Discussion Paper

Discussion Paper No. 93-03

Markov-Switching Models for Exchange-Rate Dynamics and the Pricing of Foreign-Currency Options

Juergen Kaehler and Volker Marnet

ZEW

Zentrum für Europäische Wirtschaftsforschung GmbH

International Finance Series

MARKOV-SWITCHING MODELS FOR EXCHANGE-RATE DYNAMICS AND THE PRICING OF FOREIGN-CURRENCY OPTIONS

by

Juergen Kaehler[*] and Volker Marnet[**]

[*]ZEW and University of Mannheim
[**]ZEW

January 1993

1. Introduction

The problem of decomposing an empirical frequency distribution into several components is a very old statistical problem. We are approaching the centenary of Karl Pearson's classical paper (Pearson, 1894) in which he introduced the method of moments to estimate the parameters of a two-component mixture of normal distributions and in which he applied the model to the distributions of some measures of the forehead of crabs and of the dental distance in prawns.

Over the years, the statistical analysis of finite mixture distributions has found a great variety of applications in such diverse fields as fisheries research, geology, crime and comet frequencies and medicine [1]. In finance, mixtures of normal distributions have been applied to model the price dynamics of stocks (see Barnea and Downes (1973), Ball and Torous (1983), Fielitz and Rozelle (1983), Kon (1984), Akgiray and Booth (1987), and Akgiray, Booth and Loistl (1989)) and exchange rates (see Boothe and Glassman (1987), Akgiray and Booth (1988), and Tucker and Pond (1988))[2].

From a theoretical perspective, mixtures of distributions can be motivated as models of information arrival on financial markets (see Kon (1984)). It is typically assumed that the model consists of two (or more) distributions with different variances. Drawings from the high-variance distribution represent information events while drawings from the low-variance distribution represent non-information periods which can be associated with background noise of normal trading. Alternatively, Kon (1984) suggested a three-components mixture model for stock returns based on the idea that the returns are drawn from a non-information distribution, a firm-specific information distribution, and a market-wide information distribution. Mixture models may also be related to "anomalities" of the stock market such as excess returns and volatilities on Mondays and other calendar effects. In applications to exchange rates, it has been suggested that the components of the mixture represent periods with Central Bank intervention and periods without intervention or, alternatively, that the two components of a mixture can be associated with "news" form the two countries whose exchange rate is considered (Friedman and Vandersteel (1982)).

From a statistical perspective, mixture models may be motivated as models which imply leptokurtosis. Extensive empirical analysis has revealed that leptokurtosis is a strong and robust empirical regularity of short-run price dynamics in financial markets. Under time-aggregation, however, leptokurtosis vanishes, i. e. the null hypothesis of a normal distribution can, in general, be rejected for daily and weekly data but not for monthly and quarterly data (see Fama (1976)). It can be shown (see section 2) that

---

1 See Everitt and Hand (1981) and Titterington, Smith and Makov (1985) for comprehensive surveys.

2 It is quite remarkable that the finance literature on mixture models seems to have been unaware of the statistics literature on the same subject. For instance, Ball and Torous (1983) and Kon (1984) deal with estimation problems which had long been solved.

arbitrary scale-mixtures of normal distributions imply leptokurtosis. Hence, this class of models is compatible with the strong stylized facts of leptokurtosis and convergence to normality. The latter follows, of course, from the central limit theorem.

There are a number of competing probability models, however, which also imply leptokurtosis. The work along these lines was initiated by Mandelbrot (1963) who introduced the family of stable Paretian distributions into economics and finance. However, stable Paretian distributions are not compatible with convergence to normality and, besides, have some unattractive analytical features. The other main competitors are the Student's distribution, introduced into the modelling of financial data by Praetz (1972), and the compound Poisson process, introduced by Press (1967). It is interesting to note that all four probability models can be viewed within the framework of scale-compounded normal distributions where the variance is random with an independent distribution. The mixture model attaches a multinomial distribution to the variance, the Student's distribution attaches an inverted gamma-distribution, the compound Poisson process attaches a Poisson distribution and the stable Paretian distributions attach positive stable distributions to the variance. In this way, the choice among the models can be regarded as a choice between variance functions for a normal distribution.

There is mounting empirical evidence for the rejection of the stable Paretian distribution as a valid model for price dynamics in financial markets (see Lau et al. (1990), Jansen and de Vries (1991), and Kaehler (1991). In comparative studies with daily and weekly data on stock returns and exchange-rate dynamics, it turned out that the compound Poisson process and mixtures of normal distributions are superior to stable distributions and Student's distributions in their fit to the data (see Kon (1984), Akgiray and Booth (1987,1988), Boothe and Glassman (1987), Tucker and Pond (1988), and Akgiray, Booth and Loistl (1989)).

In this paper we will consider the mixture of normal distributions and a generalization of it for the modelling of financial data and we will apply the model to daily, weekly, monthly, and quarterly exchange-rate dynamics. In Section 2 we show that heteroskedasticity is another strong empirical regularity of the data and in Section 3 we show that the mixture model does not capture this regularity. Section 4 introduces an extension of the mixture model which incorporates heteroskedasticity by letting drawings from the component distributions follow a first-order Markov chain. This Markov-switching model for mixture distributions is due to Lindgren (1978) based on the work of Baum et al.(1970). In a series of papers, Hamilton extended the model and adapted it to the modelling of interest rates, exchange-rate and the business cycle (see Hamilton (1988, 1989, 1990, 1991a,b), and Engel and Hamilton (1990)).

The stochastic specification of financial models is of fundamental importance in almost every branch of finance. We study the implications of mixture models and Markov-switching models for the pricing of foreign-currency call options in Section 4. In Section 5 we draw some conclusions from our study and suggest directions for future research.

2. Stylized Facts of Exchange-Rate Dynamics

The data to be analysed are the exchange rates of the U.S. dollar against the German mark, the British pound, the Swiss franc and the Japanese yen. The data are on a daily basis but also weekly, monthly and quarterly data are used. For these series, end-of-period data were derived from daily exchange rates. The data are from July 1st, 1974 to June 28th, 1991. Due to differences in bank holidays between countries, there are different numbers of observations in the daily data: 4260 for the mark, 4299 for the pound, 4266 for the franc, and 4226 for the yen. For all currencies, the number of observations in the weekly series is 886, in the monthly series it is 203 and in the quarterly series it is 68. Data sources are the IMF's International Financial Statistics and the monthly reports of the Swiss National Bank. The exchange-rate dynamics are analysed in the form of $x_t = 100\ (e_t - e_{t-1})$ where e_t is the logarithm of the exchange rate at time t.

Table 1 reports some descripitive statistics and tests for the daily, weekly, monthly and quarterly exchange-rate dynamics. Overall, the results are in line with earlier studies of price dynamics in financial markets (see e.g. Taylor (1986)). First, the means of the series are, in general, not significantly different from zero. It is only in the daily and weekly yen series that we find some weak evidence against a mean of zero. But note that the underlying t-test assumes normality and that this assumption is very questionable for these data as will be shown below.

Second, there is some evidence of negative skewness (β_1, defined as the 3rd standardized moment) in the daily and weekly data of the mark, the franc, and the yen. However, there is no strong evidence against the null hypothesis of a symmetric distribution $(H_0{:}\beta_1 = 0)$ in monthly and quarterly data.

Third, we found very strong leptokurtosis in all daily and weekly series. It has repeatedly been found that returns and price movements in financial markets have excess kurtosis, i.e. kurtosis which is significantly greater than 3 (the value for a normal distribution). Kurtosis β_2 , defined as the ratio of the 4th central moment to the square of the variance, increases both with excessive mass in the tails or at the centre of the distribution. A test of $H_0{:}\beta_1 = 3$ is a test of mesokurtosis with the two-sided alternatives of platykurtic $(\beta_2 < 3)$ and leptokurtic $(\beta_2 > 3)$ distributions [3]. Whereas leptokurtosis is highly significant in daily and weekly data, it is only significant (at the 5 percent level) for the monthly series of the mark, the pound and the franc and all distributions

3 Since the skewness and kurtosis statistics have unknown distributions and show strong non-normality even in large samples, we applied to both statistics some transformations (as described by D'Agostino (1986)) to improve the approximation to a standard normal distribution. We used the S_u approximation for β_1 and the Anscombe-Glynn approximation for β_2 .

Table 1
Statistical properties of exchange-rate dynamics

		mark	pound	franc	yen
day	mean	-0.008	0.009	-0.015	-0.017 *
	variance	0.473	0.545	0.660	0.406
	skewness	-0.225 ***	-0.048	-0.086 **	-0.450 ***
	kurtosis	7.423 ***	7.341 ***	8.095 ***	7.363 ***
	AD	26.675 ***	44.891 ***	32.603 ***	42.780 ***
	$Q_x(40)$	56.684 **	63.669 **	59.344 **	80.965 ***
	$Q_{xx}(40)$	567.435 ***	968.832 ***	926.047 ***	760.990 ***
week	mean	-0.039	0.044	-0.073	-0.083 *
	variance	2.153	2.137	2.891	1.786
	skewness	-0.209 **	-0.015	-0.229 ***	-0.728 ***
	kurtosis	5.380 ***	6.242 ***	4.749 ***	6.186 ***
	AD	5.992 ***	4.386 ***	4.976 ***	9.683 ***
	$Q_x(40)$	46.956	38.332	37.617	98.205 ***
	$Q_{xx}(40)$	168.753 ***	194.851 ***	172.655 ***	127.254 ***
month	mean	-0.175	0.188	-0.318	-0.379
	variance	11.770	11.329	14.707	11.381
	skewness	0.176	-0.315 *	0.215	-0.224
	kurtosis	4.084 **	3.850 **	4.003 **	3.470
	AD	0.784	0.488	0.482	2.038 **
	$Q_x(40)$	37.127	28.881	40.356	47.553
	$Q_{xx}(40)$	26.653	26.113	28.729	39.83
quarter	mean	-0.569	0.543	-0.949	-1.153
	variance	40.155	32.942	53.325	37.544
	skewness	0.174	0.053	-0.361	-0.450
	kurtosis	2.596	2.437	2.767	2.667
	AD	0.210	0.556	0.297	0.583
	$Q_x(15)$	17.748	21.545	9.370	14.202
	$Q_{xx}(15)$	15.146	11.169	10.256	9.728

Significance levels: * 10 percent, ** 5 percent, *** 1 percent.

of quarterly data are platykurtic (but not significantly). We may conclude, therefore, that leptokurtosis is a phenomenon of short-run exchange-rate dynamics and that there is convergence to normality under time-aggregation.

In order to further investigate deviations from normality, we applied the Anderson-Darling (AD) test for normality which, like the well-known Kolmogorov-Smirnov test, is based upon the vertical difference between the empirical distribution function and the theoretical distribution function. But the AD test has more power than the Kolmogorov-Smirnov test (see Stephens (1986)). As Table 1 shows, the results from the AD test are quite similar to those of the kurtosis test. Normality is overwhelmingly rejected for daily and weekly data. It is quite peculiar, however, that the only monthly series for which normality is rejected is the yen series whereas in the kurtosis test, this series was the only monthly series for which mesokurtosis could not be rejected. This demonstrates that these tests are sensitive to different distributional aspects. The H_0 of normality cannot be rejected with the AD test for any of the quarterly series.

Table 1 also reports results from a test of serial independence. We applied the Ljung-Box statistic $Q_x(M)$ which is based on autocorrelation function (ACF) of x_t where M is the highest lag in the ACF. Under the H_O of white noise, $Q_x(M)$ has asymptotically a χ^2 distribution with M degrees of freedom. Table 1 shows that there is strong serial dependence in the daily and weekly yen series. At the 5 percent significance level, Q_x is also significant for the other three daily series, but for all other series we find no evidence for serial dependence. Note, however, that the results for the daily series may be biased. In the presence of heteroscedasticity, the Bartlett standard errors of the ACF are downward biased and the Ljung-Box statistic is upward biased, i.e. we would reject the H_0 of independence too often. Diebold (1988) has shown that heteroscedasticity does indeed bias tests for serial independence with weekly data.

In order to quantify the heteroskedasticity of the series, we computed the ACF of the squared data x_t^2. McLeod and Li (1983) have established that under the H_0 of white noise, the standard errors of squared-data autocorrelations are the same as for the usual ACF. Hence also the Ljung-Box statistic $Q_{xx}(M)$ for squared data is applicable without modification. As with the test for normality, we find that there are marked differences between short-run, i.e. daily and weekly , and medium-run, i.e. monthly and quarterly exchange-rate dynamics. Whereas there is extremely strong serial dependence of volatility in daily and weekly data, this dependence disappears completely in monthly and quaterly data. Furthermore, all individual autocorrelation coeffecients for the daily data are positive and they are significant up to $M = 40$. For the weekly data, some autocorrelation coefficients are negative but all significant coefficients are positive. Hence, there is a strong clustering of small and of large exchange-rate fluctuations in the short-run data.

To summarize the statistical properties, we find very strong leptokurtosis and heteroskedasticity in daily and weekly but a convergence to Gaussian white noise under time-aggregation. In the following two sections, we shall aim to build a model compatible with these three empirical regularities.

3. Mixtures of Normal Distributions

As noted in the introduction, the mixture of normal distributions has often been applied to capture the stylized facts of price dynamics in financial markets. A finite mixture of normal distributions is defined by:

(1) $$f(x_t \mid \Theta) = \sum_{i=1}^{I} p_i f_i(x_t \mid \mu_i, \sigma_i)$$

with

(2) $$\sum_{i=1}^{I} p_i = 1 \quad \text{and} \quad 0 < p_i < 1 \quad \text{for all} i$$

and

(3) $$f_i(x_t \mid \mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left\{-\frac{(x_t - \mu_i)^2}{2\sigma_i^2}\right\}.$$

In equation (1), Θ is the parameter vector $\Theta = (p_1, \ldots, p_{I-1}, \mu_1, \ldots, \mu_I, \sigma_1, \ldots, \sigma_I)$ where p_I is redundant because of the restriction that the probabilities p_i sum to 1. Scale mixtures of normal distributions are defined as special cases of (1) with $\mu_1 = \ldots = \mu_I$.

It is straightforward to show that scale mixtures for arbitrary I are leptokurtic[4]. It is more convenient here to define leptokurtosis in terms of the fourth cumulant $\varkappa_4$ which is related to central moments by

(4) $$\varkappa_4 = \nu_4 - 3\nu_2^2,$$

where ν_k is the k-th central moment. Leptokurtosis of a random variable X can also be defined by the condition $\varkappa_4 > 0$. It is easy to show that for scale mixtures

(5) $$\nu_2^2(X) = \left(\sum_{i=1}^{I} p_i \sigma_i^2\right)^2$$

and

(6) $$\nu_4(X) = 3 \sum_{i=1}^{I} p_i \sigma_i^4.$$

Inserting both terms into (4) yields

4 Gridgeman (1970) proved only the peakedness of general scale mixtures of normal distributions. Mean mixtures of normal distributions are not generally leptokurtic. For instance, the kurtosis of a two-component normal mixture with $\sigma_1 = \sigma_2 = 1$ and $\mu_1 = -\mu_2 = 1$ is 1.25, i.e. this distribution is platykurtic.

$$\varkappa_4(X) = 3\left[\sum_{i=1}^{I} p_i\sigma_i^4 - \left(\sum_{i=1}^{I} p_i\sigma_i^2\right)^2\right] \tag{7}$$

$$= 3var(\sigma^2) > 0$$

since the variance σ^2 is non-degenerate by assumption. Furthermore, the convergence to normality under addition follows simply from the central limit theorem.

The mixture of normal distributions of equation (1) can only be estimated if I is specified. In order to identify I, we applied the Schwarz information criterion (SIC) and found that the optimal I was either 2 or 3 for the series of exchange rates. There were also some numerical problems when I was greater than 3. Distinguishing between mean mixtures, which impose the restriction $\sigma_1 = \ldots = \sigma_I$, scale mixtures, which impose the restriction $\mu_1 = \ldots = \mu_I$, and mean-scale mixtures, which impose no such restrictions, we found by applying the SIC that mean mixtures are never optimal and that mean-scale mixtures are only optimal for the two yen series. Otherwise, scale mixtures with either two or three components were optimal for the other series. However, in most three-components models there was one component with a very small probability p_i of 0.05 or less. For instance, for the weekly franc series we estimated p_2 = 0.02 along with μ_2 = 2.63 and σ_2^2 =0.05. Compared with the overall mean of μ = -0.07 and the overall variance of σ^2 =2.89, it is clear that this component picks up a bunch of strong depreciations of the Swiss franc. The small values of the p_2 and σ_2^2 are disturbing because they indicate that there might be a problem with singularities, or near singularities, in the likelihood function (see Titterington, Smith and Makov (1985)). Furthermore, one of the diagnostic tests for independence which we applied to the model (see below) could not be performed for some of the three-components models because of some low component probabilities p_i. Therefore, we report only the results from the two-component models in this paper. The estimates are given in Table 2 and asymptotic standard errors are in brackets.

Several important observations may be drawn from Table 2. First, there are no significant mean effects in the daily and weekly series of the mark, pound and franc but both yen series have components which are significantly different in their means with opposite signs. On the other hand, the components of all short-run exchange-rate series can clearly be distinguished with respect to their variances. The first component is always associated with the lower variance and σ_2^2 is larger than σ_1^2 by a factor of at least 5 in daily data and by a factor of at least 4 in a weekly data. In general the low-variance component has a higher probability than the second component, the only exception being the weekly mark series.

Table 2
Estimates of the mixture of normal distributions

		mark	pound	franc	yen
day	p	0.749 (0.053)	0.519 (0.045)	0.807 (0.030)	0.689 (0.035)
	μ_1	0.002 (0.012)	-0.010 (0.010)	0.010 (0.012)	0.032 (0.010)
	μ_2	-0.036 (0.041)	0.029 (0.022)	-0.121 (0.060)	-0.125 (0.034)
	σ_1^2	0.232 (0.024)	0.101 (0.016)	0.332 (0.021)	0.151 (0.012)
	σ_2^2	1.192 (0.145)	0.834 (0.057)	2.013 (0.207)	0.955 (0.073)
	$\tilde{\beta}_1$	-0.086	0.071	-0.192	-0.315
	$\tilde{\beta}_2$	5.330	4.954	6.077	5.578
	LR	400.05 ***	585.56 ***	524.80 ***	573.45 ***
	runs	6.704 ***	9.971 ***	7.950 ***	9.311 ***
	Markov	44.175 ***	99.056 ***	63.084 ***	86.564 ***
week	p	0.335 (0.084)	0.830 (0.141)	0.612 (0.094)	0.636 (0.073)
	μ_1	-0.015 (0.057)	0.025 (0.058)	0.071 (0.071)	0.139 (0.051)
	μ_2	-0.050 (0.076)	0.136 (0.262)	-0.300 (0.180)	-0.469 (0.147)
	σ_1^2	0.293 (0.143)	1.313 (0.319)	1.154 (0.224)	0.592 (0.108)
	σ_2^2	3.084 (0.332)	6.143 (2.625)	5.533 (0.836)	3.630 (0.504)
	$\tilde{\beta}_1$	-0.021	0.073	-0.236	-0.544
	$\tilde{\beta}_2$	4.125	5.172	4.662	5.145
	LR	75.627 ***	63.215 ***	65.848 ***	120.20 ***
	runs	3.399 ***	4.711 ***	2.247 **	5.662 ***
	Markov	11.310 ***	21.934 ***	4.994 **	31.336 ***
month	p	0.297 (0.200)	0.926 (0.173)	0.200 (0.138)	0.265 (0.184)
	μ_1	-0.225 (0.426)	0.188 (0.273)	0.047 (0.597)	0.364 (0.500)
	μ_2	-0.155 (0.365)	0.188 (3.003)	-0.409 (0.367)	-0.647 (0.520)
	σ_1^2	2.262 (2.218)	9.376 (1.982)	2.458 (2.140)	0.787 (1.366)
	σ_2^2	15.701 (3.498)	34.911 (44.730)	17.642 (2.910)	14.856 (2.992)
	$\tilde{\beta}_1$	0.015	0.000	-0.059	-0.216
	$\tilde{\beta}_2$	3.824	4.059	3.509	3.840
	LR	8.459 **	3.701	4.317	20.649 ***
	runs	-0.223	-0.234	-	-0.145
	Markov	0.058	0.046	-	0.058
quarter	p	0.955 (0.030)	0.736 (0.083)	0.935 (0.033)	0.750 (0.166)
	μ_1	-1.201 (0.752)	-2.013 (0.907)	0.149 (0.812)	1.471 (1.312)
	μ_2	12.723 (0.591)	7.683 (0.776)	-16.708 (0.763)	-9.005 (3.307)
	σ_1^2	32.604 (6.365)	17.339 (5.923)	37.541 (7.530)	16.537 (6.473)
	σ_2^2	0.734 (6.436)	5.429 (2.769)	2.044 (1.532)	15.904 (13.942)
	$\tilde{\beta}_1$	0.196	0.089	-0.379	-0.462
	$\tilde{\beta}_2$	2.921	2.225	3.053	2.763
	LR	3.870	5.129 *	6.698 **	5.258 *
	runs	0.554	-2.671 ***	0.236	2.166 **
	Markov	0.155	6.298 **	0.362	4.376 **

The results are somewhat different for medium-term exchange-rate dynamics, i.e. for monthly and quarterly data. The low-variance component has the smaller probability for these data (with the exception of the monthly pound series). It may also be surprising that the first component is associated with the high-variance state in quarterly data but note that the "ranking" of states is arbitrary and unimportant in this model. We also find some strong mean effects in the lower frequency data. All quarterly series have components which are significantly different in their means with opposite signs. Some of the mean effects, however, are disturbing. For instance, the second component for the quarterly mark series has a large mean, a small variance and a probability of 4.5 percent. With a total of 68 quarterly observations, this probability implies an expected number of approximately 3 observations from the second component in this series. This means that the second component represents a small number of large depreciations of the mark against the dollar. In more extreme cases, the mixture model may converge to a singularity where the mean of one component is equal to the value of one observation (often an extreme one) and where the variance of this components goes to zero. The likelihood function will then go to infinity and this is a major problem for maximum-likelihood estimation of the model. In order to avoid this singularity problem, one may try to keep all σ_i^2 away from zero through simple restrictions on the parameter space or through the introduction of a penalty function (which also has the interpretation of a Bayesian prior) as in Hamilton (1991a). Empirical applications often apply the restrictions of mean mixtures but we decided not to impose the restrictions of equal variances because we have good reason to believe that scale effects are more important than mean effects in our data. The only series where we got serious problems with singularities was the monthly pound series. The fully parameterized model converged for none of the starting values which we tried. We, therefore, imposed the restriction of equal variances for this series.

In order to judge whether the estimated models are compatible with the stylized facts of the data, we computed the implied skewness $\tilde{\beta}_1$ and the implied kurtosis $\tilde{\beta}_2$ of the models from

(8) $$\tilde{\beta}_1 = \frac{\sum_{i=1}^{I} p_i(3\sigma_i^2\delta_i + \delta_i^3)}{\left[\sum_{i=1}^{I} p_i(\sigma_i^2 + \delta_i^2)\right]^{3/2}}$$

and

$$\tilde{\beta}_2 = \frac{\sum_{i=1}^{I} p_i(3\sigma_i^4 + 6\sigma_i^2\delta_i^2 + \delta_i^4)}{\left[\sum_{i=1}^{I} p_i(\sigma_i^2 + \delta_i^2)\right]^2} \tag{9}$$

with $\delta_i = \mu_i - \mu$, where μ is the overall mean. The results, reported in Table 2, show a rather close agreement between the pattern of skewness and kurtosis in the data and the implied skewness and kurtosis. If we impose a 5 percent significance level, we find significant negative skewness in several daily and weekly series (see Table 1) and we find also for all those series a negative implied skewness. As regards implied kurtosis, we get leptokurtosis for all daily, weekly, and monthly series but the degree of leptokurtosis decreases under time-aggregation, as it does in the data (see Table 1). The implied kurtosis is somewhat smaller than the kurtosis of the data for daily and weekly series but there is a quite close agreement between implied leptokurtosis and actual leptokurtosis for the monthly data.

We also report in Table 2 the results from a likelihood-ratio (LR) test against the H_0 of Gaussian white noise. There is, however, a problem with the application of the LR test to mixture models since the degrees of freedom are unclear for mean-scale mixtures. We may either impose the restriction $p \equiv p_1 = 0$ (or alternatively: $p = 1$) which reduces the mixture model to Gaussian white noise or we may impose the restrictions $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$. In the first case we would have one degree of freedom (from one restriction) in the LR test and in the second case we would have two degrees of freedom. Another, related, problem with the LR test is that p is on the boundary of the parameter space under H_0 and that, therefore, the regularity conditions for the application of the χ^2 distribution are not satisfied.

We took here the pragmatic position of being on the save side with a conservative rule that the degrees of freedom are two. The choice of degrees of freedom is immaterial for daily and weekly data where the H_0 of Gaussian white noise is overwhelmingly rejected. For monthly and quarterly data, we find much weaker evidence against normality but this is no surprise in view of the stylized facts reported in the previous section.

Finally, we applied two tests of serial dependence to examine the dynamic properties of the model. In order to test for independence we introduce the unobservable state variable s_t which determines at time t from which component a realization is drawn, i.e. if $s_t = i$ then there will be a drawing from component i at t. From Bayes's theorem we get

$$p(s_t \mid x_t) = \frac{f_i(x_t \mid \mu_i, \sigma_i) \cdot p_i(s_t)}{\sum_{i=1}^{I} f_i(x_t \mid \mu_i, \sigma_i) \cdot p_i(s_t)} \tag{10}$$

and we can estimate the integer s_t by maximizing $f_i(x_t \mid \mu_i, \sigma_i) \cdot p_i(s_t)$. This gives an estimated series of states which can be tested for independence.

The first test is the multiple runs test of Barton and David (1957). A run is defined as a sequence of s_t's with the same value. The corresponding test statistic has an asymptotic standard normal distribution. A positive value of the test statistic indicates that there are less runs than expected under the H_0 of randomness. Table 2 shows that for all high-frequency series randomness is indeed rejected and that there are always less runs than expected, i.e. there is positive dependence in the states. Given that the components of the mixture are mainly different with respect to their variances, this result corresponds to the results from the ACF for squared data in Table 1. Again, there is a marked difference to low-frequency (i.e. monthly and quarterly) data where the runs test leads to a rejection of independence for the quarterly pound and yen series only[5].

The second test for independence is a conventional χ^2 test within the framework of a Markov chain for s_t. The results are very similar to the ones from the runs test. There is strong rejection of independence for all daily and weekly data but no rejection in the monthly data. For quarterly data, independence can be rejected for the pound and the yen at the 5 percent level. In all cases of rejections, the dependence is caused by the fact that the states s_t have a greater degree of persistence than expected, i.e. there is positive serial dependence of states.

We may conclude from the above analysis that the model of two mixtures of normal distributions captures well the stylized facts of non-normality and leptokurtosis in short-run exchange-rate data. With the exception of the yen series, the dominant effect is the scale effect and not the mean effect. For monthly and quarterly data, however, we find much weaker evidence for a mixture model. This, of course, is in accordance with the stylized fact of convergence to normality under time aggregation. But a general deficiency of the model is that it cannot capture the heteroskedasticity of the high-frequency data. The rejection in the test of independence for the state variable s_t indicates this deficiency. In the next section we will discuss an extension of this model which removes this deficiency by introducing a Markov chain process for the state variable.

5 The run test could not be computed for the monthly franc series because, according to (9), all observations are classified as belonging to the second component.

4. Markov-Switching Models

A natural way to add dynamics to the mixture model is to assume that the state variable s_t follows a time-homogeneous or stationary first-order Markov process, i.e.

(11) $$p(s_t = i \mid s_{t-1} = j, s_{t-2}, \ldots, s_1, x_t, x_{t-1}, \ldots, x_1) = p(s_t = i \mid s_{t-1} = j) = p_{ij}.$$

For a Markov-switching model with two states (or regimes) we get a 2x2 transition matrix of states with two independent probabilities p_{11} and p_{22}. Of course, it follows $p_{21} = 1 - p_{11}$ and that $p_{12} = 1 - p_{22}$. This Markov chain together with the mixture model (1) and the specification of the normal distribution (3) gives a seven parameter model with parameter vector $\Theta = (p_{11}, p_{22}, \mu_1, \mu_2, \sigma_1, \sigma_2, \psi)$ where $\psi = p(s_1 = 1)$. We need ψ, the probability of being in state 1 in time period $t = 1$, to start off the Markov chain and a natural choice is to set ψ equal to the stationary probability of being in state 1, i.e.

(12) $$\psi = \frac{1 - p_{22}}{2 - p_{11} - p_{22}}.$$

Estimation of the Markov-switching model is quite involved since the state variable s_t is not observable. The basic idea of the model is due to Baum et al. (1970) who suggested to estimate the model with the expectation-maximization (EM) algorithm. Furthermore, they derived the essential properties of the EM algorithm within a general model with Markov-chain dependence. They showed that, under certain regularity conditions, the EM algorithm increases the likelihood function monotonically and that it converges to the maximum-likelihood (ML) estimates. Lindgren (1978) detailed the steps needed to implement the EM algorithm for the Markov-switching model, extended the model to the case of switching regressions and examined the properties of the ML estimator. The switching-regressions model may be written as

(13) $$x_t = z_t \theta^i + u_t^i \qquad \text{with} \qquad u_t^i \sim N(0, \sigma_i^2)$$

(14) $$s_{t-d} = i \qquad \text{for} \qquad i = 1, \ldots, I$$

where z_t is a vector of exogenous or lagged endogenous variables, θ^i is a vector of regression parameters for the i –th state, u_t is a Gaussian white-noise disturbance term with state-dependent variance σ_i^2 and d is a delay parameter for the state variable s_t.

It may be noted that the switching-regressions model of (12) - (13) describes a broad class of models including the switching model of Goldfeld and Quandt (1973) and the threshold models of Tong (1983, 1990) and Priestley (1988). The specifications differ only in their assumptions about the state variable s_t. The Markov-switching model obtains from (12) - (13) when $z_t = 1$ and when we assume a Markov chain for s_t as in (11). We also assume that the delay parameter d is zero. Our motivation for ignoring mean effects of exogenous and lagged endogenous variables derives from Section 2 where we found no strong mean effects in the statistical properties of the exchange-rate data.

The estimates of the Markov-switching model are presented in Table 3 where asymptotic standard errors are given in brackets. First, we note that the large values of p_{11} and p_{22} indicate great persistence of states. We can calculate the expected duration of state $i, \lambda_i,$ from $\lambda_i = (1 - p_{ii})^{-1}$ and find that, for instance, the expected duration of state 2 is 50 days for the daily mark series. For daily data, the expected duration of states varies between 29.4 days and 50.0 days, for weekly data λ_i varies between 9.7 weeks and 31.3 weeks, for monthly data λ_i varies between 2.3 months and 12.0 months, and for quarterly data λ_i varies between 1 quarter and 14.3 quarters. The extreme values for the quarterly data were caused by the franc series where the problem of convergence to singularities occurred. For some series and states, the λ_i's are roughly consistent across time horizons. For instance, the λ_1's for the mark-dollar exchange rate are 47.6 for daily data, 9.7 for weekly data and 2.9 for monthly data. On the other hand, λ_2 is equal to 16.4 for the weekly yen series and equal to 16.7 for the monthly yen series, and this appears to be inconsistent. In general, the λ_i's of quarterly data are surprisingly high.

We also computed the stationary state probabilies $\bar{p}_1$ and $\bar{p}_2$ from $\bar{p}_1 = \psi$ (see (12)) and $\bar{p}_2 = 1 - \bar{p}_1$. It is surprising that the low-variance states, in general this is state 1, have in most eases a smaller stationary probability than the high-variance states. Only for two series is $\bar{p}_1$ greater than 0.5 (it is 0.526 for the daily pound series and 0.558 for the daily franc series, see also the last column in Table 4). This result is in conflict with the estimates from the mixture model where p_1 was greater than 0.5 for nearly all of the daily and weekly series.

From the parameter estimates of the model we can also compute the conditional probability $\pi_i(t \mid \tau) = p(s_t = i \mid x_1, \ldots, x_\tau; \theta)$ in a recursive way (see Lindgren (1978) and Hamilton (1989)). If we set $\tau = t$, we get "filter" probabilities about the probable state at time t whereas for $\tau = T$ we get "smoothed" probabilities based on the full sample. In practice, both alternatives give very similar results. In Figure 1 we plot the smoothed probabilities of state 1 for the weekly mark series together with the series x_t. It is apparent from this figure that state 1 is associated with tranquil (i.e. low variance) periods. The probability smoother identifies the period from November 1975 until November 1977 as being associated with state 1 if the criterion is that $\pi_1(t \mid T) > 0.5$. The corresponding plot of x_t shows that this was also a period of relatively small weekly exchange-rate fluctuations. The only other periods of prolonged tranquillity are the ones from July until December 1974 and from February until September 1979. This seems to indicate that the early period of the post-Bretton-Woods era was more tranquil than the more recent one. It is also interesting to note that the smoother attaches a probability of zero to $\pi_1(t \mid T)$ to both the strongest appreciation of the mark in the sample (the 7.8 percent appreciation in the wake of the Plaza agreement in September 1985) and the stronges depreciation (the 7.0 percent depreciation after the introduction of support measures for the dollar in November 1978).

Figure 1 a
Smoothed probabilities of state 1: weekly mark series

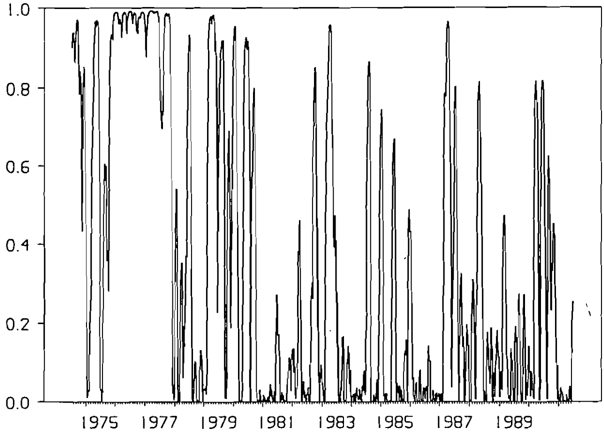
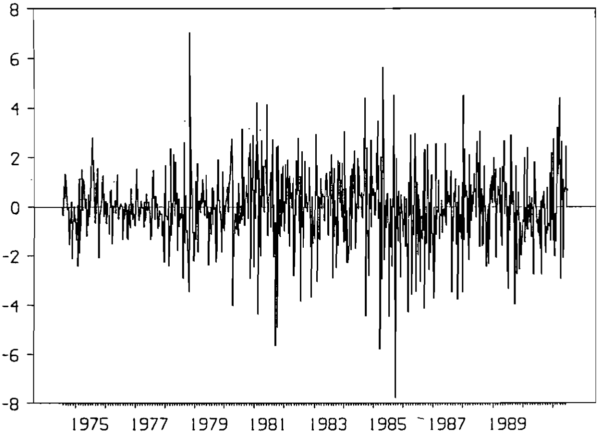


Figure 1 b
Exchange-rate dynamics: weekly mark series



As regards the estimates of means μ_i, we get results that are very similar to the ones of the mixture model. We find no strong mean effects in daily, weekly and monthly data with the exception of the daily yen series. It is only with quarterly data that we find strong mean effects with μ_1 and μ_2 having opposite signs. This result confirms the finding of Engel and Hamilton (1990) that there are significant mean effects in the three quarterly dollar exchange rate they analysed (mark, pound and French franc). It is, however, somewhat puzzling that we should find mean effects in quarterly data but not in higher-frequency data. A possible explanation for this might be significant high-order autocorrelations in the high-frequency data, a phenomenon that was also discussed in the context of business-cycle analysis (see Cochrame (1988)). Applications of the variance-ratio test to exchange-rate data seem to confirm this conjecture (see Lin and He (1991)).

As regards the estimates of variances σ_i, we find the variance-effect to be dominant in daily, weekly and monthly data but also to be prevalent in quarterly data. We noted in the previous section that the presence of singularities in the likelihood is a major problem for the estimation of mean-scale mixtures. The same problem arises in the estimation of Markov-switching models with mean-and-variance effects. But, fortunately, this problem did not bother us much in our application to exchange rates. The only series where this problem occured was the quarterly franc series.

Finally, it is interesting to compare the Markov-switching models and the mixture models with LR tests. Table 3 shows that for daily and weekly data we can reject the mixture models in favour of the corresponding Markov-switching models at very high significance levels (note that we apply a χ^2 distribution with one degree of freedom

Table 3
Estimates of the Markov-switching model

		mark	pound	franc	yen
day	p_{11}	0.979 (0.004)	0.975 (0.005)	0.977 (0.004)	0.966 (0.006)
	p_{22}	0.980 (0.005)	0.973 (0.005)	0.971 (0.006)	0.979 (0.004)
	μ_1	0.002 (0.010)	0.010 (0.009)	0.013 (0.011)	0.018 (0.008)
	μ_2	-0.018 (0.020)	0.008 (0.020)	-0.051 (0.026)	-0.039 (0.016)
	σ_1^2	0.153 (0.008)	0.147 (0.014)	0.225 (0.010)	0.072 (0.056)
	σ_2^2	0.781 (0.033)	0.793 (0.045)	1.204 (0.050)	0.610 (0.021)
	LR	513.490***	389.732***	465.268***	651.809***
week	p_{11}	0.897 (0.029)	0.897 (0.020)	0.924 (0.041)	0.922 (0.025)
	p_{22}	0.940 (0.021)	0.968 (0.004)	0.967 (0.036)	0.939 (0.028)
	μ_1	-0.062 (0.051)	-0.086 (0.062)	-0.116 (0.080)	0.087 (0.063)
	μ_2	-0.025 (0.078)	0.082 (0.059)	-0.055 (0.082)	-0.217 (0.113)
	σ_1^2	0.416 (0.066)	0.284 (0.070)	0.597 (0.185)	0.388 (0.092)
	σ_2^2	3.154 (0.240)	2.679 (0.140)	3.867 (0.560)	2.851 (0.313)
	LR	66.692***	64.356***	55.204***	78.094***
month	p_{11}	0.657 (0.369)	0.568 (0.105)	0.767 (0.075)	0.857 (0.023)
	p_{22}	0.840 (0.307)	0.908 (0.043)	0.917 (0.013)	0.940 (0.012)
	μ_1	-0.179 (0.449)	0.654 (0.457)	0.156 (0.477)	0.026 (0.210)
	μ_2	-0.174 (0.412)	0.089 (0.398)	-0.488 (0.305)	-0.541 (0.334)
	σ_1^2	2.967 (4.456)	2.108 (1.426)	3.347 (1.007)	1.568 (0.483)
	σ_2^2	15.820 (6.103)	13.171 (1.561)	18.582 (1.839)	15.116 (1.914)
	LR	0.123	0.246	5.585**	7.548***
quarter	p_{11}	0.814 (0.120)	0.807 (0.070)	0.930 (0.036)	0.647 (0.068)
	p_{22}	0.881 (0.099)	0.606 (0.216)	0.000 (0.541)	0.752 (0.148)
	μ_1	-4.115 (1.074)	-2.554 (1.044)	0.169 (0.808)	2.269 (0.706)
	μ_2	1.762 (1.651)	6.955 (1.684)	-16.696 (0.756)	-3.535 (1.437)
	σ_1^2	17.122 (6.724)	14.755 (4.570)	37.260 (7.445)	12.696 (4.288)
	σ_2^2	40.604 (10.664)	8.122 (6.641)	2.058 (1.548)	40.066 (7.082)
	LR	0.483	5.370**	0.526	3.310

to the LR statistic). However, the evidence in favour of the Markov-switching model is much weaker in the low-frequency data. Only three of the eight LR's for monthly and quarterly data are significant at the 5 percent level. It is only the monthly yen series where the LR statistic rejects the mixture model against the Markov-switching model *and* it rejects the H_0 of Gaussian white-noise against the mixture model. If we apply the LR test to a direct comparison between the Markov-switching model and Gaussian white noise, we are again confronted with the methodological problems mentioned in the last section. However, if we apply a χ^2 distribution with 2 degrees of freedom to this direct comparison we find that Gaussian white noise is rejected in six of the eight low-frequency series.

We motivated the application of mixture models to the modelling of exchange-rate dynamics in Section 3 with reference to the stylized facts of the data, i.e. with the statistical properties of leptokurtosis and convergence to normality under time-aggregation. The extension to the Markov-switching model was motivated by the fact that the mixture model cannot capture the stylized fact of heteroscedasticity whereas this properties is incorporated in Markov-switching models with scale components. We have now to ask ourselves whether leptokurtosis and convergence to normality still obtain in a Markov-switching model. The question of leptokurtosis is very easy to answer. In order to compute the moments of the distribution of X, we simply have to compute the stationary probabilities as in (12) and then proceed as in the case of a mixture model. This means that we have the same condition for leptokurtosis as in the mixture model with p_i replaced by $\overline{p}_i$ and that we may use equations (8) and (9) with $\overline{p}_i$ substituted for p_i. It follows that Markov-switching models with variance effects but without mean effects always imply leptokurtosis.

The question of convergence to normality under time-aggregation is more difficult to address since the property of independence is lost and, therefore, we cannot invoke a simple central limit theorem. There is, however, some reason to conjecture that convergence to normality obtains for non-degenerate and non-pathological Markov-switching models. Lindgren (1978) established the asymptotic independence of the X_t variables and this ought to be half the way to the proof of convergence to normality[6]. We leave it to future research to provide the remaining steps of a complete proof of convergence since we want to concentrate here on the aspects of application of the model. However, in Table 4 we offer some illustrations of asymptotic independence for the estimated models. We report there the estimates of the n-step transition probabilities $p_{ii}^n = p(s_t = i \mid s_{t-n} = i)$ for daily and weekly data. The n-step transition probabilities are obtained from the n-th power of the transition matrix. The last column

6 Lindgren (1978) proved asymptotic independence by showing that the "mixing" conditions are satisfied. The use of the term "mixing" might cause some confusion in this context since the "mixing" conditions are not related in any way to the mixing of densities as in (1).

reports the stationary transition probabilities for $n \to \infty$. Independence obtaines if $p_{11}^n + p_{22}^n = 1$, and we find this condition quite closely satisfied at a yearly time interval, i.e. for $n = 250$ with daily data and for $n = 52$ with weekly data.

Table 4
Estimates of the n-step transition probabilities

day		$n=5$	$n=20$	$n=60$	$n=250$	$n \to \infty$
mark	p_{11}^n	0.903	0.708	0.530	0.490	0.490
	p_{22}^n	0.906	0.719	0.548	0.510	0.510
pound	p_{11}^n	0.888	0.688	0.545	0.526	0.526
	p_{22}^n	0.876	0.654	0.495	0.474	0.474
franc	p_{11}^n	0.896	0.709	0.576	0.558	0.558
	p_{22}^n	0.869	0.632	0.464	0.442	0.442
yen	p_{11}^n	0.847	0.579	0.401	0.381	0.381
	p_{22}^n	0.907	0.741	0.632	0.619	0.619
week		$n=1$	$n=4$	$n=12$	$n=52$	$n \to \infty$
mark	p_{11}^n	0.897	0.678	0.444	0.369	0.369
	p_{22}^n	0.940	0.812	0.674	0.631	0.631
pound	p_{11}^n	0.897	0.663	0.369	0.235	0.234
	p_{22}^n	0.968	0.897	0.807	0.766	0.766
franc	p_{11}^n	0.924	0.696	0.479	0.305	0.303
	p_{22}^n	0.967	0.868	0.773	0.698	0.697
yen	p_{11}^n	0.922	0.704	0.532	0.439	0.439
	p_{22}^n	0.939	0.769	0.634	0.562	0.562

We may conclude from the analysis in this section that the Markov-switching model captures well the major stylized facts of the exchange-rate data. It does so especially for the short-run, i.e. daily and weekly, data.

5. Implications for the Pricing of Foreign-currency Options

The concept of choice under uncertainty is the cornerstone of financial theory. Therefore, the stochastic specification of financial models is of fundamental importance in almost any branch of modern finance. A convenient and natural choice for the underlying probability model is the normal distribution in the static context and the corresponding Wiener process in the continuous-time context.

The assumption of a Wiener process for the price (or return) process is also central for the seminal option-pricing model of Black and Scholes (1973). The fact that this assumption is at odds with the empirical regularities of stock prices has early been recognized (see the Introduction). But early attempts to adapt the Black-Scholes model to the stylized facts of financial data have only considered the effect of non-normality, i.e. leptokurtosis. A popular alternative to the Wiener process has been Merton's (1976) model of a jump-diffusion process which incorporates a compound Poisson process. Empirical applications have shown that this model provides a better fit to the data than the assumption of Gaussian white noise but when option prices are computed from the estimated parameters of compound Poisson processes, they differ only little from Black-Scholes prices (see Ball and Torons (1985)).

The issue of heteroskedasticity has only recently been addressed in the finance literature under the label of "stochastic volatility" (Jarrow and Wiggins (1989) and Taylor (1992) provide surveys of this literature). The approaches which have been applied can be grouped under two headings: the continuous-time-finance approach and the econometric approach. In the continuous-time-finance approach, the price process

$$dE/E = \alpha dt + \sigma dW \tag{15}$$

(where E the price of the underlying asset, say the exchange rate, W is standard Brownian motion, α is a constant and σ is the instantaneous standard deviation) is augmented by a specification of the volatility process as a geometric Wiener process

$$d\sigma/\sigma = \lambda dt + \gamma dV \tag{16}$$

(where λ and γ are constants and V is standard Brownian motion) or as a Ornstein-Uhlenbeck process

$$d\sigma/\sigma = \lambda(\xi - \sigma)dt + \gamma dV \tag{17}$$

(where ξ is a constant) or some variants of (16) or (17). The two-equation system of either (15)-(16) or (15) and (17) has as an additional parameter ϱ, the correlation between dW and dV.

There are several problems with this approach. First, the specification of (16) or (17) is ad hoc and only motivated by the fact that it is convenient to work with popular stochastic processes. Second, the fundamental problem with any specification of an independent stochastic volatility process is that it becomes impossible to construct a perfect-hedge portfolio because volatility is non-observable and non-traded. Therefore, the great advantage of risk-neutral evaluation is lost. It has been tried to circumvent this problem by either assuming that the volatility risk can be diversified (see e.g. Hull and White (1987)), but this appears to be arbitrary, or by putting restrictions on the utility function of investors, such as logarithmic utility functions (see e.g. Wiggins (1987)).

Whereas the continuous-time-finance approach starts from a theoretical perspective, the econometric approach starts from an empirical one. This approach has only recently been applied and it has used the generalized autoregressive conditional variance (GARCH) model of Engle (1982) and Bollerslev (1986). The aim of this approach is to find a specification of the volatility process which adequately represents the stylized facts (see e.g. Duan (1991)). A problem with this approach is that it is often unclear under which conditions the specified volatility process is compatible with the risk-neutral valuation principle. Duan (1991), however, has established such conditions for the GARCH model.

In this section, we follow the econometric approach to study the impact of leptokurtosis and heteroskedasticity on option pricing. More specifically, we compute call option prices which would obtain under a mixture model and under a Markov-switching model and compare them with Black-Scholes prices which are derived under the assumption of Gaussian white noise. Of course, we cannot hope to derive closed-form solutions for option prices of mixture models and Markov-switching model. We, therefore, have to rely on simulations which are based on the expected value of the boundary condition, i.e. we compute the call option price as

$$C = \frac{1}{R}\sum_{r=1}^{R} \max\{E_r - B; 0\} \tag{18}$$

where B is the exercise price and R = 20,000 is the numer of repetitions in every experiment. We based our simulations on the parameter estimates of the daily mark series, as reported in Tables 1-3.

Table 5 reports the results from the simulation experiments when the current spot rate E_t is varied between 1.60 and 2.00. The time to maturity is set to 20 days and the exercise price B is set to 1.80. Note that, across each row of Table 5, the computed option prices are based on the same realizations of the random variable, whereas the drawings are distinct between rows. We could, of course, compute the Black-Scholes prices from a closed-form equation but in order to reduce the impact of sample variation,

Table 5
Spot-rate effect for call options: daily mark series

E_t	Gauss	Mixture	Markov	bias $_{Mixture}$	st. error	bias $_{Markov}$	st. error
1.60	0.000000	0.000003	0.000013	0.000002	0.000275	0.000012	0.000638
1.61	0.000002	0.000002	0.000014	0.000000	0.000197	0.000013	0.000742
1.62	0.000006	0.000011	0.000030	0.000005	0.000411	0.000024	0.001022
1.63	0.000012	0.000022	0.000050	0.000010	0.000699	0.000038	0.001352
1.64	0.000010	0.000010	0.000071	-0.000000	0.000497	0.000061	0.001489
1.65	0.000031	0.000048	0.000121	0.000016	0.001198	0.000089	0.002033
1.66	0.000069	0.000092	0.000210	0.000024	0.001432	0.000142	0.002545
1.67	0.000134	0.000145	0.000308	0.000011	0.001684	0.000174	0.003025
1.68	0.000192	0.000218	0.000424	0.000026	0.002066	0.000233	0.003373
1.69	0.000368	0.000371	0.000569	0.000004	0.003007	0.000201	0.004345
1.70	0.000621	0.000639	0.000924	0.000018	0.003721	0.000303	0.005199
1.71	0.000992	0.000984	0.001366	-0.000008	0.004531	0.000374	0.006067
1.72	0.001398	0.001429	0.001717	0.000031	0.005216	0.000320	0.006803
1.73	0.002262	0.002330	0.002401	0.000068	0.006551	0.000138	0.007958
1.74	0.003329	0.003320	0.003422	-0.000009	0.007404	0.000093	0.008953
1.75	0.005169	0.005193	0.005097	0.000024	0.008793	-0.000072	0.010338
1.76	0.007003	0.007015	0.006707	0.000012	0.009833	-0.000296	0.011114
1.77	0.009605	0.009475	0.008915	-0.000130	0.011165	-0.000690	0.011801
1.78	0.012810	0.012674	0.011814	-0.000136	0.012144	-0.000995	0.012406
1.79	0.016543	0.016581	0.015555	0.000038	0.013621	-0.000988	0.012772
1.80	0.020855	0.020841	0.019793	-0.000014	0.014429	-0.001062	0.012882
1.81	0.026378	0.026222	0.025422	-0.000157	0.015999	-0.000956	0.013759
1.82	0.032268	0.032425	0.031619	0.000157	0.017135	-0.000649	0.014006
1.83	0.039193	0.038843	0.038593	-0.000350	0.017967	-0.000600	0.014581
1.84	0.045956	0.045926	0.045930	-0.000030	0.018704	-0.000027	0.015244
1.85	0.053905	0.053924	0.053992	0.000019	0.019593	0.000087	0.016332
1.86	0.062576	0.062522	0.062732	-0.000055	0.020449	0.000156	0.016879
1.87	0.070819	0.070960	0.071402	0.000140	0.020758	0.000583	0.017893
1.88	0.079893	0.080074	0.080399	0.000181	0.021302	0.000506	0.018473
1.89	0.089578	0.089608	0.090176	0.000029	0.021820	0.000598	0.019190
1.90	0.098772	0.099140	0.099507	0.000368	0.022054	0.000735	0.019760
1.91	0.108399	0.108739	0.108748	0.000340	0.022573	0.000350	0.020183
1.92	0.117965	0.117986	0.118355	0.000021	0.022773	0.000390	0.020639
1.93	0.128322	0.128344	0.128543	0.000022	0.023397	0.000221	0.021202
1.94	0.138264	0.138107	0.138489	-0.000157	0.023194	0.000225	0.021236
1.95	0.148130	0.148112	0.148246	-0.000017	0.023193	0.000116	0.021579
1.96	0.157830	0.157835	0.158127	0.000005	0.023524	0.000297	0.021620
1.97	0.167222	0.167193	0.167415	-0.000028	0.023608	0.000194	0.022045
1.98	0.178347	0.178564	0.178557	0.000217	0.023786	0.000210	0.022113
1.99	0.187682	0.187663	0.187869	-0.000019	0.023971	0.000186	0.022212
2.00	0.198572	0.198372	0.198444	-0.000200	0.024034	-0.000128	0.022345

the Black-Scholes prices, reported under the heading of "Gauss", are based on simulations, too. Note, too, that we neglect the present-value factor in the boundary condition (18) since it would have no influence on the comparison of prices between models.

Although the simulations are based on 20,000 repetitions per row, the sample variations are still sizeable for out-of-the-money options. It is, of course, inconsistent to have call option prices for the Black-Scholes model and for the mixture model which are lower at a spot rate of $E_t = 1.64$ than at $E_t = 1.63$ but here we are interested in comparisons across rows and not between rows. Under the headings of "bias" we report the differences between Black-Scholes prices and prices from the mixture model and the Markov-switching model, respectively. The corresponding standard errors are given in columns 7 and 9. Table 5 shows that the biases according to the mixture model are small and unsystematic. The standard error of biases is a multiple of the biases for all spot rates. We find, however, a systematic pattern in biases if Black-Scholes prices are compared with prices computed from the Markov-switching model. For out-of-the-money options the bias is positive and increasing if we go from $E_t = 1.60$ to $E_t = 1.71$. It then decreases and becomes negative for at-the-money options. The bias is again positive for in-the-money options with spot rates larger than 1.84. The largest bias obtains for a spot rate of 1.90 and somewhat surprisingly, the bias is again negative for a spot rate of 2.00. It is interesting to note that this pattern of biases mimics the patterns derived by Hull and White (1987) within the continuous-time-finance approach based on equations (15)-(16) and also mimics the results of Duan (1991) who found the same pattern of biases in an application of the GARCH model. Although the pattern is systematic, the biases are small and insignificant when they are compared with their standard errors. It is, therefore, doubtful whether any profitable investment strategy can be based on these biases.

Results from experiments of varying the time to maturity are reported in Table 6. We computed call option prices for at-the-money options with a spot rate and an exercise price of 1.80. The simulations were based on the same parameter estimates as in the previous experiment and the time to maturity was varied between 1 and 40 days.

A comparison of option prices derived under the assumptions of Gaussian white noise and of a mixture distribution shows that for nearly all maturities, the Black-Scholes prices are larger than the mixture-distribution prices. This corresponds to the negative value of the bias obtained in Table 5 for a spot rate of 1.80. Surprisingly, however, for a maturity of 20 days we get a positive bias in Table 6. This indicates that there is sizeable sample variation for at-the-money options although we used 20,000 repetitions. The absolute value of the bias is only a small fraction of its standard error for all maturities. The biases are, therefore, insignificant.

Table 6
Maturity effect for call options: daily mark series

Maturity	Gauss	Mixture	Markov	bias $_{Mixture}$	st. error	bias $_{Markov}$	st. error
1	0.00490	0.00448	0.00458	-0.00042	0.00326	-0.00032	0.00309
2	0.00693	0.00657	0.00648	-0.00036	0.00460	-0.00045	0.00436
3	0.00831	0.00812	0.00777	-0.00019	0.00569	-0.00054	0.00528
4	0.00957	0.00934	0.00901	-0.00024	0.00643	-0.00056	0.00597
5	0.01079	0.01061	0.01007	-0.00018	0.00731	-0.00072	0.00680
6	0.01166	0.01150	0.01086	-0.00016	0.00804	-0.00080	0.00741
7	0.01302	0.01279	0.01231	-0.00023	0.00874	-0.00071	0.00814
8	0.01377	0.01365	0.01291	-0.00012	0.00928	-0.00086	0.00849
9	0.01432	0.01410	0.01360	-0.00022	0.00978	-0.00072	0.00889
10	0.01499	0.01479	0.01398	-0.00020	0.01035	-0.00101	0.00940
11	0.01584	0.01560	0.01487	-0.00024	0.01077	-0.00097	0.00985
12	0.01642	0.01616	0.01546	-0.00026	0.01135	-0.00096	0.01024
13	0.01719	0.01711	0.01613	-0.00008	0.01189	-0.00106	0.01070
14	0.01795	0.01786	0.01695	-0.00008	0.01209	-0.00100	0.01123
15	0.01834	0.01823	0.01737	-0.00010	0.01284	-0.00097	0.01129
16	0.01913	0.01898	0.01809	-0.00015	0.01316	-0.00104	0.01184
17	0.01963	0.01936	0.01861	-0.00027	0.01342	-0.00102	0.01220
18	0.01981	0.01980	0.01870	-0.00000	0.01411	-0.00111	0.01242
19	0.02030	0.02003	0.01924	-0.00027	0.01426	-0.00106	0.01261
20	0.02100	0.02112	0.02010	0.00012	0.01494	-0.00090	0.01307
21	0.02138	0.02127	0.02045	-0.00010	0.01523	-0.00093	0.01321
22	0.02222	0.02217	0.02103	-0.00005	0.01550	-0.00119	0.01378
23	0.02275	0.02271	0.02161	-0.00004	0.01592	-0.00114	0.01419
24	0.02302	0.02296	0.02183	-0.00006	0.01609	-0.00119	0.01438
25	0.02356	0.02347	0.02252	-0.00008	0.01641	-0.00104	0.01470
26	0.02421	0.02402	0.02308	-0.00019	0.01690	-0.00113	0.01487
27	0.02451	0.02444	0.02341	-0.00007	0.01713	-0.00110	0.01503
28	0.02473	0.02467	0.02359	-0.00006	0.01710	-0.00114	0.01535
29	0.02546	0.02549	0.02409	0.00003	0.01769	-0.00137	0.01583
30	0.02559	0.02565	0.02431	0.00006	0.01803	-0.00128	0.01607
31	0.02578	0.02569	0.02463	-0.00010	0.01826	-0.00115	0.01602
32	0.02612	0.02592	0.02489	-0.00020	0.01843	-0.00123	0.01654
33	0.02700	0.02692	0.02575	-0.00009	0.01904	-0.00125	0.01679
34	0.02682	0.02669	0.02571	-0.00013	0.01903	-0.00111	0.01684
35	0.02761	0.02763	0.02614	0.00002	0.01961	-0.00147	0.01728
36	0.02808	0.02791	0.02686	-0.00017	0.01966	-0.00122	0.01748
37	0.02783	0.02783	0.02662	-0.00000	0.01980	-0.00121	0.01754
38	0.02847	0.02849	0.02709	0.00003	0.02029	-0.00138	0.01816
39	0.02852	0.02843	0.02751	-0.00009	0.02039	-0.00101	0.01777
40	0.02893	0.02905	0.02769	0.00012	0.02058	-0.00124	0.01846

If we compare Black-Scholes prices with Markov-switching prices, we find in Table 6 that the Black-Scholes model overprices the options at all maturities and this corresponds to the result in Table 5 for a spot rate of 1.80. The absolute value of the bias increases with the time to maturity but, again, the sample variation appears to be quite large. Although the bias is systematic, it is not significant in a statistical sense since the standard error of the bias is much larger than the amount of bias, in most cases the bias is less than 10 percent of its standard error.

We may conclude from the simulation experiments in this section that we find systematic differences to Black-Scholes prices if we adopt a Markov-switching model but not if we adopt a mixture model. Statistically, however, all biases are insignificant.

6. Conclusions

In this paper we have examined issues in the application of mixture models and Markov-switching models to the modelling of price dynamics in financial markets. We applied the models to exchange-rate data but the approach is readily extended to other financial prices, such as stock prices, since speculative prices share the stylized facts of leptokurtosis and heteroskedasticity.

Engel and Hamilton (1990) motivated their application of the Markov-switching model to quarterly exchange rates with a search for "long swings" in exchange rates, i.e. with a search for mean effects. Our motivation differs from it by emphazising that the most significant statistical properties of exchange-rate data are the leptokurtosis and heteroskedasticity of short-run, i.e. daily and weekly, data and by relating it to mixture models which have a long tradition in finance.

The estimation showed that Markov-switching models provide a significantly better fit than models of Gaussian white noise and mixture models and that this holds especially for short-run data. To a certain extent we can confirm the findings of mean effects in quarterly data, as in Engel and Hamilton (1990), but the dominant effect is the variance effect in daily and weekly data where we find no significant mean effects in seven of eight series.

Although we find highly significant deviations from Gaussian white noise in high-frequency data and although we were able to fit models which capture these deviations well, we find that the implications of both the mixture model and of the Markov-switching model for the pricing of call options on foreign currencies are minor. We only find systematic differences to Black-Scholes prices if we adopt the Markov-switching model but these differences are not statistically significant (and probably also not economically).

There are several points where our analysis is incomplete or where an extension would be interesting. We noted in Section 4 that currently we can only conjecture that the Markov-switching model is compatible with convergence to normality under time aggregation. We hope to provide a proof of this property in a subsequent paper. The greates gap that our paper leaves is the fact that we only assume that option pricing in the framework of the risk-neutral-valuation principle is possible if we adopt a mixture model or a Markov-switching model. We leave it to future work to show under which conditions these models allow risk-neutral valuation and we suppose that we have to impose some restricitons on the utility function or the changes in aggregate consumption in order to derive the desired result. Finally, it would be interesting to compare the performance of the Markov-switching model in modelling speculative prices with the performance of GARCH-type models.

References

Akgiray, V. and G. G. Booth (1987), "Compound Distribution Models of Stock Returns: An Empirical Comparison." *Journal of Financial Research*, 10, 269-280.

Akgiray, V. and G. G. Booth (1988), "Mixed Diffusion-Jump Process Modeling of Exchange Rate Movements". *Review of Economics and Statistics*, 70, 631-637.

Akgiray, V., G. G. Booth, and O. Loistl (1989), "Statistical Models of German Stock Returns". *Journal of Economics*, 50, 17-33.

Ball, C. A. and W. N. Torous (1983), "A Simplified Jump Process for Common Stock Returns". *Journal of Financial and Quantitative Analysis*, 18, 53-61.

Ball, C. A. and W. N. Torous (1985), "On Jumps in Common Stock Prices and Their Impact on Call Option Pricing". *Journal of Finance*, 40, 155-173.

Barnea, A. and D. H. Downes (1973), "A Reexamination of the Empirical Distribution of Stock Price Changes". *Journal of the American Statistical Association*, 68, 348-350.

Barton, D. E. and F. N. David (1957),"Multiple Runs". *Biometrika*, 44, 168-178.

Baum, L. E., T. Petrie, G. Soules, and N. Weiss (1970), "A Maximization Technique Occuring in the Statistical Analysis of Probabilistic Functions of Markov Chains". *Annals of Mathematical Statistics*, 41, 164-171.

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

Bollerslev, T. (1986), "Generalized Autoregressiv Conditional Heteroskedasticity". *Journal of Econometrics*, 31, 307-327.

Boothe, P. and D. Glassman (1987), "The Statistical Distribution of Exchange Rates". *Journal of International Economics*, 22, 297-319.

Cochrane, J. H. (1988), "How Big is the Random Walk in GNP?" *Journal of Political Economy*, 96, 893-920.

D'Agostino, R. B. (1986), "Tests for the Normal Distribution". In: R. B. D'Agostino and M. A. Stephens (eds.), *Goodness-of-Fit Techniques*. New York: M. Dekker, 367-419.

Diebold, F. X. (1988), *Empirical Modeling of Exchange Rate Dynamics*. Berlin: Springer.

Duan, J.-C. (1991), "The GARCH Option Pricing Model". Faculty of Management, Mc Gill University, Canada, mimeo.

Engel, C. M. and J. D. Hamilton (1990), "Long Swings in the Dollar: Are They in the Data and Do Markets Know It ?" *American Economic Review*, 80, 689-713.

Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation". *Econometrica*, 50, 987-1008.

Everitt, B. S. and D. J. Hand (1981), *Finite Mixture Distributions*. London: Chapman and Hall.

Fama, E. F. (1976), *Foundations of Finance*. New York: Basic Books.

Fielitz, B. D. and J. P. Rozelle (1983), "Stable Distributions and the Mixtures of Distributions Hypothesis for Common Stock Returns". *Journal of the American Statistical Association*, 78, 28-36.

Friedman, D. and S. Vandersteel (1982), "Short-Run Fluctuations in Foreign Exchange Rates". *Journal of International Economics*, 13, 171-186.

Goldfeld, S. M. and R. E. Quandt (1973), "A Markov Model for Switching Regressions". *Journal of Econometrics*, 1, 3-15.

Gridgeman, N.T. (1970), "A Comparison of Two Methods of Analysis of Mixtures of Normal Distributions". *Technometrics*, 12, 823-833.

Hamilton, J. D. (1988), "Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates". *Journal of Economic Dynamics and Control*, 12, 385-423.

Hamilton, J. D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle". *Econometrica*, 57, 357-384.

Hamilton, J. D. (1990), "Analysis of Time Series Subject to Changes in Regime". *Journal of Econometrics*, 45, 39-70.

Hamilton, J. D. (1991a), "A Quasi-Bayesian Approach to Estimating Parameters for Mixtures of Normal Distributions". *Journal of Business and Economic Statistics*, 9, 27-39.

Hamilton, J. D. (1991b), "Estimation, Inference, and Forecasting of Time Series Subject to Changes in Regime". To appear in C. R. Rao and G. S. Maddala (eds.), *Handbook of Statistics*, Vol. 10.

Hull, J. and Alan White (1987), "The Pricing of Options on Assets with Stochastic Volatilities". *Journal of Finance*, 42, 281-300.

Jansen, D. W. and C. G. de Vries (1991), "On the Frequency of Large Stock Returns: Putting Boosts and Busts into Perspective." *Review of Economics and Statistics*, 73, 18-24.

Jarrow, R. A. and J. B. Wiggins (1989), "Option Pricing and Implicit Volatilities". *Journal of Economic Surveys*, 3, 59-81.

Kaehler, J. (1991), "On the Modelling of Exchange-Rate Dynamics by Stable Paretian Distributions". Institut für Volkswirtschaftslehre und Statistik, Universität Mannheim, Discussion Paper No. 448-91.

Kon, S. J. (1984), "Models of Stock Returns - A Comparison". *Journal of Finance*, 39, 147-165.

Lau, A. H.-L., H.-S. Lau, and J. R. Wingender (1990), "The Distribution of Stock Returns: New Evidence Against the Stable Model". *Journal of Business and Economic Statistics*, 8, 217-223.

Liu, C. Y. and J. He (1991), "A Variance-Ratio Test of Random Walks in Foreign Exchange Rates". *Journal of Finance*, 46, 773-785.

Lindgren, G. (1978), "Markov Regime Models for Mixed Distributions and Switching Regressions". *Scandinavian Journal of Statistics*, 5, 81-91.

Mandelbrot, B. (1963), "The Variation of Certain Speculative Prices". *Journal of Business*, 36, 394-419.

Mc Leod, A.I. and W.K.Li (1983), "Diagnostic Checking ARIMA Time Series Models Using Squared-Residual Autocorrelations". *Journal of Time Series Analysis*, 4, 269-273.

Merton, R. C. (1976), "Option Pricing When Underlying Stock Returns are Discontinuous". *Journal of Financial Economics*, 3, 124-144.

Pearson, K. (1894), "Contributions to the Mathematical Theory of Evolution". *Philosophical Transactions of the Royal Society*, 185, 71-110.

Praetz, P. D. (1972), "The Distribution of Share Price Changes". *Journal of Business*, 45, 49-55.

Press, S. J. (1967), "A Compound Events Model for Security Prices". *Journal of Business*, 40, 317-335.

Priestley, M. B. (1988), *Non-linear and Non-stationary Time Series Analysis*. London: Academic Press.

Stephens, M. a. (1986), "Tests Based on EDF Statistics". In: R. B. D'Agostino and M. A. Stephens (eds.), *Goodness-of-Fit Techniques*. New York: M. Dekker, 97-193.

Taylor, S. J. (1986), *Modelling Financial Time Series*. Chichester: Wiley.

Taylor, S. J. (1992), "Modelling Stochastic Volatility". Department of Accounting and Finance, Lancaster University, UK, mimeo.

Titterington, D. M., A. F. M. Smith, and U. E. Makov (1985), *Statistical Analysis of Finite Mixture Distributions*. Chichester: Wiley.

Tong, H., (1983), *Threshold Models in Non-linear Time Series Analysis*. New York: Springer.

Tong, H., (1990), *Non-linear Time Series: A Dynamical Systems Approach*. Oxford: Oxford University Press.

Tucker, A. L. and L. Pond (1988), "The Probability Distribution of Foreign Exchange Prices Changes: Tests of Candidate Processes". *Review of Economics and Statistics*, 70, 638-647.

Wiggins, J. (1987), "Option Values Under Stochastic Volatility: Theory and Empirical Estimates". *Journal of Financial Economics*, 19, 351-372.