The Stochastic Volatility of Short-Term Interest Rates:
Some International Evidence

Abstract

This paper provides a practical method for efficiently estimating a stochastic volatility model of short-term riskless interest rate dynamics. Estimated interest rate dynamics are broadly similar across a number of countries and we find reliable evidence of stochastic volatility throughout. In contrast to stock returns, interest rate volatility exhibits faster mean reverting behavior and innovations in interest rate volatility are negligibly correlated with innovations in interest rates. The less persistent behavior of interest rate volatility reflects the fact that interest rate dynamics are impacted by transient economic shocks like central bank announcements (Japan and UK) and other macroeconomic news (US).
The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence

Clifford A. Ball
Owen Graduate School of Management
Vanderbilt University
Nashville, TN 37203

Walter N. Torous
Anderson School of Management
UCLA
Los Angeles, CA 90095

February 2, 1998
The Stochastic Volatility of Short-term Interest Rates:

Some International Evidence

The last decade has witnessed unprecedented growth in the market for fixed-income derivatives, domestically as well as internationally. In light of this growth, recent research has focused on empirically characterizing the dynamics of short-term riskless interest rates necessary for the valuation of many of these securities. With few exceptions (for example, Nowman (1997)), much of this evidence has been concerned with the behavior of US interest rates. The purpose of this paper is to investigate the dynamics of short-term interest rates across a number of countries.

A particularly noteworthy feature at least of the US evidence is that the volatility of short-term interest rates is itself volatile. The stochastic nature of volatility is an important property of interest rate dynamics which we wish to include in our comparison. To do so practically, this paper implements a computationally efficient and statistically reliable method for estimating stochastic volatility models in finance. We apply this method to Chan, Karolyi, Longstaff, and Sanders' (1992) (CKLS) proposed model of short-term riskless interest rate dynamics augmented with stochastic interest rate volatility.

Nowman fits the deterministic volatility version of the CKLS specification to one month UK interbank rates and concludes that, unlike one month US Treasury yields, UK interest rate volatility is not sensitive to the level of rates. This conclusion, however, may simply reflect that the behavior of interbank and Treasury rates differ regardless of their country of origin, or, alternatively, may stem from the fact that Nowman samples these interest rate series over different time periods. Additionally, Nowman's conclusions are based on
the assumption of deterministic interest rate volatility, a specification that we reject. For a cleaner comparison, we use proxies for short-term riskless rates drawn from the same market, the London interbank market, over the same sample period. Unlike Nowman, we find that the dynamics of one month Euro-dollar, Euro-mark, Euro-sterling, and Euro-yen rates are broadly similar and we provide evidence of stochastic interest rate volatility throughout. In contrast to estimated stock return dynamics, stochastic interest rate volatility is characterized by a faster mean reverting behavior. Furthermore, innovations in interest rates are only negligibly correlated with innovations in their volatility which also differs from the evidence on stock returns where a reliably negative relation between returns and volatility prevails. Finally, for all of the sampled series we confirm CKLS’s claim that interest rate volatility is sensitive to the level of interest rates, though this relation is difficult to estimate precisely.

The faster mean reverting behavior of stochastic interest rate volatility suggests that short-term interest rates are more likely to be affected by transient and less persistent economic shocks. ARCH models are predicated by the persistent nature of changing volatility and we find that if interest rate volatility is assumed to follow an EGARCH model with normal innovations then this specification is almost always rejected in favor of the stochastic volatility alternative. Intuitively, the stochastic volatility model better accommodates the transient yet significant economic shocks which impact interest rate dynamics. We find that these shocks are dominated by central bank announcements of base rate changes in the Japanese and UK data and by macroeconomic announcements in the US data. Only with extremely fat-tailed $t$-distributed innovations can the EGARCH specification accommodate these shocks and so become statistically indistinguishable from the stochastic volatility model.
The plan of this paper is as follows. Section I summarizes the short-term interest rate series and their sampling characteristics. Section II puts forward the stochastic volatility model of short-term riskless interest rate dynamics and investigates its statistical estimation. We use a filtering framework in which volatility is treated as an unobservable. Since volatility enters the interest rate model in a multiplicative fashion, we explicitly take into account the resultant non-normality of the measurement error which in quasi-maximum likelihood (QML) estimation is treated as if it were normal. The corresponding statistical improvement is achieved at the expense of a single dimensional numerical integration. We also show that the geometric scaling of interest rate data minimizes the collinearity between interest rate volatility and the level of rates, facilitating estimation of this relation. Section III presents our empirical results as well as sampling evidence which confirms the statistical superiority of the non-Gaussian estimation procedure to QML. Section IV compares our empirical results to those obtained for EGARCH specifications of interest rate volatility. Section V concludes the paper.

I. Data Description

The short-term riskless rate of interest is proxied by a one month interest rate. We use a variety of one month rates. Duffee (1996) argues that because of the increased idiosyncratic variation of US Treasury yields since the early 1980s, we should not rely exclusively on Treasury yields in calibrating models of short-term interest rate dynamics. Also, Nowman suggests that the nature of interest rate dynamics differs at least between the US and the UK.
One month US Treasury bill yields

One month US Treasury bill yields (the average of the bid and ask) are taken from the Center for Research in Security Prices' (CRSP) risk-free rate file. This monthly series is constructed by using that US Treasury bill whose maturity at month-end is closest to 30 days regardless of the bill’s original term to maturity. The sample period is June 1964 to December 1989, for a total of 307 observations, and corresponds to the sample period used by CKLS. ¹

One month UK interbank rates

Following Nowman, the short-term riskless rate of interest in the UK is proxied by the one month UK interbank (middle) rate. Weekly observations over the sample period January 1975 to September 1995, for a total of 1079 observations, are taken from Datastream. ² To investigate the estimation effects of sampling interest rate data at different frequencies, monthly observations are constructed by using corresponding end-of-month weekly observations. The resultant sample of 248 monthly observations from January 1975 to September 1995 corresponds roughly to Nowman’s sample period of March 1975 to March 1995.

One month London Euro-currency rates

We provide a further international comparison of short-term interest rate dynamics by using a number of one month rates drawn from the London Euro-currency market. Each of these rates is a London interbank rate denominated in a given currency. In particular, we use weekly observations of one month Euro-dollar, Euro-mark, Euro-sterling, and Euro-yen (middle) rates taken from Datastream ³ over the sample period January 1986 to December 1995, giving 522 observations for each series.
A. Summary Statistics

Table I gives summary statistics for the one month interest rates used to proxy the short-term riskless rate. These statistics are provided for both the level of interest rates, \( r_t \), and changes in interest rates, \( \Delta r_t \).

One month interest rate levels do not exhibit excessive skewness nor excessive kurtosis. With regard to the one month Euro-rates, the Euro-sterling rate was, on average, the highest and most volatile over the sample period. In contrast, the Euro-yen rate was, on average, the lowest. The Euro-mark, Euro-yen, and Euro-dollar rates have similar volatilities.

Changes in all of the one month interest rates are only slightly skewed but exhibit significant leptokurtosis. Weekly changes in the one month Euro-sterling rate are the most leptokurtic of the Euro-rates. Interestingly, monthly changes in the one month UK interbank rate are far more leptokurtic than corresponding weekly changes. These results are consistent with the presence of outliers in the UK data which become relatively more pronounced in the shorter monthly series. Prominent among these outliers is the doubling of the one month interbank rate to over 20% during the ERM crisis of September 1992.

II. A Stochastic Volatility Model of Short-Term Riskless Interest Rates

It is by now acknowledged that the volatility of short-term interest rate changes is itself volatile. The stochastic nature of interest rate volatility is suggested by Dybvig (1995), incorporated by Longstaff and Schwartz (1992) in their two-factor equilibrium term structure model, and modeled by Brenner, Harjes, and Kroner (1996), and Koedijk, Nissen, Schotman and Wolff (1997) in discrete-time GARCH extensions of CKLS.
We incorporate stochastic interest rate volatility into a discrete-time version of CKLS's proposed model of short-term riskless interest rate dynamics:

\[
    r_t = (a + br_{t-1}) + \sigma_{t-1} r_{t-1}^\gamma \epsilon_{1,t} \\
    \ln \sigma_t^2 - \mu = \beta(\ln \sigma_{t-1}^2 - \mu) + \xi \epsilon_{2,t}
\]

with \( \{\epsilon_{1,t}\} \) and \( \{\epsilon_{2,t}\} \) being iid standard normals. The parameters \( a \) and \( b \) characterize the interest rate's linear drift component while the parameter \( \gamma \) allows the volatility of \( r \) to functionally depend, in a power fashion, on its level. Unlike CKLS, the assumed unobservable interest rate volatility, \( \sigma_t \), evolves stochastically. In particular, the dynamics of \( \ln \sigma_t^2 \) are given by an AR(1) specification with the speed of adjustment to its unconditional mean \( \mu \) governed by the parameter \( \beta \) and its volatility characterized by the parameter \( \xi \).

We estimate the drift parameters \( a \) and \( b \) by OLS; these estimators are consistent even when volatility is stochastic.\(^4\) Letting \( x_t \equiv \ln \sigma_t^2 \) and \( \text{res}_t \equiv r_t - (\hat{a} + \hat{b} r_{t-1}) \), we take the logarithm of the squared observations and consider the following equivalent discrete-time specification:

\[
y_t \equiv \ln(\text{res}_t^2) = x_{t-1} + 2\gamma \ln(r_{t-1}) + \ln(\epsilon_{1,t}^2) \tag{1}
\]

\[
x_t - \mu = \beta(x_{t-1} - \mu) + \xi \epsilon_{2,t}. \tag{2}
\]

By squaring observations we lose all information about dependence between \( \epsilon_1 \) and \( \epsilon_2 \). This entails no loss of statistical efficiency if these error terms are uncorrelated, \( \rho = 0 \).

The Appendix confirms this assumption by providing evidence across our sampled interest rate series that \( \rho \approx 0 \). That is, upward shocks to interest rates are as likely as downward shocks to increase interest rate volatility. This is in contrast to stock returns where the empirical evidence is consistent with a reliably negative relation since downward shocks
to stock returns tend to be associated with increased stock volatility (see, for example, Nelson (1991)).

Expressions (1) and (2) give a linear time series model but with non-Gaussian errors since $\ln(\epsilon_1^2)$ is log $\chi^2$ distributed under the maintained assumption that $\epsilon_1$ is normally distributed. Treating $\ln(\epsilon_1^2)$ as if it were normally distributed, we can straightforwardly apply the Kalman filter to estimate the parameters of this model using quasi-maximum likelihood (QML) (Harvey, Ruiz, and Shephard (1994)). However, this approximate filter will be adversely affected by outlying observations in the log $\chi^2$ distribution's skewed left tail which are highly irregular under the normal approximation. To improve statistical performance while maintaining the tractability of QML, we now turn our attention to an alternative filtering procedure that explicitly incorporates this non-normality.

A. Practical Integration-based filtering of non-Gaussian time series

Given a set of $T$ observations $y_1, \ldots, y_T$, the likelihood of the parameter vector $\psi \equiv \{\mu, \beta, \xi, \gamma\}$ can be expressed as

$$\mathcal{L}(\psi) = \prod_{t=1}^{T} p(y_t \mid Y_{t-1})$$

where $Y_t \equiv \{y_t, y_{t-1}, \ldots, y_1\}$ denotes the history of the observable through time $t$ and $p(y_t \mid Y_{t-1})$ denotes the conditional density of $y_t$ given the history of the process through time $t - 1$.

This likelihood function may be evaluated using an iterative filtering procedure. Given a prior on the state $p(x_{t-1} \mid Y_{t-1})$, there are three stages to this procedure: a projection to obtain $p(x_t \mid Y_{t-1})$:

$$p(x_t \mid Y_{t-1}) = \int_{x_{t-1}} p(x_t, x_{t-1} \mid Y_{t-1}) dx_{t-1}$$
\[
= \int_{x_{t-1}} p(x_t \mid x_{t-1}, Y_{t-1})p(x_{t-1} \mid Y_{t-1})dx_{t-1}
\]
\[
= \int_{x_{t-1}} p(x_t \mid x_{t-1})p(x_{t-1} \mid Y_{t-1})dx_{t-1};
\]

followed by an integration to calculate the conditional likelihood \( p(y_t \mid Y_{t-1}) \):

\[
p(y_t \mid Y_{t-1}) = \int_{x_t} p(y_t, x_t \mid Y_{t-1})dx_t
\]
\[
= \int_{x_t} p(y_t \mid x_t)p(x_t \mid Y_{t-1})dx_t;
\]

and finally an updating to obtain \( p(x_t \mid Y_t) \):

\[
p(x_t \mid Y_t) = p(x_t \mid y_t, Y_{t-1})
\]
\[
= p(x_t, y_t \mid Y_{t-1})/p(y_t \mid Y_{t-1})
\]
\[
= p(y_t \mid x_t)p(x_t \mid Y_{t-1})/p(y_t \mid Y_{t-1}).
\]

Proceeding sequentially in this manner we may calculate \( L \). The maximum likelihood estimator \( \hat{\psi} \) is found by maximizing this function with respect to \( \psi \).

For a linear state space model with normal errors the Kalman filter may be applied and the first two moments of \( p(x_t \mid Y_t) \) are obtained from the first two moments of \( p(x_{t-1} \mid Y_{t-1}) \) by simple matrix calculations (see Harvey (1989)). This simple sequential scheme is, however, not applicable to non-Gaussian time series models like ours.

Rather than assuming normal measurement errors, as is the case with QML, we follow Frühwirth-Schnatter (1994) and approximate the prior density by a normal density with the same first and second moments as the prior. With this assumption \(^5\), the filter is then implemented in a similar fashion to the Gaussian case. The projection, based on a normal prior, preserves normality and so can be implemented analytically. The evaluation of the conditional likelihood requires numerical integration as the measurement error is \( \log \chi^2 \)
distributed, but the approximation will be highly accurate in our case as the integration is single dimensional. Computation of the first and second moments of the updated prior each requires an additional single dimensional numerical integration and so the iterative scheme may be continued.

In general, because the numerical integration is of the same dimension as the observation vector, this integration-based procedure provides a practical method of filtering univariate or multivariate financial time series of moderate dimension even for a very high-dimensional state vector. 6

A.1 Sampling Experiments - Generic Stochastic Volatility Model

A number of alternative procedures for estimating stochastic volatility models have recently been proposed. Examples include, among others, Jacquier, Polson, and Rossi (1994), and Sandmann and Koopman (1996). 7 To compare the properties of the non-Gaussian estimator we use the following generic model which has been used previously to benchmark statistical performance in the stochastic volatility literature:

\[ y_t = \sqrt{h_{t-1}} u_t \]

\[ \ln h_t = \alpha + \delta \ln h_{t-1} + \sigma_v v_t, \quad u_t, v_t \sim iid N(0,1). \]

We make our results directly comparable by calibrating this model precisely as in Jacquier, Polson, and Rossi, and Sandmann and Koopman. Parameter values \( \alpha, \delta, \) and \( \sigma_v \) are chosen to investigate the sensitivity of the estimation methods \( (i) \) to the size of the stochastic volatility component (the squared coefficient of variation \( \text{var}(h)/E(h)^2 = 10, 1, .1 \)), and \( (ii) \) to the persistence of the stochastic component (\( \delta = .9, .95, .98 \)), while fixing \( E(h) = .0009 \) or, approximately, a 20% annualized standard deviation for weekly data.
See Jacquier, Polson, and Rossi for a discussion of the relevance of these parameter values in the stochastic volatility literature. Time series of length $n=500$ are simulated and sampling distributions are based on 500 replications.

The results are tabulated in Table II. For each combination of parameter values the corresponding sampling results of Jacquier, Polson, and Rossi, and Sandmann and Koopman are reproduced in the first and second rows, respectively, while the third row gives the sampling results for the non-Gaussian estimator. The non-Gaussian estimator is as efficient as the Jacquier, Polson, and Rossi, and the Sandmann and Koopman estimation procedures across all the posited parameter values and exhibits similar biases. In particular, the sampling performance of each of the competing estimation procedures noticeably deteriorates as the size of the stochastic component (measured by $\text{var}(h)/E(h)^2$) decreases. Varying the persistence of the stochastic component (measured by $\delta$), in contrast, appears to have little effect on the sampling performance of the estimation procedures.

While these sampling results apply to the generic stochastic volatility model, they suggest that, in addition to its speed and adaptability, the non-Gaussian estimator provides an efficient means of estimating the stochastic volatility model of interest rate dynamics.

B. Geometric Scaling

Recent research has investigated whether the volatility of short-term interest rate changes is sensitive to the level of rates. For example, while CKLS find this to be the case for US data, Nowman finds no evidence of such a relation in UK data.

Empirically characterizing this relation is made difficult by the collinearity between the parameters of interest. For example, because of the collinearity between $\gamma$ and $\sigma$, Gray
(1996) fixes $\gamma = 0.5$ since the "interpretation of the individual parameter estimates (is) questionable at best." (Gray (1996), page 36). Rather than simply assuming a particular relation between the volatility and level of interest rates, we now show how the geometric scaling of interest rate data minimizes this collinearity.  

We motivate geometric scaling by couching our analysis in the deterministic volatility case

$$r_t = a + br_{t-1} + \sigma r_{t-1} \epsilon_t, \quad \epsilon_t \sim iid \ N(0, 1).$$

Since the parameters $a$ and $b$ do not affect the collinearity between $\gamma$ and $\sigma$, we assume they are known or have been previously consistently estimated. As a result, consider

$$res_t = \sigma r_{t-1} \epsilon_t,$$

or equivalently,

$$ln(res_t^2) = ln \sigma^2 + 2\gamma ln r_{t-1} + ln \epsilon_t^2.$$

The geometric mean of $\{r_1, \ldots, r_N\}$ is $\bar{r} = \left(\prod_{i=1}^{N} r_i\right)^{1/N}$. If we scale or divide the interest rate data by $\bar{r}$ then since

$$\frac{1}{N} \sum_t ln(res_t^2) = \frac{1}{N} \sum_t ln \sigma^2 + 2\gamma \frac{1}{N} \sum_t ln r_{t-1} + \frac{1}{N} \sum_t ln \epsilon_t^2,$$

it follows that $ln \sigma^2$ is consistently estimated by

$$\frac{1}{N} \sum_t ln(res_t^2) + 1.27$$

where we have used the fact that for geometrically scaled interest rate data $\sum_t ln r_{t-1} = 0$ and that the mean of $log \chi_{(1)}^2$ is -1.270 (Abramowitz and Stegun (1970), page 943). In other words, regardless of the value of $\gamma$, we obtain a consistent estimate of $ln \sigma^2$. Or, in
the case of stochastic volatility, we obtain a consistent estimate of \( \mu \). We use geometrically scaled interest rate data in our subsequent empirical analyses. \(^9\)

III. Estimation

The stochastic volatility model of short-term riskless interest rate dynamics, expressions (1) and (2), is estimated using one month US Treasury bill yields, one month UK interbank rates, as well as the one month London Euro-currency rates. Table III presents the results. For comparison purposes, we use both non-Gaussian (integration-based) and Gaussian (QML) estimation procedures. While Harvey, Ruiz, and Shephard do not use QML to estimate a model of interest rate dynamics, we do so because its tractability makes QML an obvious alternative to our proposed estimation procedure.

A. Empirical Results

Using either estimation procedure, we see clear evidence across all interest rate series of mean-reverting \((\hat{\beta} < 1)\) stochastic volatility \((\hat{\xi} > 0)\). Compared to stock returns (see, for example, Jacquier, Polson, and Rossi), stochastic interest rate volatility is less persistent. That is, stochastic interest rate volatility reverts more quickly to its long-term mean as evidenced by the relatively low \( \hat{\beta} \) estimates. While interest rate volatility would seem to be sensitive to the level of interest rates \((\hat{\gamma} > 0)\), this relation is difficult to accurately measure. Also the stochastic volatility model of short-term riskless interest rate dynamics is more precisely estimated using the non-Gaussian estimation procedure.

One month US Treasury bill yields

Once stochastic volatility is incorporated into the CKLS framework we still have evidence that the volatility of US Treasury yields is sensitive to the level of yields but no longer
to the extent estimated by CKLS; for example, using the non-Gaussian procedure we estimate $\hat{\gamma} = 0.651$. The estimated stochastic volatility process for US Treasury yields differs, at least for this sample period, from that estimated for the other interest rate series being more persistent and so reverting much more slowly to its long-term mean.

One month UK interbank rates

We confirm in the presence of stochastic volatility Nowman’s claim when using monthly observations of the one month UK interbank rate that UK interest rate volatility does not appear to be sensitive to the level of interest rates. Our estimate of $\hat{\gamma} = 0.261$, based on the non-Gaussian procedure, is similar to Nowman’s estimate of $\hat{\gamma} = 0.290$ and cannot be statistically distinguished from zero because of the estimate’s relatively large standard error. However, when we sample these rates on a weekly basis, thereby increasing the number of observations over the sample period, we reliably estimate $\hat{\gamma} = 0.486$.

This comparison suggests that it is difficult to precisely estimate $\gamma$ and only with a large amount of data, over 1000 weekly observations in our case, can accurate estimates be obtained. Accordingly, Nowman’s results should not be interpreted as evidence that UK interest rate volatility is not sensitive to the level of interest rates but rather that monthly observations over a twenty year sample period are insufficient to accurately characterize this relation.

One Month London Euro-currencies

The estimation results based on one month Euro-dollar, Euro-mark, Euro-sterling, and Euro-yen rates are broadly consistent with one another. For each of these interest rate series we have reliable evidence of mean-reverting stochastic interest rate volatility but
we cannot statistically distinguish between the corresponding estimates of the stochastic volatility's speed of adjustment, $\beta$, or between the corresponding estimates of its volatility, $\xi$. As expected, there are also no statistically significant differences in the estimated relation between short-term interest rate volatility and the level of interest rates across these series owing to the relative imprecision of the $\gamma$ estimates. For example, we cannot statistically reject that $\gamma = 0.5$ (Cox, Ingersoll, and Ross (1985)) characterizes the dynamics of each of the sampled Euro-currency rates. 12

B. Sampling Experiments - Stochastic Volatility Model of Interest Rate Dynamics

While our estimation results suggest that the stochastic volatility model of short-term riskless interest rate dynamics is more accurately estimated by the non-Gaussian estimator, we now provide Monte Carlo simulation evidence to confirm this.

The sampling experiment assumes parameter values broadly consistent with those characterizing the US Treasury bill yield results based on monthly observations ($a = 0.06$, $b = 0.92$, $\mu = -6.0$, $\beta = 0.93$, $\xi = 0.32$, and $\gamma = 0.50$). Interest rate and volatility series each of length $n = 500$ observations are simulated. After geometrically scaling this interest rate data, the parameters $a$ and $b$ are estimated using $OLS$. The stochastic volatility interest rate model, expressions (1) and (2), is then estimated using the non-Gaussian procedure and, alternatively, using QML. The experiment is repeated 500 times and the sampling distributions of the competing estimators of $\mu$, $\beta$, $\xi$, and $\gamma$ are tabulated in Table IV.

Panel A of Table IV presents the non-Gaussian results while Panel B presents the QML results. The performance of the non-Gaussian estimator is superior. In comparison, the QML estimator exhibits more bias and is far more variable.
Panels C and D of Table IV explore the sensitivity of the non-Gaussian estimator's sampling performance to varying $\gamma$ values. To do so we increase (decrease) $\gamma$ to 0.75 (0.25) but numerically determine that value of $\mu$ which over 500 replications of the simulated interest rate series gives the same sample mean and same sample standard deviation (to two significant digits) of $r$ as obtained in the sampling experiment assuming $\gamma = 0.5$. This ensures that any differences in sampling performance are not due to differences in the variability of the simulated interest rate data. From Panels C and D we see that the performance of the non-Gaussian estimator is insensitive to the assumed $\gamma$ value. In particular, $\gamma$ is estimated unbiasedly throughout although the corresponding variability varies slightly with the level of $\gamma$, being smaller for $\gamma = 0.25$ ($sd(\hat{\gamma}) = 0.077$) than for $\gamma = 0.75$ ($sd(\hat{\gamma}) = 0.105$).

To further investigate the robustness of the non-Gaussian estimator, Panel E presents the sampling results when we assume parameter values broadly consistent with those characterizing the Euro-dollar rate results based on weekly observations ($a=.005, b=.99, \mu= -8.0, \beta = 0.5, \xi = 1.2,$ and $\gamma = 0.75$). All of the parameters again appear to be estimated unbiasedly and precisely, although the variability of the $\gamma$ estimator is now larger than reported in Panel A (when $\gamma = 0.50$).

IV. A Comparison with EGARCH

ARCH models have been extensively used to characterize volatility in a variety of financial markets. For example, Longstaff and Schwartz, among others, have modeled short-term interest rate volatility using a GARCH(1,1) specification. A comparison against the stochastic volatility model provides further insights into interest rate dynamics.

For illustrative purposes, the discrete-time version of CKLS's model of short-term riskless
interest rate dynamics is now augmented by an EGARCH(1,1) (Nelson (1991)) specification for volatility:

\[ \text{res}_t = \exp \left( \frac{1}{2} \sum_{j=1}^{p} \gamma_j \epsilon_{t-j} \right) \epsilon_t \]  

\[ x_t - \mu = \beta (x_{t-1} - \mu) + g_0 \epsilon_{t-1} + g_1 (| \epsilon_{t-1} | - E [ | \epsilon_{t-1} | ] ) . \]  

As before, \( \text{res}_t \equiv r_t - (\hat{a} + \hat{b} r_{t-1}) \). We assume \( \epsilon_t \sim iid \ N(0,1) \), or alternatively, \( \epsilon_t \sim iid \ t_\nu \). Unlike the stochastic volatility model, there is now only one source of uncertainty, \( \epsilon_t \), and the stochastic nature of volatility is introduced by modeling volatility as a function of past realizations of this error.

A. Empirical Results

Table V presents the results of estimating the EGARCH model of short-term riskless interest rate dynamics, expressions (3) and (4).

Assuming either normally distributed or \( t \)-distributed innovations, ARCH effects in (log) interest rate volatility are evident. The estimated degrees of freedom, \( \hat{\nu} \), associated with the \( t \)-distributed innovations is approximately two for all but one of the sampled interest rate series (US Treasury yields), as compared to approximately ten for stock returns (Engle and Bollerslev (1986)), and is consistent with extreme outlying innovations influencing the behavior of interest rate volatility. Notice that \( \gamma \) is not estimated as precisely when \( t \)-distributed innovations are assumed. Intuitively, it is more difficult to separate out the sensitivity of interest rate volatility to the level of rates in the presence of these fat-tailed innovations; for example, high interest rate volatility may be attributable to extreme realized innovations or, alternatively, to a high level of prevailing interest rates.

A standard likelihood ratio test cannot be used to compare the stochastic volatility model (expressions (1) and (2)) to the EGARCH models (expressions (3) and (4)) since the
models are not nested. To make comparisons we follow Naik and Lee (1994) and use Vuong's (1989) modified likelihood ratio test for comparing non-nested hypotheses. Under the null hypothesis that the competing models are statistically indistinguishable, the Vuong statistic is standard normally distributed.

From Table V we see that the Vuong test rejects the EGARCH model with normally distributed innovations in favor of the stochastic volatility model for all but one of the interest rate series (US Treasury yields). With $t$-distributed innovations, we cannot statistically distinguish the stochastic volatility model from the EGARCH model.

To better understand these results, recall that ARCH models are predicated by the persistent nature of changing volatility. That is, periods of high and low volatility tend to cluster together (Engle (1982)). If short-term interest rates are impacted by more transient economic shocks then ARCH models may provide a less accurate description of their volatility dynamics since past errors will not accurately forecast volatility. Fat-tailed $t$-distributed innovations remedy this deficiency by better accommodating these transient economic shocks.

These results are illustrated in Figure 1 where we use the Euro-mark interest rate series and plot for each sample point $t$ the difference between the maximized log-likelihood values, based on data through $t-1$, of the stochastic volatility and EGARCH models evaluated at their respective sample-wide maximum likelihood estimates. Intuitively, a positive difference at $t$ is consistent with the stochastic volatility model being better able to explain the data at $t$.

Notice in Panel A of Figure 1 that for most of the sample points the stochastic volatility model's improvement over the EGARCH model with normal innovations is minimal.
However, there are a small number of very brief episodes where substantial improvement is evident. We also identify in Panel A of Figure 1 specific events, if any, associated with each of the five largest discrepancies between the likelihoods of the stochastic volatility model and the EGARCH model assuming normal innovations. In particular, we searched the Financial Times in the vicinity of these dates for mention of any noteworthy events that may have affected German credit markets. While a Bundesbank announcement (December 1990) resulted in the largest discrepancy, other economic shocks are important. These include the ERM crisis (September 1992), announcement of an OPEC production agreement (November 1988) and disclosure of dialogue with East Germany on monetary union (February 1990). In the remaining case (December 1986), no specific economic event could be found.

Apart from the ERM crisis (September 1992), the corresponding discrepancies between the stochastic volatility model and the EGARCH model with normal innovations for the Euro-pound interest rate series are dominated by announcements of base rate changes by the Bank of England, both increases (May 1989, August 1987, November 1988) as well as decreases (October 1990). Central bank actions also dominate the Euro-yen interest rate results: discount rate increases (October 1989) and decreases (January 1990, April 1995), open market operations (October 1988) and pledges by the Bank of Japan to lower rates (April 1995).

By contrast, the largest discrepancies for the euro-dollar interest rate series are due to surprises in the release of macroeconomic activity data. Examples include unexpectedly strong economic growth (November 1988), revisions of past growth rates (December 1992), and increases in consumer confidence (November 1993). These results complement those of Edison (1996) and Fleming and Remolona (1997), among others, who demonstrate
at daily and intra-day frequencies, respectively, that both short-term and long-term US Treasury rates respond systematically to surprises in the release of US macroeconomic activity data.

Alternatively, we can interpret these events as giving rise to the outliers which are captured by assuming $t$-distributed innovations in the EGARCH model. The resultant improvement in the fit of the EGARCH model for the Euro-mark interest rate series is evident in Panel B of Figure 1 where we see that the difference between the maximized log-likelihood values of the stochastic volatility model and the EGARCH model with $t$-distributed innovations is negligible across all of the sample points.

The empirical adequacy of the stochastic volatility model stems from the fact that it is also characterized by fat-tailed changes in interest rates since by construction it simultaneously incorporates interest rate and variance innovations to yield a variance mixture. Taking this mixture specification to its continuous-time limit, the resultant stochastic volatility contingent claims model may be used to consistently value fixed-income derivative securities (see, for example, Longstaff and Schwartz). Of course our estimation methodology is not limited to short-term interest rate dynamics and so the more accurate implementation of stochastic volatility contingent claims models in general can be expected.

V. Conclusions

The volatility of short-term interest rates is stochastic. This conclusion holds for all of the one month interest rate series examined. Models of short-term riskless interest rate dynamics which assume deterministic volatility are misspecified. Systematic errors in the valuation and hedging of fixed-income derivative securities will result.
Stochastic interest rate volatility differs from stochastic stock return volatility. Innovations in short-term interest rate levels are only negligibly correlated with innovations in the volatility of short-term interest rates. Interest rate volatility is also characterized by a faster mean reverting behavior reflecting the impact of transient economic shocks like central bank announcements of base rate changes. Our empirical results suggest that future work along the lines of Babbs and Webber (1994) and Balduzzi, Bertola, and Foresi (1993) to explicitly incorporate the implications of central bank actions on the behavior of short-term interest rates holds the promise of enhancing our understanding of short-term interest rate dynamics and ultimately of improving the pricing of fixed-income derivative securities.
Appendix

Consider the following discrete-time model of short-term riskless interest rate dynamics:

\[
\begin{align*}
    r_t &= a + b x_{t-1} + \exp \left( \frac{1}{2} \sigma^2_{t-1} \right) \xi_{1,t} \\
    x_t &= \alpha + \beta x_{t-1} + \xi_{2,t}
\end{align*}
\]

where as before \( x_t \) denotes the unobservable log interest rate volatility, \( x_t \equiv \ln \sigma^2_t \), but now \( \{ \xi_{1,t}, \xi_{2,t} \} \) are bivariate normally distributed with correlation \( \rho \). Clearly \( \{ x_t, r_t \} \) is a bivariate Markov process with a bivariate normal transition density.

Recall that if \( \{ X, Y \} \) is bivariate normal with mean vector \( [\mu_X, \mu_Y] \), variance vector \( [\sigma^2_X, \sigma^2_Y] \), and correlation \( \rho \), then the transition density at \( X = x, Y = y \) is

\[
f_{X,Y}(x,y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho^2}} \exp \left( -\frac{1}{2} Q \right)
\]

where

\[
Q = \frac{1}{1 - \rho^2} \left( \frac{X - \mu_X}{\sigma_X} \right)^2 - \frac{2 \rho (X - \mu_X)(Y - \mu_Y)}{\sigma_X \sigma_Y} + \frac{(Y - \mu_Y)^2}{\sigma_Y^2}.
\]

In our case, \( \mu_X \equiv \alpha + \beta x_{t-1}, \mu_Y \equiv \alpha + b r_{t-1}, \sigma^2_X \equiv \xi^2, \sigma^2_Y \equiv \exp(x_{t-1}) \sigma^2_{t-1}, \) and \( \rho \) denotes the correlation between innovations in interest rate levels and interest rate volatility. We therefore have an explicit formulation of the transition density \( p(x_t, r_t | x_{t-1}, r_{t-1}) \) where \( \{ r_t, r_{t-1} \} \) are observable while \( \{ x_t, x_{t-1} \} \) are unobservable but fall in a range of plausible values. Recall that since \( x_t \sim AR(1) \), its limiting distribution is normal with mean \( \alpha / (1 - \beta) \) and variance \( \xi^2 / (1 - \beta^2) \).

The joint system \( \{ x_t, r_t \} \) is a Markov process; however, \( \{ r_t \} \) is not and so its entire history must be used to evaluate the likelihood function. Letting \( R_t \) denote the history of the observable through time \( t \), \( R_t \equiv \{ r_1, r_2, \ldots, r_t \} \), the likelihood can be written as

\[
\mathcal{L} \equiv p(R_T) = \prod_{t=2}^{T} p(r_t | R_{t-1})
\]

and evaluated recursively as follows:

1. Given \( p(x_{t-1}, r_{t-1} | R_{t-1}) \), the projection is given by

\[
p(x_t, r_t | R_{t-1}) = \int_{x_{t-1}} p(x_t, r_t | x_{t-1}, r_{t-1}) p(x_{t-1}, r_{t-1} | R_{t-1}) dx_{t-1}
\]

and requires a single numerical integration with respect to \( x_{t-1} \).

2. The conditional likelihood is then

\[
p(r_t | R_{t-1}) = \int_{x_t} p(x_t, r_t | R_{t-1}) dx_t
\]

and also requires a single numerical integration but now with respect to \( x_t \) for each grid point of \( x_{t-1} \).

3. Given these results, the updating can be performed

\[
p(x_t, r_t | R_t) = p(x_t, r_t | R_{t-1}) / p(r_t | R_{t-1})
\]
The likelihood $\mathcal{L}$ is evaluated by repeating this recursion for each time point $t$ in the sample.

The gradient of the likelihood with respect to any parameter $\theta$ can also be calculated recursively. To see this, note that

$$
\ln \mathcal{L} \equiv \ln p(R_T) = \sum_t \ln p(r_t|R_{t-1})
$$

where

$$
p(r_t|R_{t-1}) = \int_{x_t} \int_{x_{t-1}} p(x_t, r_t|x_{t-1}, r_{t-1}) p(x_{t-1}, r_{t-1}|R_{t-1}) dx_{t-1} dx_t
$$

and

$$
\frac{\partial \ln p(R_T)}{\partial \theta} = \sum_t \frac{\partial p(r_t|R_{t-1})}{\partial \theta} / p(r_t|R_{t-1})
$$

where

$$
\frac{\partial p(r_t|R_{t-1})}{\partial \theta} = \int_{x_t} \int_{x_{t-1}} \left\{ \frac{\partial p(x_t, r_t|x_{t-1}, r_{t-1})}{\partial \theta} p(x_{t-1}, r_{t-1}|R_{t-1}) + p(x_{t-1}, r_{t-1}|R_{t-1}) \frac{\partial p(x_t, r_t|x_{t-1}, r_{t-1})}{\partial \theta} \right\} dx_{t-1} dx_t.
$$

The first part of the integrand may be viewed as filtering the partial derivative of the transition density with respect to $\theta$ and may be computed in the same recursive manner as the original filter. The second part of the integrand may be viewed as filtering the partial derivative of the prior with respect to $\theta$. Hence to compute the gradient of the likelihood requires running the original filter plus an additional filter for each parameter to be estimated.

All numerical integrations are performed with a 141-point extended Simpson's rule. The range for the unobservable log interest rate was chosen based on a series of experiments. The filter is initiated with a normal prior on the unobservable with mean equal to the estimate of the parameter $\mu$, and unit variance. The first twelve observations are used to warm-up the filter.

The estimation results with corresponding asymptotic standard errors in parentheses, obtained from the inverse of the score, are as follows. As before, all interest rate series are geometrically scaled.

<table>
<thead>
<tr>
<th>one month interest rate</th>
<th>$\mu$</th>
<th>$\beta$</th>
<th>$\xi$</th>
<th>$\gamma$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK interbank (weekly)</td>
<td>-7.865</td>
<td>0.711</td>
<td>1.129</td>
<td>0.435</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>(.131)</td>
<td>(.039)</td>
<td>(.073)</td>
<td>(.214)</td>
<td>(.062)</td>
</tr>
<tr>
<td>Euro-mark (weekly)</td>
<td>-8.161</td>
<td>0.785</td>
<td>0.987</td>
<td>0.250</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>(.203)</td>
<td>(.041)</td>
<td>(.091)</td>
<td>(.320)</td>
<td>(.082)</td>
</tr>
<tr>
<td>Euro-yen (weekly)</td>
<td>-7.241</td>
<td>0.694</td>
<td>1.106</td>
<td>0.259</td>
<td>0.224</td>
</tr>
<tr>
<td></td>
<td>(.179)</td>
<td>(.067)</td>
<td>(.131)</td>
<td>(.132)</td>
<td>(.091)</td>
</tr>
<tr>
<td>Euro-sterling (weekly)</td>
<td>-8.411</td>
<td>0.763</td>
<td>0.932</td>
<td>0.406</td>
<td>-0.029</td>
</tr>
<tr>
<td></td>
<td>(.194)</td>
<td>(.049)</td>
<td>(.094)</td>
<td>(.308)</td>
<td>(.103)</td>
</tr>
<tr>
<td>Euro-dollar (weekly)</td>
<td>-7.984</td>
<td>0.598</td>
<td>1.222</td>
<td>0.672</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>(.151)</td>
<td>(.074)</td>
<td>(.119)</td>
<td>(.227)</td>
<td>(.086)</td>
</tr>
</tbody>
</table>

Convergence was not achieved for the monthly US Treasury yield series or the monthly UK interbank series, perhaps because of the relatively small number of observations in each case.
Notice that the correlation coefficient $\rho$ is small in magnitude throughout. Only for the Euro-mark and Euro-yen rates is the correlation coefficient significantly positive; $\rho$ is indistinguishable from zero in all the other cases. Also, the estimates of the remaining parameters are remarkably similar to the corresponding estimates in Table III obtained using the computationally more efficient non-Gaussian procedure which requires at least one hundred times less computational effort than the non-linear filtering approach.
References


Babbs, Simon and Nick Webber, 1994, A theory of the term structure with an official short rate, unpublished manuscript, University of Warwick, Coventry, UK.


Naik, V. and M. Lee, 1994, The yield curve and bond option prices with discrete shifts in economic regimes, unpublished manuscript, University of British Columbia, Vancouver, B.C.


Footnotes

1. CKLS proxy the short-term riskless rate of interest by the one month Treasury yield (the average of the bid and ask) reported in CRSP's Fama 12-month Treasury bill term structure file. This file is based on the longest bill with at least 11 months and 10 days to maturity available on a given date. The yield on this bill when there is approximately one month until its maturity is taken to be the one month Treasury yield. Unfortunately, as a result, there is significant variation between the target (one month) and actual maturities. For example, over this sample period, the days-to-maturity underlying the one month yield series obtained from the Fama 12-month Treasury bill file range from 10 to 41 days. Given that with one month until its maturity this twelve-month bill is extremely off-the-run, liquidity effects may further jeopardize the accuracy of the yields used by CKLS. See Duffee for further details.

2. Datastream file LDNB1M. We use Wednesday to Wednesday observations. These data are unavailable prior to 1975.

3. Datastream files ECUS81M, ECWGM1M, ECUKL1M, and ECJAP1M, respectively. We use Wednesday to Wednesday observations.

4. This autoregression is subject to near unit root effects. We provide Monte Carlo evidence later in the paper (footnote #10) that for typical sample sizes encountered in practice the resultant biases in the OLS estimates of \( a \) and \( b \) do not significantly affect the estimation of the remaining volatility parameters.

5. The filter approximates the prior \( p(x_t|Y_{t-1}) \) which results from the convolution of the normal transition density \( p(x_t|x_{t-1}) \) with the posterior \( p(x_t|Y_{t-1}) \) by a normal density. If the posterior \( p(x_t|x_{t-1}) \) is close to a normal density this error will be small. The experimental results of Fahrmeir (1992) indicate that with increasing \( t \) the posterior tends to be normal even in cases where the observation density is extremely non-normal. We initially "warm-up" the filter for twelve observations to improve this accuracy.

6. For example, following Andersen and Lund (1996) and Balduzzi, Das, Foresi, and Sundaram (1997), our estimation procedure can accommodate a stochastic mean rate, \( \theta_t \), in addition to stochastic volatility:

\[
\begin{align*}
    r_t &= (1-b)\theta_{t-1} + b r_{t-1} + exp(\frac{1}{2} \beta_\epsilon x_{t-1}^2) \epsilon_{1,t} \\
    x_t - \mu_x &= \beta_x (x_{t-1} - \mu_x) + \zeta_x \epsilon_{2,t} \\
    \theta_t - \mu_\theta &= \beta_\theta (\theta_{t-1} - \mu_\theta) + \zeta_\theta \epsilon_{3,t}
\end{align*}
\]

with \( \epsilon_{1,t} \sim iid N(0,1) \) and \( \epsilon_{2,t}, \epsilon_{3,t} \sim BVN(0, \Sigma) \). Notice if \( b \sim 1 \), or equivalently \( 1-b \sim 0 \), so that short-term interest rate dynamics are characterized by a near unit root (Ball and Torous (1996)), then interest rates will not be informative about mean rate dynamics. The resultant borderline stationarity of interest rates makes reliable inference about their mean difficult.

7. Jacquier, Polson, and Rossi develop a Bayesian approach to estimate stochastic volatility models by using a Markov Chain Monte Carlo technique to draw directly from the posterior distribution of the model parameters and so construct finite-sample distributions for any function of these parameters. Sandmann and Koopman use Monte Carlo simulation to evaluate the residual part of the stochastic volatility model's likelihood function not captured by the Gaussian likelihood of QML.

8. We thank an anonymous referee for encouraging us to pursue this line of approach.
9. Alternatively, to investigate the collinearity between maximum likelihood estimators \( \hat{\gamma} \) and \( \hat{\sigma}^2 \) in the deterministic volatility case consider

\[
\frac{\partial^2 \ln L}{\partial \sigma^2 \partial \gamma} = -\frac{2}{\sigma^2} \sum_t \left( \frac{\text{res}_t}{r_t^2} \right)^2 \ln r_{t-1}
\]

which is the off-diagonal term of \( \ln L \)'s Hessian matrix. If this term is made close to 0 then the inverse Hessian will also be diagonal and the correlation between \( \gamma \) and \( \sigma^2 \) will be close to 0. But notice that this term is a weighted average across \( t \) of the log interest rate data. If the interest rate data are geometrically scaled, \( \sum_t \ln r_{t-1} = 0 \), then \( \sum_t \omega_t \ln r_{t-1} \approx 0 \) for \( \omega_t = (\text{res}_t/r_t^2) \), since \( \omega_t \) is an estimate of a constant variance OLS regression.

10. As a check of the appropriateness of the standard errors obtained from the non-Gaussian estimation procedure we conducted the following sampling experiment assuming the US Treasury bill yields' parameter estimates. Interest rate and volatility series each of length \( n = 307 \) observations are generated and the parameters estimated using the non-Gaussian procedure. By repeating this 500 times the sampling distributions of the parameter estimates as well as their asymptotic standard errors (from the Hessian matrix) are obtained and summarized as follows:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu = -4.50 )</td>
<td>-4.574</td>
<td>0.300</td>
<td>0.249</td>
<td></td>
</tr>
<tr>
<td>( \beta = 0.93 )</td>
<td>0.889</td>
<td>0.086</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td>( \xi = 0.32 )</td>
<td>0.345</td>
<td>0.107</td>
<td>0.103</td>
<td></td>
</tr>
<tr>
<td>( \gamma = 0.65 )</td>
<td>0.643</td>
<td>0.252</td>
<td>0.228</td>
<td></td>
</tr>
</tbody>
</table>

As can be seen the point estimates themselves are approximately unbiased with only the \( \beta \) estimate being slightly downward biased. The standard errors reported in Table II correspond well to their simulated counterparts except for the reported \( se(\hat{\beta}) = 0.043 \) which appears overly precise.

To assess whether the near-unit root in interest rate dynamics and the resultant small-sample bias in the first stage estimation of \( b \) will systematically bias the second stage results we repeated this sampling experiment but rather than estimating \( a \) and \( b \) we assumed the actual \( a \) and \( b \) values:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Mean</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu = -4.50 )</td>
<td>-4.574</td>
<td>0.300</td>
<td>0.253</td>
<td></td>
</tr>
<tr>
<td>( \beta = 0.93 )</td>
<td>0.891</td>
<td>0.083</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td>( \xi = 0.32 )</td>
<td>0.349</td>
<td>0.106</td>
<td>0.103</td>
<td></td>
</tr>
<tr>
<td>( \gamma = 0.65 )</td>
<td>0.644</td>
<td>0.245</td>
<td>0.228</td>
<td></td>
</tr>
</tbody>
</table>

While the Monte Carlo determined standard errors are now slightly smaller, there are no significant differences in the sampling results.

11. This result may reflect the Federal Reserve Board's decision to target money supply growth as opposed to interest rates between October 1979 and September 1982 and the resultant increase in interest rate volatility over that time period.

12. We further assesses whether our empirical results are sensitive to the assumed drift specification. First, Ait-Sahalia (1996) shows, at least in the deterministic volatility case, that a non-linear drift of the form \( a + b r_t + c r_t^2 + d/r_t \) is empirically more appropriate. Using geometrically scaled UK interbank rate data, we explore whether Ait-Sahalia's conclusion still holds in the presence of stochastic volatility. To do so, we simultaneously estimate in one step the parameters of the stochastic volatility model assuming a linear drift \( a + b r_t \), as well as assuming a non-linear drift \( a + b r_t + c r_t^2 + d/r_t \). The corresponding maximized log-likelihood function values \( \ln L \) and volatility parameter estimates (with asymptotic standard errors in parentheses) are as follows:
A likelihood ratio test rejects the linear specification at a 5% significance level. However, the volatility parameter estimates are insignificantly different from one another. In fact, we also see very little difference between these estimates and the corresponding estimates in Table III based on our computationally less burdensome two step procedure. Second, as noted earlier (footnote #6), our estimation procedure can also accommodate a stochastic mean rate, $\theta_t$. We use geometrically scaled US Treasury bill yields whose dynamics over our sample period are sufficiently away from a unit root to allow meaningful inference and obtain the following results (with asymptotic standard errors in parentheses):

<table>
<thead>
<tr>
<th></th>
<th>$\ln \mathcal{L}$</th>
<th>$\mu$</th>
<th>$\beta$</th>
<th>$\xi$</th>
<th>$\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>linear</td>
<td>2393.9917</td>
<td>-8.98</td>
<td>0.790</td>
<td>1.257</td>
<td>0.583</td>
</tr>
<tr>
<td></td>
<td>(.129)</td>
<td>(.050)</td>
<td>(.102)</td>
<td>(.209)</td>
<td></td>
</tr>
<tr>
<td>non-linear</td>
<td>2397.2053</td>
<td>-7.948</td>
<td>0.471</td>
<td>1.230</td>
<td>0.568</td>
</tr>
<tr>
<td></td>
<td>(.137)</td>
<td>(.046)</td>
<td>(.099)</td>
<td>(.212)</td>
<td></td>
</tr>
</tbody>
</table>

While these results are consistent with the presence of a stochastic mean rate in US Treasury bill yields, the estimates of the parameters characterizing stochastic volatility ($\gamma, \beta_x, \mu_x, \xi_x$) are once again similar to the corresponding estimates in Table III.

13. Vuong’s test statistic is a likelihood ratio statistic adjusted by the standard deviation of the difference in maximized log-likelihood functions under the competing models. In particular, let $lnf_t$ denote the maximized log-likelihood of the observable at $t$ based on its history through $t-1$ under $H_1$ (say the stochastic volatility model) and let $lnG_t$ be the corresponding maximum log-likelihood under $H_2$ (say the EGARCH model). For convenience, set $q_t = lnf_t - lnG_t$. With $T$ observations, define $LR_T = \sum_{t=1}^{T} q_t$, the maximum log-likelihood ratio and denote by $w_T$ the estimated standard deviation of $q$ based on the T observations. Under mild regularity conditions, outlined in Vuong, the test statistic $(1/\sqrt{T}) \ LR_T/w_T$ is approximately standard normally distributed when the competing hypotheses are indistinguishable. All else equal, large variability in the difference between the competing model’s log-likelihood functions diminishes our ability to reliably reject one model in favor of the other.

14. Our conclusions are not altered if another EGARCH specification with normal innovations, e.g., EGARCH(2,2), is assumed.

15. The weekly UK interbank rate results for the EGARCH model with $t$-distributed innovations are obtained after dropping the 15 September 1992 outlier and replacing it by the previous day’s observation. Without this change, the EGARCH model with $t$-distributed innovations converges to $\gamma > 1$ resulting in a non-stationary estimated interest rate process and so invalidating the Vuong test.

16. Since we rely on numerical integration, our filtering procedure can also accommodate $t$-distributed innovations. However, to underscore the adequacy of the mixture of normal innovations, in unreported results we found no evidence of statistical improvement in the stochastic volatility model’s performance for any of the interest rate series when we assumed $t$-distributed rather than normal innovations.
Table I

Summary Statistics for One Month Interest Rates Used to Proxy the Short-Term Riskless Interest Rate

The short-term riskless rate of interest, $r_t$, is proxied by a number of one month interest rates. The one month US Treasury bill yield is obtained from the CRSP risk-free rate file and the sample period coincides with that of Chan, Karolyi, Longstaff and Sanders (1992). The one month UK interbank rate, middle rate, for the sample period 1975 to 1995 is obtained from Datastream and corresponds approximately to the sample period of Newman (1996). London one month Euro-currency rates for the sample period 1985 to 1995 are also obtained from Datastream. The number of observations for each series is denoted by nobs. Summary statistics, including the sample mean, denoted by mean, sample standard deviation, denoted by std, sample skewness, denoted by skew, and sample kurtosis, denoted by kurtosis, are presented for both interest rates $r_t$ and their first differences $\Delta r_t$.  

<table>
<thead>
<tr>
<th>one month interest rate</th>
<th>sample period</th>
<th>frequency</th>
<th>Data Source</th>
<th>nobs</th>
<th>mean</th>
<th>std</th>
<th>skew</th>
<th>kurtosis</th>
<th>mean</th>
<th>std</th>
<th>skew</th>
<th>kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Treasury bill</td>
<td>1964 - 1989</td>
<td>monthly</td>
<td>CRSP</td>
<td>307</td>
<td>0.066</td>
<td>0.026</td>
<td>1.228</td>
<td>1.420</td>
<td>9.810×10^{-5}</td>
<td>0.009</td>
<td>-0.909</td>
<td>7.536</td>
</tr>
<tr>
<td>UK interbank</td>
<td>1975 - 1995</td>
<td>monthly</td>
<td>Datastream</td>
<td>248</td>
<td>0.108</td>
<td>0.032</td>
<td>0.025</td>
<td>-0.578</td>
<td>-3.769×10^{-5}</td>
<td>0.006</td>
<td>-0.780</td>
<td>207.669</td>
</tr>
<tr>
<td></td>
<td></td>
<td>weekly</td>
<td>Datastream</td>
<td>1079</td>
<td>0.109</td>
<td>0.032</td>
<td>0.020</td>
<td>-0.627</td>
<td>-2.049×10^{-4}</td>
<td>0.008</td>
<td>1.368</td>
<td>4.550</td>
</tr>
<tr>
<td>London Euro-currencies:</td>
<td></td>
<td>weekly</td>
<td>Datastream</td>
<td>522</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German Mark</td>
<td>1985 - 1995</td>
<td>weekly</td>
<td>Datastream</td>
<td>522</td>
<td>0.063</td>
<td>0.021</td>
<td>0.212</td>
<td>-1.451</td>
<td>-1.440×10^{-5}</td>
<td>0.002</td>
<td>0.307</td>
<td>8.787</td>
</tr>
<tr>
<td>Japanese Yen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.045</td>
<td>0.020</td>
<td>0.116</td>
<td>-0.584</td>
<td>-1.338×10^{-4}</td>
<td>0.002</td>
<td>-0.646</td>
<td>3.362</td>
</tr>
<tr>
<td>UK Sterling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.088</td>
<td>0.032</td>
<td>0.072</td>
<td>-1.185</td>
<td>-1.002×10^{-4}</td>
<td>0.003</td>
<td>-0.772</td>
<td>37.632</td>
</tr>
<tr>
<td>US Dollar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.062</td>
<td>0.019</td>
<td>-0.094</td>
<td>-1.018</td>
<td>-4.618×10^{-5}</td>
<td>0.002</td>
<td>-0.257</td>
<td>10.099</td>
</tr>
</tbody>
</table>
Table II

A Comparison of the Sampling Properties of the Non-Gaussian Estimation Procedure for the Generic Stochastic Volatility Model

We compare the sampling properties of the non-Gaussian estimation procedure with Jacquier, Polson, and Rossi’s (1994) Bayesian estimation procedure and with Sandmann and Koopman’s (1996) Monte Carlo based estimation procedure for the generic stochastic volatility model:

\[ \eta_t = \sqrt{h_{t-1}} \]
\[ \ln h_t = \alpha + \delta \ln h_{t-1} + \epsilon_t. \]

We choose values of the model parameters \( \alpha, \delta, \) and \( \sigma_\epsilon \) to investigate the sensitivity of these estimation procedures to the size of the stochastic volatility component by varying the squared coefficient of variation, \( \text{var}(h)/E(h)^2 \), and to the persistence of the stochastic component by varying \( \delta \). All parameter value combinations are consistent with \( E(h) = 0.009 \) or approximately a 20% annualized standard deviation in weekly data. Time series of length \( n=500 \) are simulated and the resultant sampling distributions are based on 500 replications. We report the average estimated parameter values and corresponding standard deviations are in parentheses. The results for the Bayesian estimator are reproduced from Jacquier, Polson, and Rossi (1994), Table 7, while the results for the Monte Carlo based estimator are reproduced from Sandmann and Koopman (1996), Table 2.

<table>
<thead>
<tr>
<th>( \text{Var}(h)/E(h)^2 = 10 )</th>
<th>( \alpha )</th>
<th>( \delta )</th>
<th>( \sigma_\epsilon )</th>
<th>( \alpha )</th>
<th>( \delta )</th>
<th>( \sigma_\epsilon )</th>
<th>( \alpha )</th>
<th>( \delta )</th>
<th>( \sigma_\epsilon )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayes</td>
<td>-0.80</td>
<td>0.92</td>
<td>0.56</td>
<td>-0.40</td>
<td>0.95</td>
<td>0.4835</td>
<td>-1.642</td>
<td>0.98</td>
<td>0.308</td>
</tr>
<tr>
<td>Monte Carlo</td>
<td>(.22)</td>
<td>(.03)</td>
<td>(.12)</td>
<td>(.16)</td>
<td>(.02)</td>
<td>(.06)</td>
<td>(.08)</td>
<td>(.01)</td>
<td>(.06)</td>
</tr>
<tr>
<td>Non-normal</td>
<td>-1.91</td>
<td>0.89</td>
<td>0.69</td>
<td>-0.94</td>
<td>0.94</td>
<td>0.49</td>
<td>-2.62</td>
<td>0.97</td>
<td>0.32</td>
</tr>
<tr>
<td>( \text{Var}(h)/E(h)^2 = 1 )</td>
<td>-0.736</td>
<td>0.92</td>
<td>0.363</td>
<td>-0.368</td>
<td>0.95</td>
<td>0.26</td>
<td>-1.1472</td>
<td>0.98</td>
<td>0.166</td>
</tr>
<tr>
<td>Bayes</td>
<td>-0.87</td>
<td>0.88</td>
<td>0.35</td>
<td>-0.56</td>
<td>0.92</td>
<td>0.28</td>
<td>-0.22</td>
<td>0.97</td>
<td>0.23</td>
</tr>
<tr>
<td>Monte Carlo</td>
<td>(.34)</td>
<td>(.05)</td>
<td>(.07)</td>
<td>(.34)</td>
<td>(.05)</td>
<td>(.07)</td>
<td>(.14)</td>
<td>(.02)</td>
<td>(.08)</td>
</tr>
<tr>
<td>Non-normal</td>
<td>-0.60</td>
<td>0.90</td>
<td>0.34</td>
<td>-0.33</td>
<td>0.95</td>
<td>0.25</td>
<td>-0.16</td>
<td>0.97</td>
<td>0.16</td>
</tr>
<tr>
<td>( \text{Var}(h)/E(h)^2 = .1 )</td>
<td>-0.706</td>
<td>0.92</td>
<td>0.135</td>
<td>-0.353</td>
<td>0.95</td>
<td>0.0964</td>
<td>-1.141</td>
<td>0.98</td>
<td>0.061</td>
</tr>
<tr>
<td>Bayes</td>
<td>-1.54</td>
<td>0.78</td>
<td>0.15</td>
<td>-1.12</td>
<td>0.84</td>
<td>0.12</td>
<td>-0.66</td>
<td>0.91</td>
<td>0.14</td>
</tr>
<tr>
<td>Monte Carlo</td>
<td>(.35)</td>
<td>(.19)</td>
<td>(.08)</td>
<td>(.15)</td>
<td>(.16)</td>
<td>(.07)</td>
<td>(.03)</td>
<td>(.12)</td>
<td>(.10)</td>
</tr>
<tr>
<td>Non-normal</td>
<td>-0.88</td>
<td>0.85</td>
<td>0.10</td>
<td>-0.64</td>
<td>0.89</td>
<td>0.09</td>
<td>-0.45</td>
<td>0.92</td>
<td>0.07</td>
</tr>
<tr>
<td>( \text{Var}(h)/E(h)^2 = .01 )</td>
<td>-1.22</td>
<td>0.83</td>
<td>0.16</td>
<td>-0.79</td>
<td>0.89</td>
<td>0.12</td>
<td>-0.42</td>
<td>0.94</td>
<td>0.08</td>
</tr>
</tbody>
</table>

33
Table III

Parameter Estimates of the Stochastic Volatility Model of Short-Term Riskless Interest Rate Dynamics

We estimate the parameters of the stochastic volatility model of short-term riskless interest rate dynamics:

\[
\begin{align*}
\ln(\text{res}_t^2) &= x_{t-1} + 2\gamma \ln(r_{t-1}) + \ln(\sigma^2_t) \\
x_t - \mu &= \beta(x_{t-1} - \mu) + \xi_{t,t}
\end{align*}
\]

where \( x_t \equiv \ln(\sigma^2_t) \) is the logarithm of the unobserved volatility \( \sigma^2_t \) and is assumed to follow an autoregressive specification, \( \text{res}_t \equiv \Delta r_t - a - b r_{t-1} \), and the parameters \( a \) and \( b \) are estimated by OLS with their corresponding standard errors in parentheses. The short-term riskless interest rate is proxied by one month rates of interest described in Table I. All interest rate data are geometrically scaled. The model is cast in state-space form and the resultant likelihood function is evaluated using an integration-based filter which assumes that \( \ln(\sigma^2_t) \equiv \log \chi^2 \) distributed and the prior on the state is normally distributed. Under the QML method the resultant likelihood function is evaluated assuming \( \ln(\sigma^2_t) \) is normally distributed. The first twelve observations of the samples are used to initialize both filtering procedures. Corresponding asymptotic standard errors, obtained from the inverse of the Hessian matrix evaluated at the respective optima, are in parentheses.

<table>
<thead>
<tr>
<th>one month interest rate</th>
<th>frequency</th>
<th>( a )</th>
<th>( b )</th>
<th>( \mu )</th>
<th>( \beta )</th>
<th>( \xi )</th>
<th>( \gamma )</th>
<th>( \mu )</th>
<th>( \beta )</th>
<th>( \xi )</th>
<th>( \gamma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Treasury bill</td>
<td>monthly</td>
<td>.063</td>
<td>.943</td>
<td>.056</td>
<td>.928</td>
<td>.0318</td>
<td>.651</td>
<td>.452</td>
<td>.892</td>
<td>.310</td>
<td>.919</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.022)</td>
<td>(.019)</td>
<td>(.281)</td>
<td>(.043)</td>
<td>(.094)</td>
<td>(.255)</td>
<td>(.215)</td>
<td>(.082)</td>
<td>(.144)</td>
<td>(.307)</td>
</tr>
<tr>
<td>UK interbank</td>
<td>monthly</td>
<td>.031</td>
<td>.969</td>
<td>-6.024</td>
<td>.476</td>
<td>1.354</td>
<td>.261</td>
<td>.965</td>
<td>.671</td>
<td>.879</td>
<td>.150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.018)</td>
<td>(.017)</td>
<td>(.216)</td>
<td>(.150)</td>
<td>(.206)</td>
<td>(.319)</td>
<td>(.227)</td>
<td>(.238)</td>
<td>(.388)</td>
<td>(.354)</td>
</tr>
<tr>
<td></td>
<td>weekly</td>
<td>.006</td>
<td>.994</td>
<td>-7.879</td>
<td>.711</td>
<td>1.128</td>
<td>.486</td>
<td>.797</td>
<td>.731</td>
<td>1.175</td>
<td>.617</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.004)</td>
<td>(.004)</td>
<td>(.134)</td>
<td>(.053)</td>
<td>(.103)</td>
<td>(.202)</td>
<td>(.150)</td>
<td>(.069)</td>
<td>(.175)</td>
<td>(.232)</td>
</tr>
<tr>
<td>London Euro-currencies</td>
<td>weekly</td>
<td>.003</td>
<td>.997</td>
<td>-8.169</td>
<td>.759</td>
<td>1.033</td>
<td>.425</td>
<td>.848</td>
<td>.527</td>
<td>1.880</td>
<td>.618</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.004)</td>
<td>(.003)</td>
<td>(.211)</td>
<td>(.052)</td>
<td>(.140)</td>
<td>(.299)</td>
<td>(.202)</td>
<td>(.072)</td>
<td>(.158)</td>
<td>(.296)</td>
</tr>
<tr>
<td>German Mark</td>
<td></td>
<td>-.002</td>
<td>.998</td>
<td>-7.197</td>
<td>.623</td>
<td>1.201</td>
<td>.310</td>
<td>.782</td>
<td>.996</td>
<td>.066</td>
<td>-.052</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.005)</td>
<td>(.004)</td>
<td>(.165)</td>
<td>(.080)</td>
<td>(.126)</td>
<td>(.127)</td>
<td>(.194)</td>
<td>(.008)</td>
<td>(.039)</td>
<td>(.214)</td>
</tr>
<tr>
<td>Japanese Yen</td>
<td></td>
<td>.003</td>
<td>.996</td>
<td>-8.414</td>
<td>.759</td>
<td>.044</td>
<td>.398</td>
<td>.848</td>
<td>.843</td>
<td>.682</td>
<td>.384</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.004)</td>
<td>(.003)</td>
<td>(.192)</td>
<td>(.068)</td>
<td>(.128)</td>
<td>(.274)</td>
<td>(.217)</td>
<td>(.068)</td>
<td>(.176)</td>
<td>(.303)</td>
</tr>
<tr>
<td>UK Sterling</td>
<td></td>
<td>.005</td>
<td>.994</td>
<td>-8.011</td>
<td>.562</td>
<td>1.280</td>
<td>.754</td>
<td>.810</td>
<td>.701</td>
<td>1.032</td>
<td>.775</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.004)</td>
<td>(.004)</td>
<td>(.103)</td>
<td>(.108)</td>
<td>(.174)</td>
<td>(.214)</td>
<td>(.182)</td>
<td>(.121)</td>
<td>(.261)</td>
<td>(.257)</td>
</tr>
</tbody>
</table>

34
Table IV

Sampling Properties of the non-Gaussian and Quasi-Maximum Likelihood Estimation Procedures

We simulate interest rate and volatility series each of length $n = 500$ observations according to

$$
\Delta r_t = (a + br_{t-1})\Delta t + \sigma_t r_{t-1} \sqrt{\Delta t} \epsilon_{t,1} \\
\ln \sigma_t^2 - \mu = \beta (\ln \sigma_{t-1}^2 - \mu) \Delta t + \xi \sqrt{\Delta t} \epsilon_{t,2}
$$

for a variety of parameter values. After geometrically scaling the interest rate data, the parameters $a$ and $b$ are estimated by OLS. The remaining parameters of the stochastic volatility model are estimated using the non-Gaussian procedure and using quasi-maximum likelihood (QML). The experiment is repeated five hundred times and the sampling distributions of the estimators are presented. Panel A tabulates the results of the non-Gaussian procedure while Panel B tabulates the QML results for parameter values consistent with our estimation results using monthly data: $a = 0.06$, $b = 0.92$, $\mu = -6.0$, $\beta = 0.93$, $\xi = 0.32$, and $\gamma = 0.50$. We also investigate the sensitivity of the non-Gaussian procedure to varying $\gamma$ values, all else being equal. Panel C reports the sampling results when we increase $\gamma$ to 0.75 but use that value of $\mu$ which over five hundred replications gives the same sample mean and sample standard deviation of $r$ as obtained in the sampling experiment assuming $\gamma = 0.5$. Panel D reports these results when we similarly decrease $\gamma$ to 0.25. Panel E reports the sampling properties of the non-Gaussian procedure for parameter values consistent with our estimation results using weekly data: $a = 0.005$, $b = 0.99$, $\mu = -8.0$, $\beta = 0.50$, $\xi = 1.2$, and $\gamma = 0.75$.

### A: Non-Gaussian procedure for representative monthly parameters

<table>
<thead>
<tr>
<th>Assumed Parameter Value</th>
<th>Sample Mean</th>
<th>Sample Standard Deviation</th>
<th>Sample Skewness</th>
<th>Sample Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a = 0.06$</td>
<td>0.105</td>
<td>0.026</td>
<td>0.212</td>
<td>0.310</td>
</tr>
<tr>
<td>$b = 0.92$</td>
<td>0.909</td>
<td>0.025</td>
<td>-0.239</td>
<td>0.370</td>
</tr>
<tr>
<td>$\mu = -6.0$</td>
<td>-6.051</td>
<td>0.502</td>
<td>0.033</td>
<td>-1.406</td>
</tr>
<tr>
<td>$\beta = 0.93$</td>
<td>0.864</td>
<td>0.090</td>
<td>-4.868</td>
<td>42.253</td>
</tr>
<tr>
<td>$\xi = 0.32$</td>
<td>0.323</td>
<td>0.094</td>
<td>0.894</td>
<td>2.419</td>
</tr>
<tr>
<td>$\gamma = 0.5$</td>
<td>0.494</td>
<td>0.089</td>
<td>0.143</td>
<td>-0.379</td>
</tr>
</tbody>
</table>

### B: QML for representative monthly parameters

<table>
<thead>
<tr>
<th>Assumed Parameter Value</th>
<th>Sample Mean</th>
<th>Sample Standard Deviation</th>
<th>Sample Skewness</th>
<th>Sample Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu = -6.0$</td>
<td>-6.197</td>
<td>0.655</td>
<td>-0.113</td>
<td>0.466</td>
</tr>
<tr>
<td>$\beta = 0.93$</td>
<td>0.807</td>
<td>0.191</td>
<td>-2.841</td>
<td>12.785</td>
</tr>
<tr>
<td>$\xi = 0.32$</td>
<td>0.397</td>
<td>0.261</td>
<td>1.512</td>
<td>2.286</td>
</tr>
<tr>
<td>$\gamma = 0.5$</td>
<td>0.467</td>
<td>0.118</td>
<td>0.042</td>
<td>0.524</td>
</tr>
</tbody>
</table>

35
Table IV (continued)

C: Non-Gaussian procedure assuming $\gamma = 0.75$

<table>
<thead>
<tr>
<th>Assumed Parameter Value</th>
<th>Sample Mean</th>
<th>Sample Standard Deviation</th>
<th>Sample Skewness</th>
<th>Sample Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu = -4.78$</td>
<td>-4.893</td>
<td>0.582</td>
<td>-0.061</td>
<td>-5.43</td>
</tr>
<tr>
<td>$\beta = 0.93$</td>
<td>0.864</td>
<td>0.090</td>
<td>-4.611</td>
<td>38.534</td>
</tr>
<tr>
<td>$\xi = 0.32$</td>
<td>0.323</td>
<td>0.095</td>
<td>0.863</td>
<td>2.239</td>
</tr>
<tr>
<td>$\gamma = 0.75$</td>
<td>0.734</td>
<td>0.105</td>
<td>0.014</td>
<td>-0.495</td>
</tr>
</tbody>
</table>

D: Non-Gaussian procedure assuming $\gamma = 0.25$

<table>
<thead>
<tr>
<th>Assumed Parameter Value</th>
<th>Sample Mean</th>
<th>Sample Standard Deviation</th>
<th>Sample Skewness</th>
<th>Sample Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu = -7.27$</td>
<td>-7.271</td>
<td>0.435</td>
<td>0.010</td>
<td>0.888</td>
</tr>
<tr>
<td>$\beta = 0.93$</td>
<td>0.865</td>
<td>0.089</td>
<td>-4.765</td>
<td>40.286</td>
</tr>
<tr>
<td>$\xi = 0.32$</td>
<td>0.323</td>
<td>0.094</td>
<td>0.844</td>
<td>2.357</td>
</tr>
<tr>
<td>$\gamma = 0.25$</td>
<td>0.233</td>
<td>0.077</td>
<td>0.204</td>
<td>0.226</td>
</tr>
</tbody>
</table>

E: Non-Gaussian procedure for representative weekly parameters

<table>
<thead>
<tr>
<th>Assumed Parameter Value</th>
<th>Sample Mean</th>
<th>Sample Standard Deviation</th>
<th>Sample Skewness</th>
<th>Sample Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a=0.005$</td>
<td>0.014</td>
<td>0.003</td>
<td>0.475</td>
<td>0.627</td>
</tr>
<tr>
<td>$b=0.99$</td>
<td>0.988</td>
<td>0.003</td>
<td>-0.380</td>
<td>0.619</td>
</tr>
<tr>
<td>$\mu = -8.0$</td>
<td>-8.016</td>
<td>0.397</td>
<td>-0.231</td>
<td>0.919</td>
</tr>
<tr>
<td>$\beta = 0.5$</td>
<td>0.484</td>
<td>0.099</td>
<td>-0.520</td>
<td>0.451</td>
</tr>
<tr>
<td>$\xi = 1.20$</td>
<td>1.173</td>
<td>0.118</td>
<td>-0.168</td>
<td>0.241</td>
</tr>
<tr>
<td>$\gamma = 0.75$</td>
<td>0.734</td>
<td>0.190</td>
<td>-0.171</td>
<td>1.652</td>
</tr>
</tbody>
</table>
Table V

Parameter Estimates of the EGARCH Model of Short-Term Riskless Interest Rate Dynamics

We estimate the parameters of the EGARCH model of short-term riskless interest rate dynamics:

\[ r_{st} = \exp \left( \frac{1}{2} x_t \right) r_t^\gamma \]

\[ x_t - \mu = \beta (x_{t-1} - \mu) + \sigma_t \epsilon_{t-1} + g_1 (| \epsilon_{t-1} | - E [ | \epsilon_{t-1} | ]) \]

where \( r_{st} = \Delta r_t - (a + b \epsilon_{t-1}) \) and the parameters \( a \) and \( b \) are estimated by OLS. Both normally distributed and \( t \)-distributed innovations with \( \nu \) degrees of freedom are assumed. The short-term riskless interest rate is proxied by one-month rates of interest described in Table I. All interest rate data are geometrically scaled. Asymptotic standard errors obtained from the inverse of the Hessian matrix evaluated at the respective optima, are in parentheses. Convergence was not achieved for the EGARCH model when using the monthly UK interbank rate series. \( V_{\text{Vuong}} \) is the value of the Vuong (1989) statistic for testing the EGARCH model against the stochastic volatility model.

<table>
<thead>
<tr>
<th>one month interest rate</th>
<th>frequency</th>
<th>( \mu )</th>
<th>( \sigma )</th>
<th>( \beta )</th>
<th>( g_1 )</th>
<th>( \gamma )</th>
<th>( V_{\text{Vuong}} )</th>
<th>( \mu )</th>
<th>( \beta )</th>
<th>( g_1 )</th>
<th>( \gamma )</th>
<th>( \nu )</th>
<th>( V_{\text{Vuong}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Treasury bill</td>
<td>monthly</td>
<td>-3.999</td>
<td>0.904</td>
<td>0.106</td>
<td>0.249</td>
<td>-1.12</td>
<td>1.39</td>
<td>-5.067</td>
<td>0.960</td>
<td>0.183</td>
<td>0.206</td>
<td>-4.48</td>
<td>6.598</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.327)</td>
<td>(.012)</td>
<td>(.098)</td>
<td>(.068)</td>
<td>(.241)</td>
<td></td>
<td>(.543)</td>
<td>(.016)</td>
<td>(.179)</td>
<td>(.088)</td>
<td>(.779)</td>
<td>(2.361)</td>
</tr>
<tr>
<td>UK Interbank</td>
<td>weekly</td>
<td>-4.593</td>
<td>0.638</td>
<td>-0.207</td>
<td>0.618</td>
<td>0.654</td>
<td>3.14*</td>
<td>-10.549</td>
<td>0.962</td>
<td>-0.019</td>
<td>0.154</td>
<td>.477</td>
<td>2.021</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.243)</td>
<td>(.056)</td>
<td>(.256)</td>
<td>(.054)</td>
<td>(.218)</td>
<td></td>
<td>(.740)</td>
<td>(.013)</td>
<td>(.020)</td>
<td>(.021)</td>
<td>(.436)</td>
<td>(1.18)</td>
</tr>
<tr>
<td>London Euro-currencies:</td>
<td>weekly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German Mark</td>
<td></td>
<td>-7.106</td>
<td>0.709</td>
<td>0.110</td>
<td>0.615</td>
<td>0.529</td>
<td>3.32*</td>
<td>-9.077</td>
<td>0.813</td>
<td>0.086</td>
<td>0.390</td>
<td>0.069</td>
<td>2.822</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.709)</td>
<td>(.040)</td>
<td>(.049)</td>
<td>(.082)</td>
<td>(.016)</td>
<td></td>
<td>(.764)</td>
<td>(.060)</td>
<td>(.041)</td>
<td>(.062)</td>
<td>(.285)</td>
<td>(.416)</td>
</tr>
<tr>
<td>Japanese Yen</td>
<td></td>
<td>-6.312</td>
<td>0.639</td>
<td>0.229</td>
<td>0.341</td>
<td>0.146</td>
<td>5.16*</td>
<td>-8.492</td>
<td>0.854</td>
<td>0.132</td>
<td>0.210</td>
<td>0.319</td>
<td>2.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.092)</td>
<td>(.113)</td>
<td>(.057)</td>
<td>(.108)</td>
<td>(.066)</td>
<td></td>
<td>(.375)</td>
<td>(.066)</td>
<td>(.038)</td>
<td>(.049)</td>
<td>(.145)</td>
<td>(.928)</td>
</tr>
<tr>
<td>UK Sterling</td>
<td></td>
<td>-6.947</td>
<td>0.833</td>
<td>0.036</td>
<td>0.444</td>
<td>0.990</td>
<td>4.37*</td>
<td>-8.881</td>
<td>0.902</td>
<td>0.002</td>
<td>0.181</td>
<td>0.249</td>
<td>2.241</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.167)</td>
<td>(.054)</td>
<td>(.089)</td>
<td>(.066)</td>
<td>(.220)</td>
<td></td>
<td>(.397)</td>
<td>(.037)</td>
<td>(.027)</td>
<td>(.057)</td>
<td>(.314)</td>
<td>(.273)</td>
</tr>
<tr>
<td>US Dollar</td>
<td></td>
<td>-6.655</td>
<td>0.860</td>
<td>0.077</td>
<td>0.397</td>
<td>0.114</td>
<td>3.89*</td>
<td>-8.998</td>
<td>0.825</td>
<td>0.050</td>
<td>0.164</td>
<td>0.762</td>
<td>2.074</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.163)</td>
<td>(.034)</td>
<td>(.047)</td>
<td>(.056)</td>
<td>(.165)</td>
<td></td>
<td>(.242)</td>
<td>(.071)</td>
<td>(.036)</td>
<td>(.052)</td>
<td>(.220)</td>
<td>(.286)</td>
</tr>
</tbody>
</table>

\* denotes significance at the 5% level.
Figure 1: We use the Euro-mark interest rate series and plot for each sample point $t$ the difference between the maximized log-likelihood values, based on data through $t$-1, of the stochastic volatility and EGARCH models evaluated at their respective sample-wide maximum likelihood estimates. Panel A displays the results assuming an EGARCH model with normal innovations and Panel B displays the results assuming an EGARCH model with $t$-distributed innovations. A positive (negative) difference at $t$ is consistent with the stochastic volatility model being better (less) able to explain the data at $t$.

A: Stochastic Volatility Model versus EGARCH Model with Normally Distributed Innovations

B: Stochastic Volatility Model versus EGARCH Model with $t$-Distributed Innovations