The Dynamics of the Short-Term Interest Rate in the UK

Antonios Antoniou

Durham Business School
Mill Hill Lane
Durham DH1 3LB
United Kingdom
Phone: +44 191 334 5290
antonios.antoniou@durham.ac.uk

Alejandro Bernales S.

Inter-American Development Bank
1300 New York Avenue, N.W.
Washington, DC 20577
Phone: +1 202 623 2575
abernale@uc.cl

Diether W. Beuermann

Inter-American Development Bank
1300 New York Avenue, N.W.
Washington, DC 20577
Phone: +1 202 623 2172
dietherbeu@aloe.ulima.edu.pe

(December, 2005)

Abstract

We estimate and test different continuous-time short-rate models for the UK. The preferred model encompasses both the "level effect" of Chan, Karolyi, Longstaff and Sanders (1992a) and the conditional heteroskedasticity effect of GARCH type models. Our findings suggest that including a GARCH effect in the specification of the conditional variance, almost halves the dependence of volatility on rate levels. We also find weak evidence of mean-reversion and volatility asymmetries in the stochastic behavior of rates. Extensive diagnostic tests suggest that the Constant Elasticity of Variance model of Cox (1975), with an added GARCH effect, provides a reliable description of short-rate dynamics. We demonstrate that the most important feature in short-rate modeling is the correct specification of the conditional variance of changes in rates; suggesting that the conditional mean characterization is of second order.

JEL: C22

Keywords: Short-rate, level effect, GARCH effect.

Findings and interpretations correspond to the authors and do not represent the opinion of the University of Durham and/or The Inter-American Development Bank.

I. Introduction

Models of the term structure of interest rates are widely used in pricing interest rate derivatives and instruments with embedded options, such as callable bonds and mortgage-backed securities. Many such models are based on the simplifying assumption that changes in the entire term structure are driven by changes in a single underlying random factor, often taken to be the "short" or "instantaneous" rate of interest. Therefore, these term structure models are called "Single-Factor Models".

In this study, we provide an analysis of the short-term interest rate in the UK covering a 20 years period. The study's contribution is to add to the literature in this area by examining different proposed single-factor models and extending them in order to capture time-varying volatility dynamics. We provide the first application of a time-varying volatility version of Nowman (1997) exact discrete model, and compare it with the extended version of the Chan et. al. (1992a, hereafter CKLS) discrete approximation, proposed by Brenner et.al. (1996, hereafter BHK). The study compares the performance of the various models using an extended set of diagnostic and prediction tests in and out of sample.

The models commence with the generalized diffusion process proposed by CKLS which nests, trough parameters restrictions, many of the traditional term structure models.[1] These are continuous time models in which volatility is parameterized only as a function of interest rate levels; we refer to them as "Level Models".[2]

The study then examines the class of models which incorporates persistence in volatility dynamics, such as the GARCH and GJR-GARCH models.[3] The analysis then extends these models to incorporate both levels and time-varying volatility dynamics as proposed by BHK; we refer to them as "Level-News Models". Finally, we extend the CKLS nested models in order to incorporate time-varying volatility dynamics, and propose a new version of single factor model that appears to account for UK interest rates dynamics reliably.

The question about which model best captures the dynamics of the short-term rate is an empirical issue that has been widely addressed for the "Level Models", but very little research has been done regarding the "Level-News Models".[4] Furthermore, empirical tests have tended to analyze either "Level Models" or GARCH models, but very few studies provide a comparison of these two classes of models. Moreover, the few existing tests comparing different classes of models within a nested framework are heavily biased towards the US market.[5] In this sense, we add to the international literature evidence arising from a variety of models in the UK.

The rest of the study is organized as follows. Chapter II reviews the relevant literature. Chapter III discusses the econometric approach. Chapter IV presents the data set used. Chapter V reports empirical results and their interpretation. Finally, Chapter VI concludes.

[1] See section 2.1 for a detailed outline of the CKLS nested models.
[2] For a discussion of continuous time processes see Hull, John (2003) "*Options, Futures and Other Derivatives*". Fifth edition, Chapters 11 and 23.
[3] See Engle (1982), Bollerslev (1986), Engle and Bollerslev (1986), and Glosten, Jagannathan and Runkle (1993).
[4] See for example, Chan et.al. (1992a, 1992b), Nowman (1997), Byers and Nowman (1998).
[5] See for example, BHK, Andersen and Lund (1997), Koedijk et.al. (1992), and Bali (2003).

II. Literature Review

2.1. The CKLS Nested Models

CKLS seminal paper nested into one "unrestricted" stochastic differential equation (SDE), several well-known single-factor models of the short-rate. CKLS SDE is

$$dr = (\alpha + \beta r)dt + \sigma r^{\gamma} dZ \tag{1}$$

where r is the short-term interest rate at time t, α and β are the parameters that describe the conditional mean of changes in rates, σ is the volatility of the interest rate, γ measures the sensitivity of the volatility of rates on the level of rates (elasticity parameter), and Z is a Brownian motion.[6]

The stochastic process (1) implies that changes in interest rates have a drift rate of $(\alpha + \beta r)$ and a variance rate of $(\sigma^2 r^{2\gamma})$. Therefore, we can see that the conditional mean and variance of changes in the short-term rate depend on the level of r. By definition, $\alpha > 0$ and $\beta < 0$, and we can rewrite (1) as

$$dr = -\beta(-\frac{\alpha}{\beta} + r)dt + \sigma r^{\gamma} dZ \tag{2}$$

The model implies that the long-run mean of interest rates is $-\frac{\alpha}{\beta}$, and that the speed of mean reversion towards the long-run mean is given by β. The more negative the β is, the faster r responds to deviations from the long-run mean.[7]

The SDE given in (1) defines a broad class of interest rate processes. Several short-rate models can be obtained from (1) by placing the appropriate restrictions on the parameters α, β, σ and γ. These nested models and the corresponding parameter restrictions are summarized in Table 2.1.1.

TABLE 2.1.1

Parameter Restrictions Imposed by Alternative Models of The Short-Term Interest Rate

Alternative models of the short-term riskless rate of interest r can be nested with appropriate parameter restrictions within the unrestricted model

$$dr = (\alpha + \beta r)\,dt + \sigma r^{\gamma}\,dZ$$

Model Name	Model Specification	α	β	σ	γ
Merton (1973)	$dr = \alpha\,dt + \sigma\,dZ$		0		0
Vasicek (1977)	$dr = (\alpha + \beta r)\,dt + \sigma\,dZ$				0
Cox et.al. (1985)	$dr = (\alpha + \beta r)\,dt + \sigma r^{0.5}\,dZ$				0.5
Dothan (1978)	$dr = \sigma r\,dZ$	0	0		1
GBM (1973)	$dr = \beta r\,dt + \sigma r\,dZ$	0			1
Brennan-Schwartz (1980)	$dr = (\alpha + \beta r)\,dt + \sigma r\,dZ$				1
Cox et.al. (1980)	$dr = \sigma r^{1.5}\,dZ$	0	0		1.5
CEV (1975)	$dr = \beta r\,dt + \sigma r^{\gamma}\,dZ$	0			

[6] A Brownian motion is a continuous-time stochastic process with the properties that between any two dates s and t (s > t), the increment $Z_s - Z_t$ has a normal distribution with mean zero and variance of s – t and the increment is independent of the value of the process at all dates prior to t.

[7] For example, if β = -0.9, it will take 1/0.9 = 1.11 periods for the short-rate to revert towards the long-run mean.

CKLS nested models correspond to the "Level Models". These models differ in how they parameterize expected changes and volatilities of interest rates; in particular, they differ in how they assume expected changes and volatilities of interest rates are related to their levels. These models implicitly assume that the conditional volatility of interest rates is a function of the level of rates, and the strength of the relation between volatility and levels of rates is given by the elasticity parameter γ.[8]

For example, the models of Merton (1973) and Vasicek (1977) assume $\gamma = 0$, therefore the volatility is not related to the level of rates and rates have a constant volatility (i.e. Homoskedastic Models). Cox et.al. (1985, hereafter CIR-SR) assumes that the volatility of rates responds proportionally to the square root of their levels ($\gamma = 0.5$). The models of Dothan (1978), the Geometric Brownian Motion (GBM) - used in the option pricing formula of Black and Scholes (1973) - , and Brennan and Schwartz (1980) propose characterizations in which the volatility responds directly to rate levels ($\gamma = 1$). Cox et.al. (1980, hereafter CIR-VR) assumes $\gamma = 1.5$. Finally, the constant elasticity of variance model (hereafter CEV) introduced by Cox (1975), does not impose any restrictions on γ. Whether the assumptions of these "Level Models" are realistic or not is an empirical issue that is surveyed in the next section.

2.2. Empirical Evidence on "Level Models"

The right hand side of (1) has two parts. The first is the conditional mean of changes in interest rates with a drift rate of $(\alpha + \beta r)$, and the second is the conditional volatility of rate changes with a variance rate of $(\sigma^2 r^{2\gamma})$. Although some controversy has emerged regarding the appropriate functional form of the drift in (1), most studies focus their attention on the ability of each model to capture the volatility of interest rates.[9] The reason for this is that volatility is a key variable governing the value of contingent claims such as interest rate options; moreover, optimal hedging strategies for risk-averse investors depend critically on the volatility. In this sense, given that the functional form of the drift has been catalogued of second order in explaining interest rate dynamics, the present study focuses on the correct specification of the volatility.[10]

CKLS conclude that the models that best describe the dynamics of interest rates are those that allow the conditional volatility of rate changes to be highly dependent on the level of rates. Their γ estimate was 1.5. Therefore, at least for the US, commonly used models such as Vasicek (1977) and CIR-SR (1985) perform poorly relative to less well-known models such as Dothan (1978) and CIR-VR (1980). CKLS pointed out that the critical parameter in capturing interest rate dynamics was γ, and that evidence of a mean reversion, measured by the β parameter, was at best weak. Moreover, Chan et.al. (1992b) found in an application to Japan that, in spite of the relatively low volatility of rates, the dependence of the conditional volatility on rate levels was stronger than in the US with a γ estimate of 2.44. Overall, the general conclusion was that models assuming $\gamma < 1$ perform poorly when compared with models suggesting $\gamma > 1$.

An international study by Tse (1995) applied the CKLS framework to eleven countries and showed that the conditional volatility of interest rates can be very sensitive to their levels

[8] This has been catalogued as the "level effect" in the literature. For example, if $\gamma = 1.5$, it implies that a 1% increase in levels of rates will mean a 1.5% increase in volatility of rates during that observation period. (At higher γ, the volatility is more sensitive to interest rate levels).

[9] See for example, Ait-Sahalia (1996), Stanton (1997) and Conley et.al.(1997). For a discussion on whether the correct functional form of the drift rate in interest rate processes should be linear or not.

[10] See Nowman (1997).

(high γ), as in the cases of France, Holland and the USA. However, other countries like Canada, Italy, Switzerland and the UK exhibit low elasticities of the variance to the level of rates (low γ). Therefore, the Vasicek model may be preferred for these countries. For Australia, Belgium, Germany and Japan, they found a moderate elasticity of variance, suggesting that there was no clear-cut statistical evidence for the choice between the CIR-SR model and the GBM. Another international study by Dahlquist (1996) found that in the unrestricted model (1), the value of γ was less than one for all the countries studied except for Sweden, where the value of γ was 1.15. And the hypothesis of constant variance ($\gamma = 0$) could not be rejected for Germany and the UK. Dahlquist generally advocates the Brennan-Schwartz (1980) model for Denmark and Sweden, and the Vasicek (1977) and CIR-SR models for Germany and the UK respectively. Nowman (1997) also finds support for the CIR-SR model in British interest rate, reporting an insignificant value for γ of only 0.29, but Nowman confirms the CKLS result for the US with a γ estimate of 1.36.[11]

Evidence for Australia has been provided by Gray (1996) who found a γ estimate of about 1.5. His findings have been supported by Brailsford and Maheswaran (1998) who obtain a γ estimate of 1.7 and conclude that models allowing volatility to be highly sensitive to rate levels perform the best in Australia. Nowman (1998), using Euro-Currency interest rate series and an alternative econometric approach based on maximum likelihood estimation, found γ estimates of 1.05 and 0.98 for US and Japan respectively; these estimates contrast with CKLS and Chan et.al.(1992b) results. More recently, Nowman (2002) using four alternative time-series for Japan obtained estimates of γ ranging from 0.06 to 0.35 and suggests a low elasticity. These mixed results suggest that the estimates are sensitive to the data series and the estimation method used. Also, Nowman (1998) finds that France and Italy exhibit high elasticity of variance with γ estimates of 2.8 and 2.2 respectively. Mc Manus and Watt (1999) provided evidence for Canada and US, their γ estimates were 0.44 and 1.53 respectively and they support previous findings of Tse (1995) and CKLS.

Finally, Hiraki and Takezwa (1997) extended the analysis over a range of maturities varying from seven days to twelve months in Japan, and found that volatility is sensitive to the levels of rates with the same strength across maturities; reporting γ estimates of 0.4 and 0.5 for weekly and biweekly data respectively. However, Byers and Nowman (1998) in an application to the UK and the US found that the strength of the elasticity was different across maturities; in particular, they suggest that the elasticity is weaker as maturity increases, with γ estimates ranging from 1.8 to 1.3 in the UK, and from 1.95 to 1.16 in the US.

In general, evidence suggests that the conditional volatility of changes in interest rates is sensitive to the level of rates. However, the strength of this relationship appears to be country specific. For example, most studies agree that elasticity is high for the US $(\gamma > 1)$. Canada exhibits consistent evidence suggesting low elasticity of variance $(\gamma < 0.5)$. The UK exhibits mixed results, while most of the studies suggest low elasticity, Byers and Nowman (1998) found a high elasticity.

In Japan, Chan et.al. (1992b) found a high elasticity while Tse (1995) and Nowman (1998) found a moderate elasticity, but Hiraki and Takezwa (1997) and Nowman (2002) suggest low elasticities. The same controversy arises from Australia, for which Gray (1996) and

[11] Nowman (1997) used the same data set for the US as in the CKLS study. However, he proposed an alternative discrete approximation discussed later in the text.

Brailsford and Maheswaran (1998) using weekly and daily data respectively found high elasticities. However, Tse (1995) using monthly data found a moderate elasticity. These mixed results suggest that the elasticity estimates are highly sensitive to the interest rate series and data frequency used in the empirical application. Results for Italy found high elasticity using maximum likelihood estimation (ML) and low elasticity using generalized methods of moments estimation (GMM), suggesting that estimates could be very sensitive to the econometric approach used.

In short, the empirical evidence on "Level Models" supports a relationship between the volatility of rates and the level of rates. However, this dependence appears to be country specific. Moreover, estimates of elasticity parameters might be sensitive to the interest rate series, data frequency and econometric techniques applied in empirical work. But is it correct to allow volatility of the short-rate to vary only with the level of the short-rate? Does this assumption is very restrictive or holds empirically? This issue is explored in the next section.

2.3. Incorporating Time-Varying Volatility in Short-Rate Dynamics

So far we have discussed the properties and empirical evidence about "Level Models". However, these models have been generally criticized for their restrictive assumptions about the behavior of volatility.[12] For instance, a high value of γ implies a high degree of sensitivity to the level of interest rates yet periods of high interest rates but relative stability have been observed in various markets.[13] Similarly, periods of low interest rates but relatively high volatility have also been observed.[14] These observations contradict the basic assumption of volatility being solely determined by the level of rates. Further, BHK point out that volatility is restricted to be a function of only the level of interest rates, yet in practice news arrival is likely to have a significant impact on movements in rates.

In order to capture this "news effect", GARCH models have been applied to interest rates data as follows

$$r_{t+1} - r_t = \alpha + \beta r_t + \varepsilon_{t+1} \qquad \varepsilon_{t+1} \sim N(0, \sigma^2_{t+1}) \tag{3}$$

$$\sigma^2_{t+1} = \alpha_0 + \alpha_1 \varepsilon^2_t + b\sigma^2_t \tag{4}$$

this model is the GARCH (1,1), in which this period's volatility is a function of last period's unexpected news and conditional volatility. The model can still capture mean reversion in the series by including r_t as a regressor. The conditional heteroskedasticity captures leptokurtosis in the unconditional distribution of interest rates. However, the major problem with fitting GARCH models to short-term interest rate data is t hat the estimated parameters often imply an explosive volatility process such that the conditional variance process (σ^2_{t+1}) is not covariance-stationary. These conditions are met when the sum of the conditional variance parameters $(\alpha_1 + b)$ exceed unity (with $\alpha_0 > 0$). For example, Engle et.al. (1987) find $(\alpha_1 + b) > 1$, Hong (1988) finds $(\alpha_1 + b) = 1.073$, Engle et.al. (1990) find $(\alpha_1 + b) = 1.0096$, BHK find $(\alpha_1 + b) = 1.0011$, Koedijk et.al. (1997) find $(\alpha_1 + b) = 1.14$ and Bali (2003) finds $(\alpha_1 + b) = 1.048$; all for US. Given these results, it appears that GARCH modeling for interest rate dynamics is not the most appropriate representation. Also, this kind of models completely ignores the "level effect" (discussed in the previous section) that has been found significant in explaining volatility of interest rates. Finally, GARCH models permit negative interest rates.

[12] See Engle and Ng (1993a)
[13] For example, the period between 1983 and 1984 in the US.
[14] For example, the period between 1992 and early 1993 in the US.

In this sense, BHK developed a model which incorporates the effects of both levels and information shocks. They introduce a time-varying parameter model which nests both the "Level Models" (as in equation 1) and the GARCH model (as in equations 3 and 4). This class of models is called "Level-News Models". The model specification is

$$r_{t+1} - r_t = \alpha + \beta r_t + \varepsilon_{t+1} \tag{5}$$

$$E(\varepsilon_{t+1}) = 0 \qquad E(\varepsilon_{t+1}^2) = \sigma_{t+1}^2 r_t^{2\gamma} \tag{6}$$

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + b\sigma_t^2 \tag{7}$$

the model incorporates volatility conditioned on both the level of interest rates and a GARCH process.[15] Information shocks will enter the system through the lagged squared error term (ε_t^2) which flows through to have an impact on volatility. The model collapses to the "Level Models" as in equation (1) when the conditional variance equation (7) is a constant, i.e., when $\alpha_1 = b = 0$. Similarly, when $\gamma = 0$ then the model collapses to the GARCH process as in equations (3)-(4).

Following the work of Nelson (1991), Engle and Ng (1993b) and Glosten et.al. (1993), BHK allow for an asymmetry between negative and positive shocks in volatility through modifying the conditional variance specification given in (7) to include an asymmetric term. The alternative conditional variance specification is given by

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \alpha_2 \varepsilon_t^2 \eta_t + b\sigma_t^2 \tag{8}$$

where: $\eta_t = 1$ if $\varepsilon_t < 0$, and $\eta_t = 0$ otherwise, the parameter α_2 in (8) is a measure of the difference in the slope coefficient between negative and non-negative shocks. In equation (8), if $\alpha_2 > 0$ and statistically significant then negative shocks have a larger impact on volatility than positive shocks.

BHK showed that, at least for the US, the dependence of interest rates volatility on levels has been exaggerated in the literature. Their findings contrast with CKLS conclusion that, "the relation between interest rate volatility and the level of rates is the most important feature of any dynamic model of the short-term risk less rate" (p.1217). By comparing "Level Models" with "Level-News Models" they found that, while the "level effect" is important, adequately modeling the volatility parameter as a function of unexpected news is equally important. They showed that the estimates of γ in "Level-News Models" are lower than the γ estimates in "Level Models". Also, they suggest that "Level-News Models" provide a better description of volatility in interest rates. Therefore, they conclude that "Level Models" are misspecified in the way they model volatility. In this sense, the new class of "Level-News Models" appeared to be a potentially fruitful path by which to improve "Level Models" and capture adequately interest rate dynamics. The findings of BHK for the US have been supported in similar studies by Andersen and Lund (1997) and Koedijk et.al. (1997). These studies arrive to the same conclusions as BHK.

Brailsford and Maheswaran (1998) applied the BHK framework to Australia and support the view that allowing GARCH effects in interest rate modeling, reduces significantly the impact of the "level effect". They also find that "Level-News Models" are superior in capturing the dynamics of the volatility process. Their findings encourage the great potential of this relatively new class of models.

Recently, Bali (2003) applied the BHK framework to the US and assessed forecasting power of "Level-News Models". His findings suggest that "Level-News Models" outperform "Level Models" in forecasting volatility of interest rate changes. Moreover, using Monte Carlo

---

[15] Note that the "Level-News Models" are single-factor models, because only one source of uncertainty, r_t, appears in the mean equation (5) and this same source of uncertainty drives the GARCH behavior of the parameter σ^2_{t+1}.

simulations for yields of three and six month zero-coupon bonds, he found that incorporating the level and news effects in volatility improves the pricing performance of the interest rate models. Therefore, it appears that "Level-News Models" are not only better in fitting historical data and forecasting the volatility of rates, but also in pricing financial instruments. These encouraging results for the US motivate further research on this type of models in non US countries.

III. The Econometric Approach

3.1. The Estimation Method

Following BHK and Nowman (1997), maximum likelihood estimation (ML), which assumes a conditional normal distribution of error terms, is used. The utilization of ML instead of a distribution-free estimator such as the generalized method of moments (GMM) of Hansen (1982), is justified by three main reasons. First, as BHK suggest, statistical tests based on ML estimators tend to be more powerful than tests based on GMM estimators.[16] This might be potentially important in small samples. Second, Broze et.al. (1995) proved in their proposition 3.3 that the GMM estimator is not well behaved when $\gamma > 1$. Empirical evidence for the UK is mixed, with some studies suggesting a low elasticity of variance with $\gamma < 0.5$, while Byers and Nowman (1998) found high elasticity with $\gamma > 1$. Therefore, using ML avoids the possibility of using a not well behaved estimator. Third, Nowman (1997) showed that the ML estimator is more efficient than the GMM estimator, permitting us to perform more precise estimations and tests.[17]

3.2. The Short-Rate Models

The discrete approximation of the continuous-time process (1) used in CKLS and BHK is

$$r_{t+1} - r_t = \alpha + \beta r_t + \varepsilon_{t+1} \qquad E(\varepsilon_{t+1}) = 0 \tag{9}$$

where ε_{t+1} is the innovation or information shock at time *t+1*, and E(·) is the conditional expectations operator. We should note that the discretized process in (9) is only an approximation of the continuous-time specification in (1). The reason is that changes in rates are measured over discrete intervals of time and we use discrete data to estimate a model settled in continuous-time. Therefore, the "temporal aggregation bias", described in Grossman et.al. (1987) arises. However, the amount of approximation error introduced can be shown to be of second order importance if rate changes are measured over short periods of time.[18]

Alternatively, instead of using the CKLS discrete approximation (9), Nowman (1997) derived an exact discrete model for estimating the SDE (1).[19] This exact ML estimator has the potential to reduce the "temporal aggregation bias". The model specification is

$$r_{t+1} = e^{\beta} r_t + \frac{\alpha}{\beta}(e^{\beta} - 1) + \varepsilon_{t+1} \qquad E(\varepsilon_t \varepsilon_s) = 0 \qquad (t \neq s) \tag{10}$$

[16] We understand by poor power properties, that hypothesis tests under reject the null when in fact it is false. Therefore, an adequate power implies not rejecting the null in concordance with the probability of committing error type II adopted for the test.
[17] We understand by efficiency, that an estimator exhibits minimum variance.
[18] See Campbell (1986) and CKLS.
[19] See Appendix I for a mathematical derivation of Nowman's exact discrete model.

Now, from Nowman's exact discrete model (10), we estimate the relevant parameters by assuming the following conditional variances, $E(\varepsilon_{t+1}^2) = h_{t+1}$, specifications:

Model 1: (CKLS)

$$E(\varepsilon_{t+1}^2) = h_{t+1} = \frac{\sigma^2}{2\beta}(e^{2\beta} - 1)r_t^{2\gamma} \tag{11}$$

Model 2: (GARCH)

$$E(\varepsilon_{t+1}^2) = h_{t+1} = \alpha_0 + \alpha_1\varepsilon_t^2 + bh_t \tag{12}$$

Model 3: (GJR-GARCH)

$$E(\varepsilon_{t+1}^2) = h_{t+1} = \alpha_0 + \alpha_1\varepsilon_t^2 + \alpha_2\varepsilon_t^2\eta_t + bh_t \tag{13}$$

where: $\eta_t = 1$ if $\varepsilon_t < 0$, and $\eta_t = 0$ otherwise.

Model 4: (BHK1)

$$E(\varepsilon_{t+1}^2) = h_{t+1} = \frac{\sigma_{t+1}^2}{2\beta}(e^{2\beta} - 1)r_t^{2\gamma} \tag{14}$$

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1\varepsilon_t^2 + b\sigma_t^2 \tag{15}$$

Model 5: (BHK2)

$$E(\varepsilon_{t+1}^2) = h_{t+1} = \frac{\sigma_{t+1}^2}{2\beta}(e^{2\beta} - 1)r_t^{2\gamma} \tag{16}$$

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1\varepsilon_t^2 + \alpha_2\varepsilon_t^2\eta_t + b\sigma_t^2 \tag{17}$$

where: $\eta_t = 1$ if $\varepsilon_t < 0$, and $\eta_t = 0$ otherwise.

Model 1 embodies the assumption that it mimics the continuous-time model in (1). The specification closely parallels the continuous-time dynamics of (1) since it allows the variance of interest rate changes to depend directly on the level of the interest rate; this is the CKLS model that belongs to the "Level Models" class.

Model 2 is the well-known GARCH (1,1) model that accounts for time-varying volatility driven by the parameters α_1 and *b*. Model 3 is the GJR-GARCH model that is obtained by including an asymmetric term in the conditional variance specification; the parameter α_2 measures the differential impact between positive and negative shocks. If α_2 is found positive and statistically significant, it will imply that bad news have a greater impact on rates volatility than good news.

Model 4 is the first type of "Level-News Models" analyzed; it combines the "level effect" and the "news effect" by including time-varying volatility in the specification. We will refer to this model as the BHK1 model. Finally, Model 5 adds an asymmetric term in the conditional variance specification and nests all of the previous models. Consequently, it provides an ideal vehicle for model comparisons. The relative importance of the level of interest rates versus information shocks can be evaluated by imposing suitable parameter restrictions in the coefficients of Model 5 (called BHK2 model). For example, setting $\alpha_2 = 0$ produces the BHK1 model. Alternatively, setting $\alpha_1 = \alpha_2 = b = 0$ produces the CKLS model. Setting $\gamma = 0$ yields the GJR-GARCH model. Furthermore, all of the CKLS nested models can also be obtained from the BHK2 model by imposing the appropriate parameter restrictions illustrated in Table 3.2.1

TABLE 3.2.1
Parameter Restrictions Imposed by Alternative Models of The Short-Term Interest Rate

Alternative models of the short-term riskless rate of interest r can be nested with appropriate parameter restrictions within the unrestricted model BHK2

$$r_{t+1} = e^{\beta} r_t + \alpha/\beta\,(e^{\beta}-1) + \varepsilon_{t+1} \qquad E(\varepsilon_t\varepsilon_s) = 0 \quad (t\neq s)$$
$$E(\varepsilon^2_{t+1}) = h_{t+1} = \sigma^2_{t+1}/2\beta\,(e^{2\beta}-1)\, r_t^{2\gamma}$$
$$\sigma^2_{t+1} = \alpha_0 + \alpha_1\,\varepsilon^2_t + \alpha_2\,\varepsilon^2_t\,\eta_t + b\,\sigma^2_t$$

Model	α	β	γ	α_0	α_1	α_2	b
BHK1						0	
CKLS					0	0	0
GJR-GARCH			0				
GARCH			0			0	
Merton		0	0		0	0	0
Vasicek			0		0	0	0
CIR SR			0.5		0	0	0
Dothan	0	0	1		0	0	0
GBM	0		1		0	0	0
Brennan-Schwartz			1		0	0	0
CIR VR	0	0	1.5		0	0	0
CEV	0				0	0	0

As we can appreciate, BHK2 is the "unrestricted" model from which all evaluations of the competing models can be obtained. BHK1, CKLS, GJR-GARCH, GARCH and all of the CKLS nested models are restricted versions of BHK2.

3.3. The Diagnostic Tests

The models we estimate are evaluated with the likelihood ratio test statistic, the Bergstrom's (1990) "*S*" statistic and the Ljung-Box (1978) "*Q*" statistic. In addition, we use the Wooldridge's (1990) Conditional Moment Test which allow us to identify possible sources of misspecification in the competing models. In general, the null of this test states that the model is correctly specified regarding the source of misspecification being tested. Our sources of misspecification for the conditional moment tests will be the lagged interest rate level,

$$\lambda_{1,t} = r_t \tag{18}$$

this moment will identify short-rate models that misrepresent the dependence of volatility on interest rate levels.

The second indicator is

$$\lambda_{2,t} = \varepsilon_t\,\eta_t \tag{19}$$

where $\eta_t = 1$ if $\varepsilon_t < 0$, and $\eta_t = 0$ otherwise. In the same way, this moment will identify models misrepresenting the asymmetric response of volatility to news.

The third set of indicators focuses on serial correlation in squared standardized residuals,

$$\lambda_{3,t} = v_t \tag{20}$$
$$\lambda_{4,t} = v_{t-1} \tag{21}$$
$$\lambda_{5,t} = v_{t-2} \tag{22}$$
$$\lambda_{6,t} = v_{t-3} \tag{23}$$

here, the misspecification indicators are lagged values of the generalized residuals.[20] These moments test for remaining GARCH effects in the models.[21]

3.4. Forecasting Ability Test

To evaluate the volatility forecasting ability of the competing models, we proxy for the "true" or ex-post volatility as the absolute value of changes in interest rate via

$$v_{t+1} = \left| r_{t+1} - r_t \right| \tag{24}$$

then, we use the models' parameter estimates and conditional volatility specifications to construct a one step-ahead forecast of volatility, $\sqrt{\hat{h}_{t+1}}$, for each observation in the sample and for 22 observations out of sample. Using these, we calculate two mean squared forecast errors measures, one for observations in-sample and another for the out-of-sample observations. The mean-square-forecast error is

$$MSFE = \frac{1}{K} \sum_{t=0}^{K-1} \left(v_{t+1} - \sqrt{\hat{h}_{t+1}} \right)^2 \tag{25}$$

where K is equal to the sample size for the in-sample measure, and 22 for the out-of-sample measure. Finally, we compute the proportion of the variance of absolute rate changes that can be explained by the models` conditional volatility estimates, denoted by R^2 and given by

$$R^2 = 1 - \frac{MSFE}{\frac{1}{K} \sum_{t=0}^{K-1} \left(v_{t+1} - v_{t+1} \right)^2} \tag{26}$$

We compute two R^2 measures, one for in-sample and another for the out-sample volatility forecasting ability.

IV. The Data

The short-term interest rate used in this study is the one-month Euro-Currency rate on UK currency deposits (middle rate) obtained from *Datastream*^{TM}. We use weekly data on a Wednesday to avoid missing observations and any week-day effect. The data cover the period from June 1983 to January 2003 giving a total of 1023 observations.

Table 4.1 reports descriptive statistics. It displays the mean, standard deviation, and first six autocorrelations of the rate and change in rates. We also report the augmented Dickey-Fuller (ADF) statistic of Said and Dickey (1984) for the presence of a unit root. The average level of the rates is 8.48% with a standard deviation of 3.2% on an annual basis. The autocorrelations for the level fall off slowly and those of the first differences are small and not systematically positive or negative. The ADF statistic does not reject the null hypothesis of a unit root at the 5% significance level for the rates, but it does for the first differences, which suggests that the series is integrated of order one (difference stationary series.).

---

[20] Generalized residuals are functions of the data and the model parameters which are constructed to have zero conditional expectation if the model is correctly specified. In our case, if the variance equation of a competing model is correctly specified, we would expect $E(\varepsilon^2_{t+1} - h_{t+1} \mid \omega_t) = 0$, where ω_t is the information set available at time t. Hence, $v_t = \varepsilon^2_t - h_t$, will be our generalized residual.

[21] See BHK for a didactic explanation about the process of the Conditional Moment Test of Wooldridge (1990).

TABLE 4.1.

Summary Statistics

Means, standard deviations, and autocorrelations of weekly 1-month UK Euro-Currency rates and first differences are computed from June 1983 to January 2003.The variable r(t) denotes the 1-month UK Euro-Currency rate and Δ r(t) is the weekly change. ρ_j denotes the autocorrelation coefficient of order j. N represents the number of observations used. ADF denotes the Augemented Dickey-Fuller unit root statistic with a 5 percent critical value of -3.417.

Variable	N	Mean	Standard Deviation	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	ρ_6	ADF
r(t)	1023	8.482%	3.199%	0.997	0.994	0.991	0.987	0.984	0.980	-2.423
Δ r(t)	1022	-0.005%	0.257%	-0.074	0.063	0.073	-0.017	0.003	0.080	-10.4

In addition, we also collect 22 weekly observations out of sample, which will be used to asses volatility forecasting ability. These observations are from 22 January 2003 to 18 June 2003.

V. Empirical Results

5.1. Estimation Results and Model Comparisons

Table 5.1.1 reports maximum likelihood parameter estimates and diagnostic tests for the CKLS, GARCH, GJR-GARCH, BHK1 and BHK2 models estimated with the Nowman (1997) exact discrete model.

The models presented in Table 5.1.1, with exception of the GARCH model, exhibit an insignificant drift parameter (α). Also, mean reversion does not appear to be an important feature for UK interest rates dynamics. Even though, in most instances, the sign is the correct one, β is statistically indistinguishable from zero.[22]

This is important, given that most theoretical models of interest rates place strong emphasis on the mean reversion feature. From an economic viewpoint, mean reversion makes sense. When, for example, interest rates are high, economic activity and the demand for loans decline. This will put downward pressure on interest rates. The exact opposite happens when interest rates are low. Empirically, however, it turns out that this feature is not of particular consequence. Therefore, this initial finding suggests that any significant differences between the alternative models are caused by how they treat volatility.

As in Byers and Nowman (1998), estimates from the CKLS model suggest a high dependence of the volatility on the level of rates, with a γ estimate of 1.5386.[23] So, for the UK, the variance of unexpected interest rate changes is proportional to the cube of the level of interest rates. As a result, the CKLS model implies that as interest rates increase, volatility increases dramatically. The GARCH model, on the other hand, does not permit volatility to depend on interest rate levels, but instead allows volatility to change as news hit the market.[24] Therefore, in this model, no apparent relationship between the conditional volatility and the level of rates exists. In fact, the correlation between rate levels and GARCH volatility is only 0.311, while the correlation between rate levels and the CKLS volatility is 0.993. Notice, however, that the GARCH model exhibits an explosive behavior of the conditional variance, as $\alpha_1 + b = 1.2453$, both parameters being statistically significant. This finding is consistent with previous studies

[22] Mean reversion exists if $\beta < 0$, so a test for mean reversion is a test of whether $\beta = 0$ against the alternative that $\beta < 0$. However, under the null of no mean reversion, r_{t+1} has a stochastic trend, implying that the usual t-test is inappropriate.

[23] This estimate is not statistically different from the 1.5 estimate of CKLS.

[24] Since GARCH models assume $\gamma = 0$, no relation between the volatility and the level of rates is observed.

TABLE 5.1.1

Gaussian Estimation Using Nowman (1997) Exact Discrete Model

Statistical Models of the Short-Term Interest Rate:

Weekly 1-Month UK Euro-Currency Rates, 15/06/1983 - 15/01/2003

	CKLS	GARCH	GJR-GARCH	BHK1	BHK2
Panel 1.					
α	0.0002	0.0483	-0.0416	-0.0001	0.0011
	(0.991)	(<0.001)**	(0.175)	(0.99)	(0.932)
β	-0.0005	-0.0089	0.0067	-0.0004	-0.0004
	(0.859)	(<0.001)**	(0.099)	(0.817)	(0.837)
α_0	0.0001	0.0025	0.0022	0.0002	0.0002
	(0.013)*	(<0.001)**	(<0.001)**	(0.007)**	(<0.001)**
α_1		0.6152	0.7578	0.021	0.028
		(<0.001)**	(0.315)	(<0.001)**	(0.146)
α_2			0.1902		-0.0209
			(0.81)		(0.351)
b		0.6301	0.5494	0.5637	0.5471
		(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**
γ	1.5386			0.7759	0.7983
[Standard Error]	[0.134]			[0.091]	[0.036]
	(<0.001)**			(<0.001)**	(<0.001)**
Panel 2.					
LL	169.6	236.41	245.42	303.57	309.03
$\chi^2_{(\text{\# restrictions})}$	278.85	145.23	127.22	10.92	
	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	
$S\ (\varepsilon_t/h_t^{0.5})$	21.69	24.07	22.19	15.64	15.47
	(0.041)*	(0.02)*	(0.036)*	(0.209)	(0.217)
$Q\ (\varepsilon_t/h_t^{0.5})$	22.06	19.27	11.61	16.39	15.84
	(0.037)*	(0.082)	(0.071)	(0.174)	(0.199)
$Q\ (\varepsilon^2_t/h_t)$	26.09	4.3	4.01	5.67	6.21
	(0.01)**	(0.977)	(0.983)	(0.932)	(0.905)
Rate Level (λ_1)	4.215	0.0002	0.001	0.091	0.107
	(0.04)*	(0.988)	(0.979)	(0.763)	(0.744)
Asymmetry (λ_2)	0.014	0.0003	0.0005	0.028	0.002
	(0.905)	(0.986)	(0.983)	(0.867)	(0.964)
GARCH ($\lambda_3 - \lambda_6$)	17.357	0.002	0.001	0.201	0.088
	(0.002)**	(0.999)	(0.999)	(0.995)	(0.999)
$R^2_{(\text{in-sample})}$	0.247	0.288	0.274	0.285	0.293
$R^2_{(\text{out-sample})}$	0.42	0.31	0.291	0.358	0.396

Columns 1, 2, 3, 4 and 5 of Panel 1 report the maximum likelihood estimates of the model,

(1) $$r_{t+1} = e^{\beta} r_t + \alpha/\beta\ (e^{\beta} - 1) + \varepsilon_{t+1} \qquad E(\varepsilon_t \varepsilon_s) = 0 \quad (t \neq s)$$

(2) $$E(\varepsilon^2_{t+1}) = h_{t+1} = \sigma^2_{t+1}/2\beta\ (e^{2\beta} - 1)\ r_t^{2\gamma}$$

(3) $$\sigma^2_{t+1} = \alpha_0 + \alpha_1\, \varepsilon^2_t + \alpha_2\, \varepsilon^2_t\, \eta_t + b\, \sigma^2_t$$

where: $\eta_t = 1$ if $\varepsilon_t < 0$, else $\eta_t = 0$. P-values are in parentheses.(*Significant at the 5% level, **Significant at the 1%level.)

Panel 2 reports: log-likelihood values (LL); log-likelihood ratio tests of the alternative models against the "unrestricted" BHK2; Bergstrom's (1990) S tests for up to twelfth order serial correlation in standardized residuals ($S\ (\varepsilon_t/h_t^{0.5})$); Ljung-Box tests for up to twelfth order serial correlation in standardized and squared standardized residuals ($Q\ (\varepsilon_t/h_t^{0.5})andQ\ (\varepsilon^2_t/h_t)$,respectively); λ_1-λ_6 are a set of Wooldridge's (1990) conditional moment tests discussed in the text; the R^2 measures of volatility forecasting ability in- and out-of-sample. P-values are in parentheses. (*Significant at the 5% level, **Significant at the 1%level.)

trying to fit GARCH models to interest rate series, and supports the inappropriateness of this type of models for interest rates.[25]

By augmenting the conditional variance specification of the GARCH model with an asymmetric term, we obtain a GJR-GARCH. The β estimate suggests that mean reversion is insignificant. The lagged information shocks parameter (α_1) increases from 0.6152 in the GARCH model to 0.7578 in the GJR-GARCH model, but it becomes insignificant. This means that including an asymmetric term increases the impact of lagged information shocks, but makes it statistically indistinguishable from zero. Moreover, the asymmetry coefficient (α_2) is also insignificant, suggesting that the inclusion of it in the conditional variance specification is not only irrelevant, but also distorts the effect of information shocks that was previously found significant. An insignificant α_2 coefficient also suggests that asymmetric response of interest rates volatility to news in the UK is, at best, weak. The lagged conditional variance parameter (*b*) remains statistically significant, but its value drops from 0.6301 to 0.5494. In short, we can imply that significant estimates of the GARCH parameters α_1 and *b*, suggest that information shocks play a major role in the determination of the conditional variance for UK interest rates. Likewise, serial correlation in the conditional variances is an important feature of the volatility process.

We have seen that both, the level effect (measured by γ) and the GARCH effect (measured by α_1 and *b*) are statistically significant. However, the relative significance of these effects is an important question that can be addressed by considering the gains in terms of the value of the log-likelihood between the CKLS and the GARCH models.[26] Calculated changes in the log-likelihood are positive and substantial, from 169.6 in CKLS to 236.41 in GARCH, and 245.42 in GJR-GARCH. This suggests that GARCH models provide a better fit than CKLS. Further useful insights can be gained from Figure 5.1.1, which graphs ex-post volatility (measured by the absolute value of changes in rates), along with the conditional volatilities from the CKLS and the GARCH models.

Figure 5.1.1

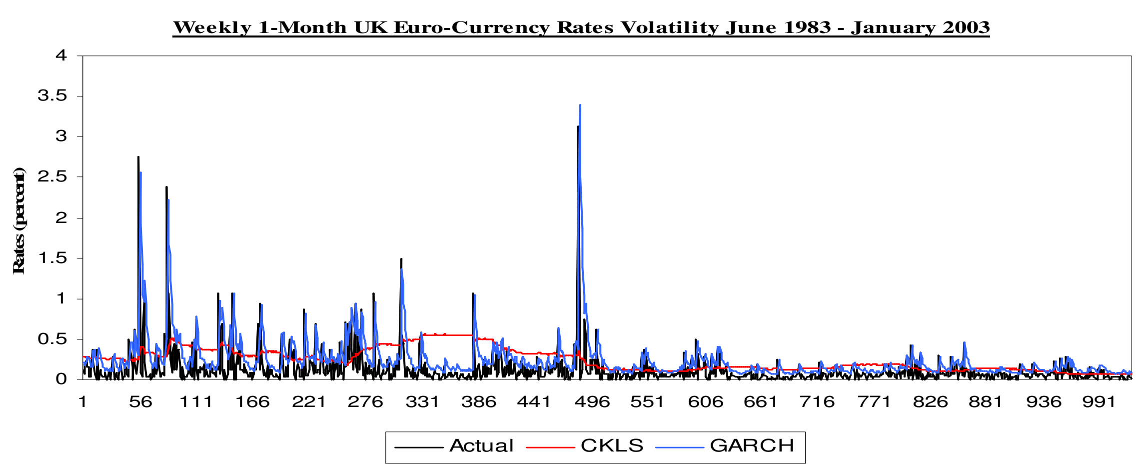

Figure 5.1.1. Volatility forecasts using the CKLS and GARCH models: This figure displays the time–series plot of the ex-post volatility, which is measured as the absolute value of weekly changes in the one-month UK Euro-Currency interest rates, and the volatility forecast, which is the square root of the conditional variance implied by the estimates of the CKLS and GARCH models from June 1983 to January 2003.

[25] See Engle et.al. (1987), Hong (1988), Engle et.al. (1990), BHK, Koedijk et.al. (1997), and Bali (2003).
[26] This is not a formal test since CKLS and GARCH models are not nested.

From Figure 5.1.1, we can observe that CKLS misrepresents realized volatility. In contrast, GARCH tracks realized volatility much better. From this, if one had to choose a model for the pricing of interest rate derivative securities, GARCH would be a better candidate. This provides evidence against models that rely exclusively on interest rate levels such as the CKLS. However, before we advocate any competing model for the short-rate, we must take care about the misspecification tests provided in Panel 2 of Table 5.1.1. First, if the model is correctly specified, the standardized residuals should have zero mean and unit variance, being serially uncorrelated. Also, a correctly specified model, should account for serial correlation in second moments, which are measured by the squared standardized residuals. As we can appreciate, for the CKLS model, the Bergstrom (*S*) test rejects the null that serial correlations up to the twelfth lag in standardized residuals are jointly zero; this is confirmed by the Ljung-Box (*Q*) test. Also, the *Q* test for up to twelfth order serial correlation in squared standardized residuals, suggests that the model fails to account fully for time variation in second moments. Moreover, the Wooldridge's conditional moment tests (CM) for rate levels (λ_1) indicates that the CKLS model fails to capture the dependence of volatility on levels, with a p-value of 0.04. This observation is surprising, since the CKLS model is supposed to capture the level effect, but it appears that ignoring GARCH effects causes a misspecification in the conditional volatility that leads to failure in adequately capturing the relationship between volatility and the level of rates. Finally, the CM test for remaining GARCH effects ($\lambda_3 - \lambda_6$), suggests that CKLS fails to capture the serial correlation in conditional variances.

CM tests on the GARCH and GJR-GARCH models reveal that both of them capture the level effect (λ_1) and the GARCH effect ($\lambda_3 - \lambda_6$). However, they fail to capture the serial correlation in standardized residuals, evidenced by the *S* tests which are significant, rejecting the null of white noise. We should note, however, that the CM tests for the asymmetry effect (λ_2), reveal that all of the models regardless on whether they incorporate an asymmetric term in their conditional volatility specifications, adequately account for asymmetries in interest rates dynamics. This is not surprising since we have found earlier that the asymmetric term in the GJR-GARCH model was insignificant; suggesting a weak asymmetric response of interest rates volatility in the UK. Evidence until now suggests that both, pure level models (the CKLS) and pure time-varying volatility models (the GARCH models) are misspecified in the way they model volatility of interest rates for the UK. Therefore, our results suggest that models incorporating both the dependence of volatility on levels and serial correlations in conditional variances may be superior to either the levels or GARCH models.

Results for the BHK1 and BHK2 models are presented in the last two columns of Table 5.1.1. Consider first the BHK1 model. Conditional volatilities are plotted in Figure 5.1.2. next page. These estimated volatilities seem to track realized volatility better than the CKLS model.

In the BHK1 model, the variance process differs from both the CKLS and the GARCH processes. The GARCH parameters, α_1 and *b*, are independently and jointly significantly different from zero, implying that the volatility parameter (σ) is time varying. Similarly, γ is highly significant, implying that the variance is an increasing function of levels. In fact, the correlation between the conditional volatility and the interest rate level is 0.447, which lies between the 0.311 of the GARCH model and the 0.993 of the CKLS model. Also, this model passes all of the volatility-related misspecification tests. It captures the dependence of volatility on levels (λ_1), the asymmetric response of volatility to news (λ_2), the GARCH effect ($\lambda_3 - \lambda_6$) and the serial correlation in standardized and squared standardized residuals; evidenced by insignificant CM, *S* and *Q* tests. This shows that BHK1 is statistically preferable to both the CKLS and GARCH models.

Figure 5.1.2

Weekly 1-Month UK Euro-Currency Rates Volatility June 1983 – January 2003

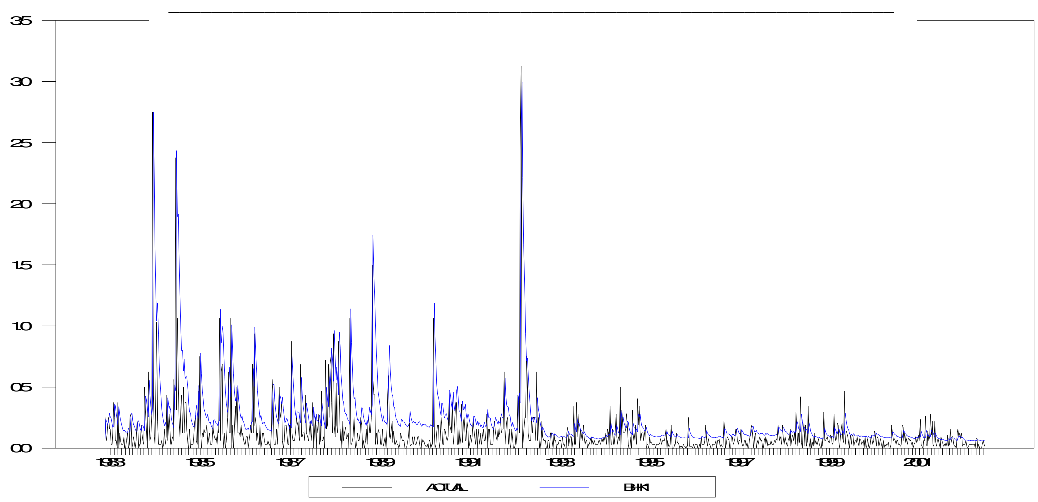

Figure 5.1.2. Volatility forecast using the BHK1 model: This figure displays the time–series plot of the ex-post volatility, which is measured as the absolute value of weekly changes in the one-month UK Euro-Currency interest rates, and the volatility forecast, which is the square root of the conditional variance implied by the estimates of the BHK1 model from June 1983 to January 2003.

An important observation is that α_1 has dropped substantially compared to the GARCH and GJR-GARCH models; causing $\alpha_1 + b$ to drop from 1.2453 in the GARCH model, to 0.5847 in BHK1. Implying that, in contrast to the GARCH model, the conditional variance process is finite and stationary.[27] Therefore, common findings in the literature of explosive conditional variance processes could be due in part, a consequence of a misspecification error caused by ignoring the relationship between volatility and rate levels.[28] Another interesting observation is that the γ estimate has dropped substantially from 1.5386 in the CKLS model, to 0.7759 in BHK1. This implies that extant findings of a high dependence of volatility on rate levels have been caused by a misspecification error, originated by ignoring the GARCH effect. Our findings suggest that, at least for the UK, the combination of level and news effects is important to fully capture the dynamic behavior of interest rates. Recall that the CKLS model fails to capture the dependence between volatility and rate levels, but after including the GARCH component - as in the BHK1 model - this dependence is fully captured with a γ estimate of almost half. This implies that ignoring GARCH effects overstates the value of γ and causes misspecification errors in short-rate models. Moreover, our γ estimate in the BHK1 model is statistically different from either 0.5 or 1. This suggests that, at least for the UK, time-varying volatility versions of theoretical models such as the CIR-SR that assumes γ = 0.5 or the Dothan (1978), GBM, and Brennan-Schwartz (1980) models which assume γ = 1, may not be reliable. It could be, however,

[27] This is only a conjecture since in the BHK1 model volatility persistence is no longer measured by $\alpha_1 + b$. Now volatility is a function of both the volatility parameter, σ^2_t, and interest rate levels.

[28] See Engle et.al. (1987), Hong (1988), and Engle et.al. (1990) for findings of explosive behaviors in conditional variances processes when fitting GARCH models to interest rates series.

that a time-varying volatility version of the CEV model (which leaves γ free), performs well for the UK. Our findings contrast with BHK results for the US in which they find - in the BHK1 model - a γ estimate of 0.459, not significantly different from 0.5, advocating an extended version of the CIR-SR model.

We now extend BHK1 in order to incorporate an asymmetric term; this yields BHK2 model, presented in the last column of Table 5.1.1. While BHK2 statistically dominates BHK1 (the likelihood ratio test χ^2 evidences a rejection of BHK1 against BHK2), most of the results discussed for BHK1 are unaffected. The model also passes all of the volatility-related misspecification tests. We should note, however, that the asymmetry coefficient, α_2, is insignificant; implying again a weak asymmetric effect of interest rates volatility in the UK. Perhaps more important, the γ estimate is similar to the BHK1 estimate. It rises slightly to 0.7983, but it is still statistically different from either 0.5 or 1; again some statistical evidence in support of a time-varying volatility version of the CEV model for the UK.

5.2. The CKLS Nested Models

We now turn our analysis to the CKLS nested models.[29] Parameter estimates and diagnostic tests are presented in Table 5.2.1. A common feature between the eight well-known nested models is that the mean equation parameters (α and β) are insignificant; suggesting again that any difference between the short-rate models will be given by the stochastic specification of the conditional variance. Firstly, we should note that all of these models are rejected against the "unrestricted" BHK2 by the likelihood ratio test (χ^2); implying that the imposed restrictions are not valid. Secondly, the *S* and *Q* tests for standardized residuals show that models assuming a γ parameter of either less or more than one, fail to account for the serial correlation in residuals; rejecting the null hypothesis of white noise. Moreover, all of the models fail to account for the serial correlation in second moments; as evidenced by the significant *Q* tests for squared standardized residuals. Also, all of them - except for the Dothan model - fail to capture the dependence between volatility and rate levels (λ_1); with significant CM tests.

None of these models account for the GARCH effect in interest rates volatility ($\lambda_3 - \lambda_6$), but all of them account for the asymmetric effect (λ_2). Another interesting observation is that the CEV model, which leaves the γ parameter free but imposes the restriction of $\alpha = 0$, reaches a γ estimate of 1.5384. This estimate in very similar to the CKLS model estimate of 1.5386; implying that the inclusion or not of α, the drift parameter, in the model does not affect the estimates of γ.

In short, evidence from the well-known CKLS nested models, confirms our findings that failure to account for the GARCH effect in the stochastic behavior of interest rates, originates diverse sources of misspecifications. However, the asymmetric effect is still captured without including a GARCH or Asymmetric GARCH term in the conditional variance specifications. This confirms our intuition that no asymmetric response of interest rates volatility to news in the UK exists or is at best weak.

[29] Note that all of these models belong to the "Level Models" class.

TABLE 5.2.1

Gaussian Estimation Using Nowman (1997) Exact Discrete Model

Statistical Models of the Short-Term Interest Rate:

Weekly 1-Month UK Euro-Currency Rates, 15/06/1983 - 15/01/2003

	Merton	Vasicek	CIR SR	Dothan	GBM	Brennan-Schwartz	CIR VR	CEV
Panel 1.								
α	-0.0054	0.0141	0.0071			0.0024		
	(0.517)	(0.646)	(0.773)			(0.877)		
β		-0.0023	-0.0015		-0.0005	-0.0008		-0.0005
		(0.33)	(0.695)		(0.488)	(0.711)		(0.602)
α_0	0.0661	0.0662	0.0066	0.0007	0.0007	0.0007	0.0001	0.0001
	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(0.014)*
α_1								
α_2								
b								
γ			0.5	1	1	1	1.5	1.5384
[Standard Error]								[0.079]
								(<0.001)**
Panel 2.								
LL	-62.11	-61.69	56.09	136.68	136.86	136.88	156.14	169.6
$\chi^2_{(\text{\# restrictions})}$	742.28	741.44	505.88	344.7	344.33	344.31	305.78	278.85
	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**
S $(\varepsilon_t/h_t^{0.5})$	28.05	28.07	23.27	20.2	20.06	20.14	33.33	21.68
	(0.005)**	(0.005)**	(0.026)*	(0.064)	(0.066)	(0.065)	(<0.001)**	(0.041)*
Q $(\varepsilon_t/h_t^{0.5})$	28.36	28.37	23.58	20.4	20.4	20.46	21.77	22.05
	(0.005)**	(0.005)**	(0.023)*	(0.06)	(0.06)	(0.059)	(0.04)*	(0.037)*
Q (ε^2_t/h_t)	72.73	72.84	58.66	41.68	41.4	41.36	27.1	26.09
	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(0.008)**	(0.01)**
Rate Level (λ_1)	8.803	5.131	12.462	0.979	10.542	6.938	5.425	5.015
	(0.003)**	(0.024)*	(<0.001)**	(0.322)	(0.001)**	(0.008)**	(0.02)*	(0.025)*
Asymmetry (λ_2)	0.16	0.103	0.024	0.002	0.001	0.015	0.079	0.048
	(0.69)	(0.748)	(0.878)	(0.965)	(0.975)	(0.904)	(0.78)	(0.827)
GARCH ($\lambda_3 - \lambda_6$)	21.356	19.158	16.953	16.649	18.135	17.767	16.202	17.846
	(<0.001)**	(<0.001)**	(0.002)**	(0.002)**	(0.001)**	(0.001)**	(0.003)**	(0.001)**
$R^2_{(\text{in-sample})}$	0.236	0.236	0.259	0.262	0.262	0.262	0.249	0.247
$R^2_{(\text{out-sample})}$	0.401	0.401	0.407	0.413	0.413	0.413	0.42	0.42

Columns 1, 2, 3, 4, 5, 6, 7 and 8 of Panel 1 report the maximum likelihood estimates of the model,

(1) $$r_{t+1} = e^{\beta} r_t + \alpha/\beta\,(e^{\beta} - 1) + \varepsilon_{t+1} \qquad E(\varepsilon_t \varepsilon_s) = 0 \quad (t \neq s)$$

(2) $$E(\varepsilon^2_{t+1}) = h_{t+1} = \sigma^2_{t+1}/2\beta\,(e^{2\beta} - 1)\, r_t^{2\gamma}$$

(3) $$\sigma^2_{t+1} = \alpha_0 + \alpha_1 \varepsilon^2_t + \alpha_2 \varepsilon^2_t \eta_t + b\,\sigma^2_t$$

where: $\eta_t = 1$ if $\varepsilon_t < 0$, else $\eta_t = 0$. P-values are in parentheses.(*Significant at the 5% level, **Significant at the 1%level.)

Panel 2 reports: log-likelihood values (LL); log-likelihood ratio tests of the alternative models against the "unrestricted" BHK2; Bergstrom's (1990) S tests for up to twelfth order serial correlation in standardized residuals (S $(\varepsilon_t/h_t^{0.5})$); Ljung-Box tests for up to twelfth order serial correlation in standardized and squared standardized residuals (Q $(\varepsilon_t/h_t^{0.5})$ and Q (ε^2_t/h_t), respectively); λ_1-λ_6 are a set of Wooldridge's (1990) conditional moment tests as discussed in the text; the R^2 measures of volatility forecasting ability in- and out-of-sample.

P-values are in parentheses. (*Significant at the 5% level, **Significant at the 1%level.)

5.3. An Alternative Measure of Model Performance

An alternative model evaluation criterion is the volatility forecasting power. The R^2 measures of forecasting ability are reported in the last two lines of the previous Tables. Surprisingly, the out-of-sample forecasting performance is better than the in-sample forecasting performance for all of the competing models. This observation is not expected; since the normal thing is a relatively lower out-of-sample forecasting ability. However, this could be due to the fact that during the out-of-sample period volatility has been relatively low compared with part of the in-sample period, especially the first half; this becomes evident from Figure 5.1.2. Therefore, the in-sample R^2 measure is lower due to the greater volatility experimented during the sample period used for the estimation.

As expected, the models accounting for the level and GARCH effects - like the BHK2 and BHK1 - have the best volatility forecasting ability in-sample; and models assuming no relationship between volatility and either levels or information shocks - like the Merton or Vasicek models - perform the worst. But, rather surprisingly, when assessing the out-of-sample forecasting power a different picture emerges. While BHK2 and BHK1 have a relatively good performance, models suggesting a high dependence of volatility on rate levels and ignoring GARCH effects like the CKLS, CIR VR, and CEV (that were found to be misspecified and rejected against BHK2) perform the best. This could be specific to the out-of-sample evaluation period, during which interest rates have been relatively stable and spikes in volatility haven't been observed. Therefore, from an out-of-sample perspective, level models may well account for the volatility dynamics by parameter zing them only as a function of rate levels. We should note, however, that during periods of relatively high volatility, models like BHK2 and BHK1 are expected to perform the best. In general, we can conclude that periods of low volatility can be well modeled by "Level Models", but it should be safer to rely on "Level-News Models" as they have been found to be statistically superior.

5.4. Extending the CKLS Nested Models

So far, our findings show that modeling the volatility of interest rates as a function of information shocks and levels is the most appropriate specification for short-rate dynamics in the UK. However, we have not been able to find any alternative to the BHK2 that, while incorporating news and level effects, offers a more parsimonious specification for short-rate dynamics. Remember that the only model correctly specified over the basis of volatility-related diagnostic tests was the BHK1, but it was rejected by the likelihood ratio test (χ^2). In this section, we explore whether there is an alternative model that offers a reliable description of interest rate dynamics in the UK. In addressing this question, we extend all of the CKLS nested models of Table 5.2.1 such that they additionally include GARCH effects in their conditional variances specifications.[30]

Table 5.4.1 reports parameter estimates and diagnostic tests for the extended models that, at least, passed all of the volatility-related tests. These models are the GJR-Dothan, GJR-GBM, G-GBM, GJR-Brennan-Schwartz, G-Brennan-Schwartz, GJR-CEV, G-CEV, and one new model assuming $\alpha = \beta = 0$ while leaving all of the remaining parameters free, we will call it the NO-DRIFT model.[31]

[30] The only model that was not extended is the Vasicek (1977), since an extension of this model in order to include time-varying volatility parameters will yield the GARCH model itself.

[31] We call "GJR-MODEL NAME" to models extended with GARCH and Asymmetric effects. And "G-MODEL NAME" to models extended with solely a GARCH effect.

TABLE 5.4.1

Gaussian Estimation Using Nowman (1997) Exact Discrete Model
Statistical Models of the Short-Term Interest Rate:
Weekly 1-Month UK Euro-Currency Rates, 15/06/1983 - 15/01/2003

	GJR-Dothan	GJR-GBM	G-GBM	GJR-Brennan Schwartz	G-Brennan Schwartz	GJR-CEV	G-CEV	NO-DRIFT
Panel 1.								
α				-0.0038	0.0032			
				(0.73)	(0.761)			
β		-0.0003	-0.0006	0.0002	-0.001	-0.0002	-0.0004	
		(0.708)	(0.302)	(0.889)	(0.528)	(0.737)	(0.468)	
α_0	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	0.0002
	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(0.006)**
α_1	0.0145	0.014	0.0075	0.0142	0.0075	0.0279	0.0209	0.0285
	(<0.001)**	(0.81)	(0.012)**	(<0.001)**	(<0.001)**	(0.178)	(0.104)	(0.002)**
α_2	-0.0133	-0.0129		-0.0131		-0.0208		-0.0211
	(<0.001)**	(0.125)		(<0.001)**		(0.258)		(<0.001)**
b	0.4711	0.4736	0.56	0.4733	0.5598	0.5465	0.5637	0.5469
	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**	(<0.001)**
γ	1	1	1	1	1	0.7996	0.7763	0.7981
[Standard Error]						[0.069]	[0.063]	[0.041]
						(<0.001)**	(<0.001)**	(<0.001)**
Panel 2.								
LL	305.88	305.99	298.1	306.06	298.15	309.02	303.57	308.94
$\chi^2_{(\text{\# restrictions})}$	6.29	6.08	21.86	5.94	21.76	0.01	10.92	0.17
	(0.098)	(0.048)*	(<0.001)**	(0.015)*	(<0.001)**	(0.92)	(0.004)**	(0.919)
$S\ (\varepsilon_t/h_t^{0.5})$	16.36	15.79	18.73	15.72	18.91	15.44	15.64	15.88
	(0.176)	(0.201)	(0.095)	(0.204)	(0.091)	(0.218)	(0.208)	(0.197)
$Q\ (\varepsilon_t/h_t^{0.5})$	16.26	16.06	19.65	15.98	19.79	15.82	16.39	15.96
	(0.179)	(0.189)	(0.074)	(0.192)	(0.071)	(0.199)	(0.174)	(0.193)
$Q\ (\varepsilon^2_t/h_t)$	9.12	9.06	7.08	9	7.07	6.21	5.67	6.21
	(0.693)	(0.697)	(0.852)	(0.703)	(0.853)	(0.905)	(0.932)	(0.905)
Rate Level (λ_1)	1.956	0.114	0.171	0.124	0.194	0.111	0.084	2.474
	(0.162)	(0.735)	(0.679)	(0.725)	(0.66)	(0.738)	(0.772)	(0.116)
Asymmetry (λ_2)	1.678	0.062	0.048	0.0005	0.051	0.039	0.041	3.499
	(0.195)	(0.801)	(0.826)	(0.982)	(0.821)	(0.843)	(0.839)	(0.061)
GARCH ($\lambda_3 - \lambda_6$)	0.624	0.001	0.02	0.061	0.304	0.0003	0.011	1.03
	(0.96)	(0.999)	(0.999)	(0.999)	(0.989)	(0.999)	(0.999)	(0.905)
$R^2_{(\text{in-sample})}$	0.289	0.29	0.284	0.29	0.284	0.293	0.285	0.292
$R^2_{(\text{out-sample})}$	0.411	0.411	0.37	0.411	0.37	0.396	0.358	0.395

Columns 1, 2, 3, 4, 5, 6, 7 and 8 of Panel 1 report the maximum likelihood estimates of the model,

(1) $$r_{t+1} = e^{\beta} r_t + \alpha/\beta\,(e^{\beta} - 1) + \varepsilon_{t+1} \qquad E(\varepsilon_t \varepsilon_s) = 0 \quad (t \neq s)$$

(2) $$E(\varepsilon^2_{t+1}) = h_{t+1} = \sigma^2_{t+1}/2\beta\,(e^{2\beta} - 1)\, r_t^{2\gamma}$$

(3) $$\sigma^2_{t+1} = \alpha_0 + \alpha_1 \varepsilon^2_t + \alpha_2 \varepsilon^2_t \eta_t + b\,\sigma^2_t$$

where: $\eta_t = 1$ if $\varepsilon_t < 0$, else $\eta_t = 0$. P-values are in parentheses.(*Significant at the 5% level, **Significant at the 1%level.)

Panel 2 reports: log-likelihood values (LL); log-likelihood ratio tests of the alternative models against the "unrestricted" BHK2; Bergstrom's (1990) S tests for up to twelfth order serial correlation in standardized residuals ($S\ (\varepsilon_t/h_t^{0.5})$); Ljung-Box tests for up to twelfth order serial correlation in standardized and squared standardized residuals ($Q\ (\varepsilon_t/h_t^{0.5})$ and $Q\ (\varepsilon^2_t/h_t)$, respectively); λ_1-λ_6 are a set of Wooldridge's (1990) conditional moment tests as discussed in the text; the R^2 measures of volatility forecasting ability in- and out-of-sample.

P-values are in parentheses. (*Significant at the 5% level, **Significant at the 1%level.)

The first column of Table 5.4.1 reports the Dothan model augmented with GARCH and asymmetric terms. We should note that this model is not rejected against the BHK2 by the likelihood ratio test (χ^2), which implies that the restrictions imposed are valid. These restrictions are $\alpha = \beta = 0$ and $\gamma = 1$, three in total. The parameters in the model are highly significant. This supports the view that the conditional mean parameters (α and β) are of second order in explaining short-rate dynamics, and can easily be ignored. Also, we can argue that restricting γ to be 1, in combination with GARCH and asymmetric terms, adequately captures short-rate dynamics in the UK. The model accounts for serial correlation in standardized and squared standardized residuals, the level effect (λ_1), the asymmetric effect (λ_2) and the GARCH effect ($\lambda_3 - \lambda_6$).

Columns 2 troughs 5 report the extended versions of the GBM and Brennan-Schwartz models. Although all of them pass the volatility-related diagnostic tests, they are rejected against the BHK2 by the likelihood ratio test (χ^2). We should note that all of these models assume $\gamma = 1$, just as the GJR-Dothan that was not rejected, but the difference is that they add one or both of the conditional mean parameters α and β. Therefore, we can imply that α and β are not only insignificant in explaining short-rate dynamics, but also they represent a source of misspecification when included in short-rate models; that leads to rejection of these models against the unrestricted BHK2.

Column 6 shows the GJR-CEV model. The only difference between this model and the BHK2 is that the former assumes $\alpha = 0$. We should note that the estimates from this model are almost the same as in BHK2, especially the γ parameter with a value of 0.7996. Also, the log-likelihood value is technically the same and the model is not rejected against BHK2, passing all the diagnostic tests. This confirms the intuition regarding the insignificance of the drift parameter, α, in explaining short-rate dynamics in the UK. Column 7 reports the G-CEV model. Although it passes the diagnostic tests, it is rejected against BHK2. This suggests that restricting jointly the drift term, α, and the asymmetric term, α_2, to be zero is not appropriate. However, we should note that the γ estimate of this model is qualitatively the same as the GJR-CEV, BHK2 and BHK1 models.

Finally, column 8 shows the NO-DRIFT model. This model passes all the diagnostic tests and is not rejected against BHK2. This implies that restricting the conditional mean parameters (α and β) to be both zero offers a reliable description of short-rate dynamics in the UK. This supports the insignificance of these two parameters and confirms the weak evidence of mean-reversion documented in the literature.[32] We should also note that the γ estimate is again qualitatively the same as in BHK2, being statistically different from either 0.5 or 1. This suggests that the most reliable value for γ in the UK is approximately 0.8, and that the CKLS model exaggerates the dependence of volatility on rate levels; evidenced by its estimate of 1.5386, which almost doubles the value of the most appropriate estimate given by 0.7983 in BHK2.

In short, it appears that the most reliable specification for short-rate dynamics in the UK is given by the combination of level and news effects. In addition, restrictions imposed on α and β parameters are valid, permitting us to reach more parsimonious and equally reliable models for the short-rate. This supports extant findings regarding weak evidence of mean reversion and enforces the view that interest rates volatility should not be modeled only as a function of the level of rates.

[32] See CKLS, BHK and Nowman (1997) for findings of weak mean-reversion in interest rate processes.

VI. Conclusion

In this study, we estimate and compare three classes of single-factor short-term interest rate models in the UK. In the first, volatility is only a function of interest rate levels. We call these "Level Models". In the second, volatility is only a function of information shocks to the interest rate market. These are GARCH type models. In the third, volatility is a function of both rate levels and information shocks. We call these "Level-News Models".

Our findings show that "Level Models" exaggerate the dependence of volatility on rate levels. They overstate the value of the parameter that measures the strength of this relation (γ). Therefore, ignoring GARCH effects appears to cause an omitted variables problem that leads "Level Models", such as the CKLS, to be misspecified. Moreover, Wooldridge's (1990) conditional moment tests reveal that "Level Models" fail to capture both the dependence of volatility on rate levels, and the serial correlation in conditional variances. On the other hand, GARCH models produce estimates suggesting an explosive behavior of the conditional variance. Therefore, ignoring the "level effect" also causes an omitted variables problem that leads GARCH processes to be explosives. Furthermore, Bergstrom's (1990) "*S*" tests reveal that GARCH models do not produce white noise standardized residuals; implying that they are misspecified.

We support the view that the combination of level and news effects yields the most appropriate description of the volatility of interest rates. This is evidenced by the "Level-News Models" that have been found to be correctly specified, capturing both the dependence of volatility on rate levels and the serial correlation in conditional variances. However, we disagree with the BHK statement that: "While the sensitivity of interest rate volatility to interest rate levels is important, adequately modeling volatility as a function of unexpected information shocks is equally important" (1996, p. 85). We show that GARCH models perform relatively better than "Level Models", suggesting that modeling volatility as a function of unexpected information shocks is more important than modeling it only as a function of rate levels.

We obtain – in the "Level-News Models" – an estimate for the parameter that measures the sensitivity of interest rate volatility with respect to the interest rate level, γ , of about 0.8. This estimate is significantly different from either 0.5 or 1, which does not support extended versions of theoretical short-rate models that restrict γ to be either 0.5 or 1. This contrasts with the results of BHK for US in which they find a γ estimate of 0.459 (statistically insignificantly different from 0.5), and support a time-varying volatility version of the CIR-SR model. We find, however, that an extended version of the CEV model, that includes a GARCH effect in the conditional variance specification, produces a reliable description of short-rate dynamics in the UK.

We find that the parameter that measures the speed of mean reversion, β , is statistically indistinguishable from zero. This provides weak evidence of mean reversion in interest rates dynamics and is consistent with extant literature.

We obtain insignificant estimates for the parameter that measures the asymmetric response of volatility to news, α_2 ; implying that volatility asymmetries in UK interest rates are at most weak. Furthermore, conditional moment tests reveal that the asymmetric effect is adequately captured by all of the models tested, no matter if they include or not GARCH or asymmetric terms in their conditional variances specifications. This finding strongly supports the weak asymmetric response of volatility to news in the UK.

Finally, we propose a short-rate model which restricts the mean reversion parameter, β , and the drift parameter, α , to be both zero. The proposed model appears to be well specified and is not rejected against the "unrestricted" BHK2; providing a parsimonious and reliable description of short-rate dynamics for the UK. This suggests that the parameters that describe the conditional mean of interest rate changes (α and β), are insignificant in capturing short-rate dynamics. Therefore, it appears that the most important feature in short-rate modeling is the correct specification of the conditional variance; leaving little relevance for the conditional mean characterization.

References

Aït-Sahalia, Y. (1996). "Testing Continuous-Time Models of the Spot Interest Rate." *Review of Financial Studies*, vol. 9, no. 2 (Summer), 385–426.

Andersen, T.G., Lund, J., (1997). "Estimating continuous-time stochastic volatility models of the short-term interest rate." *Journal of Econometrics*, Vol. 77, 343–377.

Bali, T.G., (2003). "Modeling the stochastic behavior of short-term interest rates: Pricing implications for discount bonds." *Journal of Banking & Finance*, Vol. 27, 201–228.

Ball, C.A., Torous, W.N., (1994). "A Stochastic Volatility Model for the Short-Term Interest Rates and the Detection of Structural Shifts." Working Paper, Vanderbilt University.

Bergstrom, A. R., (1984). "Continuous-time stochastic models and issues of aggregation over time." *Handbook of Econometrics*, Vol II. (Elsevier Science, Amsterdam).

Bergstrom, A. R., (1990). *Continuous Time Econometric Modelling*. Oxford: Oxford University Press.

Black, F. and Scholes, M. (1973). "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, vol. 81: 637 – 654.

Bliss, R.R., and D.C. Smith. (1998). "The Elasticity of Interest Rate Volatility: Chan, Karolyi, Longstaff, and Sanders Revisited." *Journal of Risk*, vol. 1, no. 1 (Fall):21–46.

Bollerslev, T., (1986). "Generalised Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics*, Vol. 33, 307-327.

Brailsford, T.J, Maheswaran, K., (1998). "The Dynamics of the Australian Short-Term Interest Rate." *Australian Journal of Management*, Vol. 23, 213-234.

Brennan, M., and E. Schwartz. (1979). "A Continuous-Time Approach to the Pricing of Bonds." *Journal of Banking and Finance*, vol. 3, no. 2 (July):133–155.

Brenner, R.J., R.H. Harjes, and K.F. Kroner. (1996). "Another Look at Models of the Short-Term Interest Rate." *Journal of Financial and Quantitative Analysis*, vol. 31, no. 1 (March):85–107.

Broze, L., Scaillet, O., Zakoian, J.-M., (1995). "Testing for continuous-time models of the short-term interest rate." *Journal of Empirical Finance*, Vol. 2, 199–223.

Byers, S.L., Nowman, K.B., (1998). “Forecasting UK and US Interest Rates Using Continuous Time Term Structure Models.” *International Review of Financial Analysis*, Vol. 7, 191–206.

Campbell, J.Y., (1986). “A Defense of Traditional Hypotheses About the Term Structure of Interest Rates.” *Journal of Finance*, Vol. 41, 183-193.

Chan, K., G. Karolyi, F. Longstaff, and A. Sanders. (1992a). “An Empirical Comparison of Alternative Models of the Short-Term Interest Rate.” *Journal of Finance*, vol. 47, no. 3 (July):1209–27.

Chan, K., G. Karolyi, F. Longstaff, and A. Sanders. (1992b). “The Volatility of Japanese Interest Rates: A Comparison of Alternative Term Structure Models.” *Pacific-Basin Capital Markets Research*, vol. 3:119–136.

Chapman, D.A., and N.D. Pearson. (2001). “Recent Advances in Estimating Term Structure Models” *Financial Analysts Journal*, Vol. 57, No. 3, July/August 2001:77-95.

Conley, T.G., L.P. Hansen, E.G.J. Luttmer, and J.A. Scheinkman., (1997). “Short-Term Interest Rates as Subordinated Diffusions.” *Review of Financial Studies*, Vol. 10, no. 3 (Fall), 525–577.

Cox, J. (1975). “Notes on Option Pricing I: Constant Elasticity of Variance Diffusions.” Working Paper, Stanford University.

Cox, J.C., J. Ingersoll, and S.A. Ross. (1980). “An Analysis of Variable Rate Loan Contracts.” *Journal of Finance*, vol. 35:389–403.

Cox, J.C., J. Ingersoll, and S.A. Ross. (1985). “A Theory of the Term Structure of Interest Rates.” *Econometrica*, vol. 53, no 2 (March):385–407.

Dahlquist, M., (1996). “On alternative interest rate processes.” *Journal of Banking & Finance*, Vol. 20, 1093-1119.

Dothan, U. (1978). “On The Term Structure of Interest Rates.” *Journal of Financial Economics*, vol. 7, no. 1 (January):229–264.

Duffee, G.R. (1993). “On the Relation Between the Level and Volatility of Short-Term Interest Rates: A Comment on CKLS.” Working Paper, Federal Reserve Board, Washington.

Engle, R.F., (1982). “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom inflation.” *Econometrica*, Vol. 50, 987-1008.

Engle, R.F., Bollerslev, T., (1986). “Modeling the persistence of conditional variances” *Econometric Reviews*, Vol. 5, 1–50.

Engle, R.F., Lilien, D., Robins, R., (1987). “Estimating time varying risk premia in the term structure: The ARCH-M model.” *Econometrica*, Vol. 55, 391–407.

Engle, R.F., Ng, V., (1993a). “Time-varying volatility and the dynamic behavior of the term structure.” *Journal of Money, Credit and Banking*, Vol. 25, 336-349.

Engle, R.F., Ng, V., (1993b). “Measuring and testing the impact of news on volatility.” *Journal of Finance*, Vol. 48, 1749-1778.

Engle, R.F., Ng, V., Rothschild, M., (1992). “Asset pricing with a factor ARCH covariance structure: Empirical estimates for treasurybills.” *Journal of Econometrics*, Vol. 45, 213–238.

Glosten, L.R., Jagannathan, R., Runkle, D.E., (1993). “On the relation between the expected value and the volatility of the nominal excess returns on stocks.” *Journal of Finance*, Vol. 48, 1779-1801.

Gray, S.F. (1996). “Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process.” *Journal of Financial Economics*, vol. 42, no. 1 (September):27–62.

Grossman, S.J., Melino, A., Schiller, R.J., (1987). “Estimating the Continuous-Time Consumption-Based Asset Pricing Model.” *Journal of Business and Economic Statistics*, Vol. 5, 315-327.

Hansen, L.P., (1982). “Large Sample Properties of Generalized Method of Moments Estimation.” *Econometrica*, Vol. 4, 1029-1054.

Hiraki, T., Takezawa, N., (1997). “How Sensitive is Short-Term Japanese interest rate volatility to the level of the interest rate?” *Economics Letters*, Vol. 56, 325-332.

Hong, C., (1988). “The integrated generalized autoregressive conditional heteroskedastic model. The process, estimation and Monte Carlo methods.” Working Paper, University of California, San Diego.

Koedijk, K.G., Nissen, F.G.J.A., Schotman, P.C., Wol., C.C.P., (1997). “The dynamics of short-term interest rate volatility reconsidered.” *European Finance Review*, Vol. 1, 105–130.

Litterman, R., and J. Scheinkman. (1991). “Common Factors Affecting Bond Returns.” *Journal of Fixed Income*, vol. 3, no. 3 (December):54–61.

Ljung, G.M., Box, G.P.E., (1978). “On a measure of lack of fit on time series models.” *Biometrika*, Vol. 66, 66-72.

Mc Manus, D., Watt, W., (1999). “Estimating One-Factor Models of Short-Term Interest Rates.” Bank of Canada Working Paper 99-18.

Merton. R. C. (1973). “The Theory of Rational Option Pricing.” *Bell Journal of Economics*, vol. 4, no. 1: 141- 183.

Nelson, D.B., (1991). “Conditional heteroskedasticity in asset returns: A new approach.” *Econometrica*, Vol. 59, 347-370.

Nowman. K.B. (1997). "Gaussian Estimation of Single-Factor Continuous Time Models of The Term Structure of Interest Rates." *Journal of Finance*, vol. 52, no. 4 (September): 1695 – 1706.

Nowman. K.B. (1998). "Continuous-Time Short Term Interest Rate Models." *Applied Financial Economics*, Vol. 8, 401 – 407.

Nowman. K.B. (2002). "The volatility of Japanese interest rates: Evidence for Certificate of Deposit and Gensaki rates." *International Review of Financial Analysis*, Vol. 11, 29–38.

Said, S.E., Dickey, D.A., (1984). "Testing for unit roots in autoregressive moving average models of unknown order." *Biometrika*, Vol. 71, 599-607.

Stanton, R. (1997). "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk." *Journal of Finance*, Vol. 52, no. 5 (December),1973–2002.

Tse, Y.K. (1995). "Some international evidence on the stochastic behavior of interest rates." *Journal of lnternational Money and Finance*, Vol. 14, No. 5, pp. 721-738.

Vasicek, O. (1977). "An Equilibrium Characterization of the Term Structure." *Journal of Financial Economics*, vol. 5, no. 2 (November):177–188.

Wooldridge, J.M., (1990). "A unified approach to robust, regression-based specification tests." *Econometric Theory*, Vol. 6, 17-43.

Appendix I: Derivation of Nowman (1997) Exact Discrete Model

Consider the CKLS stochastic differential equation:

$$dr(t) = \{\alpha + \beta r(t)\}\, dt + \sigma\, r^{\gamma}(t)\, dZ \tag{1}$$

Now assume as an approximation to the true underlying moment given by equation (1) that over the interval [0 , T], r(t) satisfies the stochastic differential equation:

$$dr(t) = \{\alpha + \beta r(t)\}\, dt + \sigma\, \{r(t'- 1)\}^{\gamma}\, dZ \tag{2}$$

where t' is the smallest integer greater than or equal to t. Nowman assumed that volatility changes at the beginning of a unit observation period and then remains constant trough the period until the next period.

Now, by splitting integers in equation (2), it follows that r(t) satisfies the stochastic integral equation:

$$r(t) - r(t'- 1) = \int_{t'-1}^{t} [\alpha + \beta r(s)] ds + \sigma\{r(t'-1)\}^{\gamma} \int_{t'-1}^{t} dZ(s) \tag{3}$$

for all t in [t'-1, t'] where

$$t'-1 < t \leq t' \quad \text{and} \quad \int_{t'-1}^{t} dZ(s) = Z[t`-1, t] \tag{4}$$

Under these assumptions, Nowman (1997) used the exact discrete model of Bergstrom (1984, Theorem 2) to obtain the discrete model corresponding to equation (3) given by

$$r(t) = e^{\beta} r(t\text{-}1) + \frac{\alpha}{\beta}(e^{\beta}-1) + \varepsilon_t \qquad (t = 1,2,...,T) \tag{5}$$

where ε_t (t = 1,2,...,T) satisfies the conditions

$$E(\varepsilon_t \varepsilon_s) = 0 \quad (t \neq s) \tag{6}$$

$$E(\varepsilon^2_t) = \int_{t-1}^{t} e^{2(t-\tau)\beta} \sigma^2 \{r(t-1)\}^{2\gamma} d\tau = \frac{\sigma^2}{2\beta}(e^{2\beta}-1)\{r(t-1)\}^{2\gamma} = m^2_{tt} \tag{7}$$

Comparison with the CKLS discrete approximation

Note that we can get the CKLS approximation as a special case of Nowman's exact discrete model derived in (5)-(7).

Using a Taylor series expansion for e^{β} yields: $e^{\beta} = 1 + \beta + \frac{\beta^2}{2!} + \ldots$, and using only the first two terms of the Taylor series expansion to replace e^{β} in (5) we get:

$$r(t) = (1 + \beta)\, r(t\text{-}1) + \frac{\alpha}{\beta}(1 + \beta - 1) + \varepsilon_t \tag{8}$$

$$r(t) = r(t\text{-}1) + \beta\, r(t\text{-}1) + \alpha + \varepsilon_t \tag{9}$$

$$r(t) - r(t\text{-}1) = \alpha + \beta\, r(t\text{-}1) + \varepsilon_t \tag{10}$$

Equation (10) is the CKLS discrete approximation. Following the same procedure for the conditional variance specification given in (7), we get:

$$m^2_{tt} = \frac{\sigma^2}{2\beta}\,(1 + 2\beta - 1)\,\{r(t\text{-}1)\}^{2\gamma} = \sigma^2\, r(t\text{-}1)^{2\gamma} \tag{11}$$

Equation (11) is the conditional variance specification of CKLS discrete approximation.