. Option standard errors are computed using the asymptotic formula:

$$V(C(\hat{\theta})) = \left(\frac{\partial C(\hat{\theta})}{\partial \theta}\right)' V(\hat{\theta}) \left(\frac{\partial C(\hat{\theta})}{\partial \theta}\right), \tag{32}$$

where $V(\hat{\theta})$ denotes the covariance matrix of the parameter estimates, and $\partial C(\hat{\theta})/\partial\theta$ denotes the partial derivatives of the option price with respect to the parameters of the volatility process evaluated at $\theta = \hat{\theta}$, where $\hat{\theta}$ are the parameters of the restricted volatility specification in Table V.

We calculate derivatives numerically. Using the *same* sequence ϵ_t as for the estimation of the option value a new sample path r_t $(t = 1, \ldots, m)$ is constructed with parameter vector $\overline{\theta_j} = \hat{\theta} + \ell_j$, where ℓ_j is a vector with zeros apart from a small number δ_j at position j corresponding with the jth element of θ. The new value for θ is used to compute a new option value at the expiration date. Using the same random numbers the procedure is repeated using $\underline{\theta_j} = \hat{\theta} + \ell_j$. The numerical central derivative of the option with respect to θ_j is then estimated as the average over N simulations:

$$\frac{\partial C(\hat{\theta})}{\partial \theta_j} = \frac{1}{N}\sum_{k=1}^{N} \frac{C(\overline{\theta_j})_k - C(\underline{\theta_j})_k}{2\delta_j}. \tag{33}$$

References

Aït-Szahalia, Y. (1996a) Nonparametric pricing of interest rate derivative securities, *Econometrica* **64**, 527–560.

Aït-Szahalia, Y. (1996b) Testing continuous-time models of the spot interest rate, *Review of Financial Studies* **9**, 385–426.

Andersen, T. and Lund, J. (1996a) The short rate diffusion revisited: An investigation guided by the efficient method of moments, Working Paper, Northwestern University.

Andersen, T. and Lund, J. (1996b) Stochastic volatility and mean drift in the short rate diffusion: Sources of steepness, level and curvature in the yield curve, Working Paper, Northwestern University.

Balduzzi, P., Das, S. R., and Foresi, S. (1995) The central tendency: A second factor in bond yields, Working Paper, New York University.

Black, F. (1995) Interest rates as options, *Journal of Finance* **50**, 1371–1376.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**, 307–328.

Bollerslev, T. (1987) A conditional heteroskedastic time series model for speculative prices and rates of return, *Review of Economics and Statistics* **69**, 542–547.

Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992) ARCH modelling in finance: Review of the theory and the empirical evidence, *Journal of Econometrics* **52**, 5–59.

Bollerslev, T., Engle, R. F., and Nelson, D. B. (1994) ARCH models, in R. F. Engle and D. L. McFadden, (eds.), *Handbook of Econometrics*, Volume 4, North-Holland, Amsterdam.

Bollerslev, T. and Wooldridge, J. M. (1992) Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances, *Econometric reviews* **11**, 143–172.

Bomhoff, E. J. and Schotman, P. C. (1988) The term structure in the United States, Japan and West Germany, *Carnegie Rochester Conference Series on Public Policy* **28**, 269–314.

Brenner, R. J., Harjes, R. H., and Kroner, K. (1995) Another look at alternative models of the short-term interest rate, *Journal of Financial and Quantitative Analysis* **1**, 85–107.

Broze, L., Scaillet, O., and Zakoian, J. (1995) Testing for contiuous-time models of the short-term interest rate, *Journal of Emprical Finance* **2**, 199–223.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., and Sanders, A. B. (1992) An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance* **52**, 1,209–1,227.

Conley, T. G., Hansen, L. P., Luttmer, E. G. J., and Scheinkman, J. A. (1995) Short-term interest rates as subordinated diffusions, Working Paper.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985) A theory of the term structure of interest rates, *Econometrica* **53**, 385–408.

DeMunnik, J. F. J. and Schotman, P. C. (1994) Cross sectional versus time series estimation of term structure models; Empirical results for the Dutch bond market, *Journal of Banking and Finance* **18**, 997–1,025.

Drost, F. C. and Nijman, T. E. (1993) Temporal aggregation of GARCH processes, *Econometrica*, **61**, 909–927.

Engle, R. F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**, 987–1,007.

Engle, R. F. (1984) Wald, likelihood ratio, and Lagrange multpier tests in econometrics, in Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics*, Volume 2, North-Holland, Amsterdam.

Engle, R. F. and Ng, V. K. (1993) Measuring and testing the impact of news on volatility, *Journal of Finance* **48**, 1,749–1,778.

Geweke, J. (1994) Monte Carlo simulation and numerical integration, Working Paper 562, Federal Reserve Bank of Minneapolis.

Longstaff, F. A. and Schwartz, E. S. (1992) Interest rate volatility and the term structure: A two-factor general equilibrium model, *Journal of Finance* **47**, 1,259–1,282.

Nelson, D. B. (1990) ARCH models as diffusion approximations, *Journal of Econometrics* **45**, 93–98.

Nelson, D. B. (1991) Conditional heteroskedasticity in asset pricing: A new approach, *Econometrica* **59**, 347–370.

Pagan, A. R., Hall, A. D., and Martin, V. (1994) Exploring the relations between the finance and econometrics literatures on the term structure, Working Paper, Australian National University.

Pfann, G. A., Schotman, P. C., and Tschering, R. (1996) Nonlinear interest rate dynamics and implications for the term structure, *Journal of Econometrics* **74**, 149–176.

Silverman, B. W. (1986) *Density Estimation for Statistics and Data Analysis*, Chapmann and Hall, London.

Stanton, R. (1995) A nonparametric model of term structure dynamics and the market price of interest rate risk, Working Paper, U.C. Berkeley.

Tauchen, G. E. (1996) New minimum chi-square methods in empirical finance, in D. M. Kreps and K. F. Wallis (eds.), *Advances in Economics and Econometrics: Theory and Applications*, Cambridge University Press.

Torous, W. and Ball, C. A. (1995) Regime shifts in the short term riskless interest rates, IFA Working Paper 216.

Wooldridge, J. M. (1994) Estimation and inference for dependent processes, in R. F. Engle and D. L. McFadden (eds.), *Handbook of Econometrics*, Volume 4, North-Holland, Amsterdam.