 Journal of Business & Economic Statistics, July 1993, Vol. 11, No. 3

The Message in Weekly Exchange Rates in the European Monetary System: Mean Reversion, Conditional Heteroscedasticity, and Jumps

Peter J. G. Vlaar and Franz C. Palm
Department of Economics, University of Limburg, 6200 MD Maastricht, The Netherlands

Weekly rates of the European Monetary System (EMS) vis-à-vis the Deutsche mark from April 1979 to March 1991 are modeled as a combined MA (1)–GARCH(1, 1)–jump process. The moving average (MA) part accounts for mean reversion required for the rates to stay inside the target zone. The generalized autoregressive conditional heteroscedasticity (GARCH) part accounts for changing volatility, whereas the jump process models parity changes and other erratic movements. Using an adjusted Pearson chi-squared goodness-of-fit test, we find similar results for the Bernoulli and the Poisson jump processes. In those cases in which the Bernoulli–normal distribution does not pass the goodness-of-fit test, a mixture of three normals does. Finally the MA(1)–GARCH(1, 1)–Bernoulli jump models are jointly estimated assuming a constant contemporaneous correlation matrix for the disturbances and a common jump probability for all the currencies.

KEY WORDS: Bernoulli jump process; GARCH; Mean reversion; Target zone for exchange rates.

In March 1979 the European Monetary System (EMS) was founded. A main element of the EMS is the Exchange Rate Mechanism (ERM) in which each currency has a central rate expressed in terms of the European Currency Unit. These central rates determine a grid of bilateral central rates, around which fluctuation margins are established. To keep the exchange rate within these margins, participating countries are obliged to intervene in the foreign-exchange market when a bilateral exchange rate reaches the boundary of this band. For this purpose, special credit facilities have been established. Instead of intervening, it is also possible to realign the parities, provided that all of the members of the EMS agree.

In empirical work, the ERM has not yet received much attention. Most empirical studies on exchange rates are concerned with U.S. dollar rates. These exchange rates exhibit leptokurtic behavior (fatter tails than usual) and clusters of high and low volatility, but not significant serial correlation. These stylized facts can be reproduced by means of an autoregressive conditional heteroscedasticity (ARCH) (Engle 1982), generalized autoregressive heteroscedasticity (GARCH) (Bollerslev 1986), or exponential GARCH (EGARCH) (Nelson 1991) specification, as done, for instance, by Diebold (1988), Hsieh (1989), or Baillie and Bollerslev (1989).

Especially for high frequency financial data, however, the GARCH specification cum normal innovations cannot fully explain the leptokurtic behavior. That is why several authors used distributions other than the normal, such as the student-t, a discrete mixture of normals, a generalized error distribution, or a normal–Poisson distribution (e.g., Baillie and Bollerslev 1989; Boothe and Glassman 1987; Hsieh 1989; Jorion 1988; Nelson 1991; Tucker and Pond 1988).

Since the ERM currencies have to stay within a target zone, we should expect their statistical properties to differ from those of the free-float rates. As long as no realignment occurs, the changes of the rates have to be small and mean-reverting to stay within the band. At times of realignments, however, large depreciations can occur resulting in much fatter tails of the distribution of the ERM currencies than for the free-float currencies. Moreover, since all realignments within the EMS meant an appreciation of the Deutsche mark (D-mark), the distributions of all ERM currencies expressed in terms of the D-mark are positively skewed.

In the literature, the effects of a target zone on exchange rates have been mostly modeled along the lines of Krugman (1991). In his model, the expected future spot rates are affected by the target zone because central banks will intervene whenever the exchange rate reaches the boundary of the band. This type of model heavily relies on the assumption that the intervention policy is fully credible and known to the public. Moreover, the Krugman model does not account for parity realignments. Since in reality the nature of the policy interventions is not known, it is hardly surprising that these models are not fully supported by empirical evidence.

Another line of research on the ERM currencies was followed by Nieuwland, Verschoor, and Wolff (1991). They analyzed weekly D-mark rates of several ERM currencies by estimating a model with an ARCH(1) specification and a mixed normal–Poisson distribution, as in the work of Jorion (1988). The Poisson process generates jumps, which might reproduce the discontinuities arising from (anticipations of) parity adjustments.

In this article, we generalize these models in four directions, using weekly observations on the D-mark rates of the Belgian franc, the Dutch guilder, the French franc, the Danish kroner, the Irish pound, the Italian lira, and, for comparison reasons, the British pound and the U.S. dollar, for the period April 4, 1979, to March 27, 1991.

First, a moving average (MA) specification is included to allow for mean reversion of the exchange rates. Second, the ARCH(1) specification is replaced by the more general GARCH(1, 1) specification. Third, the normal–Poisson mixture is compared to mixtures of two, three, and four normal distributions, which can be more easily estimated and interpreted than a Poisson mixture. The inclusion of a jump process reduces the influence of outliers on the MA–GARCH specification and accounts for skewness and high excess kurtosis. Economic explanations for the presence of outliers are speculative attacks inside the band and realignments. An adjusted goodness-of-fit test is proposed and used to check the fit of the combined GARCH–jump processes. Fourth, the six ERM currencies are estimated in a multivariate setting. An MA(1)–GARCH(1, 1)–Bernoulli–normal model is estimated assuming a constant correlation matrix and an identical jump probability for all currencies. Because the univariate models used are consistent with the multivariate model, the appropriateness of the multivariate model is also an indication of that of the univariate models. Moreover, the multivariate model can be used to determine a mean-variance efficient portfolio of EMS currencies.

The structure of the article is as follows. Section 1 describes the data, and information on the differenced series is given; Section 2 contains the statistical models; Section 3 provides empirical results; and finally Section 4 summarizes the main results. The Appendix describes the moments of the Bernoulli–normal mixture distributions.

1. THE DATA

The data consist of 626 weekly Wednesday closing rates from the London Eurocurrency market in terms of the D-mark from April 4, 1979, to March 27, 1991. All data are taken from Datastream. As the data series were reported in terms of the U.K. pound, we computed D-mark rates, assuming perfect arbitrage. The countries under consideration are those participating in the ERM of the EMS since the beginning (3/13/79)—that is, Belgium, France, Denmark, Ireland, Italy, and the Netherlands, and, for reasons of comparison, the United States and the United Kingdom. For the U.K. pound, only the free-float period (until October 8, 1990) is considered.

Before specifying the model, we take a closer look at the statistical properties of the ln D-mark rates. The first row of Table 1 shows that the unit-root hypothesis is rejected for the deviation of the exchange rate from the central parity (s-p). If we do not account for realignments, however, the unit-root hypothesis can no longer be rejected. That is why the model will be estimated in first differences.

The skewness is significantly positive for all ERM currencies. This might be a result of the asymmetry in the movements of the parity adjustments. Compared to the free-float currencies, due to large outliers the excess kurtosis is extremely high for all ERM currencies, especially for the Belgian franc, the French franc, and the Irish pound. A high kurtosis could result from a time-varying variance. A time-varying second moment might be detected by $Q_{yy}(25)$—a Box–Pierce test applied to the squared data (Bollerslev 1988). This test statistic requires a finite fourth moment, however, which might be doubtful for our data given the high kurtosis.

Table 1. Summary Statistics of ln D-mark Rates

Statistics		BFr	DFl	FFr	DKr	IPd	ILi	BPd	US$
ADF	$s-p$	−5.84	−4.39	−4.32	−5.04	−4.94	−3.97		
ADF	s	−2.09	−2.79	−1.68	−3.60	−1.29	−1.97	−.65	−.86
Mean ($\times 10^2$)	Δs	.04	.01	.06	.05	.06	.08	.04	.01
Standard deviation ($\times 10^2$)	Δs	.51	.24	.48	.48	.49	.58	1.15	1.57
Skewness	Δs	5.07	.26	6.17	1.41	4.29	2.08	.40	.44
Exc. kurtosis	Δs	67.02	5.55	64.96	24.49	48.37	16.50	1.99	1.02
$\rho_y(1)$	Δs	−.15	−.19	.02	−.21	−.11	−.14	.14	.07
$Q_y(25)$	Δs	74.56	73.67	27.90	66.78	43.34	63.90	45.46	23.30
$Q_{yy}(25)$	Δs	10.39	136.00	1.38	68.02	4.56	61.34	52.23	42.58
$Q_{\|y\|}(25)$	Δs	186.21	332.80	23.52	81.99	57.16	151.24	70.90	47.50

NOTE: The data consist of 626 weekly Wednesday spot rates, expressed in domestic currency per D-mark, from April 4, 1979, to March 27, 1991. For the British pound, the sample runs until October 3, 1990. s is the ln D-mark rate and p is the ln central parity. ADF is the augmented Dickey–Fuller test statistic with a constant and one lag of the differenced series. The 5% critical value is −2.87.

That is why we also report the Box–Pierce test for the absolute value of the exchange-rate changes $[Q_{|y|}(25)]$, which requires only a finite second moment. It turns out that the absolute changes are indeed significantly correlated for all currencies except the French franc, whereas the null hypothesis of no serial correlation in the squared data was not rejected for the currencies with the highest excess kurtosis.

Finally, the first-order serial correlation coefficient $\rho_y(1)$ and the Box–Pierce statistic of the raw data $[Q_y(25)]$ are highly significant for all currencies except for the French franc and the U.S. dollar, indicating serial correlation in the series. The negative autocorrelation for the ERM currencies is probably due to the stabilizing effects of the intervention policy and partly produces mean reversion.

2. THE MODEL

The results for the ERM countries presented in Section 1 differ from those for the free-float currencies. First of all, first differences of ERM rates exhibit significant negative autocorrelation. This is a result of the mean reversion needed to restrict the rates to stay within the target zone. In the model, mean reversion is accounted for by the MA parameters.

A second aspect in which ERM exchange rates differ from free-float series concerns the skewness. ERM exchange rates in terms of the D-mark are not symmetric. As a consequence, symmetric distributions such as the normal, student-t, or generalized error distribution are unlikely to give appropriate results. We combine normal distributions and a stochastic jump process to account for skewness and leptokurtosis. For the jump intensity (λ), we concentrate on the Bernoulli distribution (Ball and Torous 1983), but the Poisson distribution is also computed for comparison reasons. The jump size is assumed to be normally distributed with expectation ν and variance δ^2. These distributions can explain the skewness and the leptokurtic behavior of a series, as will be shown in the Appendix. Jumps in the ERM exchange rates occur in abnormal circumstances such as parity adjustments but also, for instance, as a result of expected policy changes after an election (Ungerer, Hauvonen, Lopez-Claros, and Mayer 1990, table 1), speculative attacks, or changes in interest rates.

A third feature of our models is the presence of conditional heteroscedasticity in the data that will be accounted for by the GARCH specification.

Finally, we should consider the interpretation of an intercept denoted by μ. Since we estimate the model in first differences, this parameter represents the slope of a deterministic time trend. Within the ERM, the presence of a time trend would require frequent parity adjustments, or jumps in the opposite direction within the band. Although it is unlikely that a time trend is significant, we estimate the models with intercept because it represents part of the first moment of the data. If we would not include this term into our specification, this first moment would affect ν. In that case, ν could not fully explain the third moment, and first and third moments would not be variation-free.

To specify the log-likelihood functions, we first write the model as

$$\Delta s_t = \mu + \lambda\nu + \varepsilon_t + \sum_{i=1}^{r} \psi_i \varepsilon_{t-i}, \tag{1}$$

where ε_t is the disturbance, which has expectation 0 (see the Appendix). The normal–Bernoulli log-likelihood function has the following form:

$$\begin{aligned}\ln(L_b) = {} & -(T/2)\ln(2\pi) + \sum_{t=1}^{T} \ln\{[(1-\lambda)/h_t] \\ & \times \exp[-(\varepsilon_t + \lambda\nu)^2/2h_t^2] + [\lambda/(h_t^2+\delta^2)^{1/2}] \\ & \times \exp[-(\varepsilon_t - (1-\lambda)\nu)^2/2(h_t^2+\delta^2)]\}.\end{aligned} \tag{2}$$

The normal–Poisson log-likelihood $[\ln(L_p)]$ can be written as

$$\begin{aligned}\ln(L_p) = {} & -T\lambda - (T/2)\ln(2\pi) \\ & + \sum_{t=1}^{T} \ln\left\{\sum_{j=0}^{\infty} \lambda^j/[j!(h_t^2+\delta^2 j)^{1/2}]\right. \\ & \left.\times \exp[-(\varepsilon_t - (j-\lambda)\nu)^2/2(h_t^2+\delta^2 j)]\right\}.\end{aligned} \tag{3}$$

The GARCH (p, q) specification is the same for the two models:

$$h_t^2 = \alpha_0 + \sum_{i=1}^{p} \beta_i h_{t-i}^2 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2. \tag{4}$$

In the economic literature, stochastic jumps are mostly modeled by means of a Poisson distribution (Akgiray and Booth 1988; Ball and Torous 1985; Feinstone 1987; Hsieh 1989; Jorion 1988; Nieuwland et al. 1991; Tucker and Pond 1988). As can be seen from Formula (3), a difficulty with a Poisson function is that it contains an infinite sum. This sum has to be truncated for the process to become estimable. Ball and Torous (1985) gave an upper limit for the truncation error. We used a truncation after 11 terms, which seems to be appropriate for most applications. Given an appropriate truncation, we can estimate Model (3), together with the mean and variance specifications (1) and (4), by maximum likelihood. Even if we truncate after only five terms, maximum likelihood estimation [combining (1), (3), and (4)] is quite complicated. The solution turns out to be sensitive to the choice of the starting values. Ball and Torous (1985) suggested using the Bernoulli distribution to get starting values for the Poisson distribution. This procedure is followed here.

Both the GARCH specification and the jump process (see the Appendix) can explain the leptokurtic behavior of the series. Since the statistical and economic motivations for GARCH effects and jumps are quite different, we choose a model specification that accounts for the two simultaneously. After a jump has taken place (for instance a parity adjustment), volatility will be high, but gradually it will return to normal values when a new equilibrium is reached. If we did not include GARCH, the large volatility following a jump would mistakenly be taken for additional jumps. The jump intensity would rise, as can be seen, for instance, in the work of Jorion (1988, tables 3 and 4) and Nieuwland et al. (1991, tables 4 and 6).

3. EMPIRICAL RESULTS

The lengths of the lag structures for the MA and GARCH specifications have been determined using a likelihood-ratio test with a 5% marginal significance level. The critical values of this statistic are adjusted because nonnegativity constraints (Kodde and Palm 1986) of the GARCH parameters lead to a test of a one-sided hypothesis.

Positivity of α_i and β_i is enforced by estimating $\alpha_i^{1/2}$ and $\beta_i^{1/2}$, although this is not strictly necessary for GARCH models of order higher than (1, 1) (Nelson and Cao 1992). For the Bernoulli model, λ is forced between 0 and 1 by estimating k in the expression $\lambda = (1 + \exp(k))^{-1}$.

Tables 2 and 3 contain the empirical results of MA(r)–GARCH(p, q) models for each currency. Three different distributions are used to estimate these models. The first is a normal distribution, the second a normal–Bernoulli mixture, and the last distribution is a normal–Poisson mixture. Since the results for the normal–Poisson and normal–Bernoulli mixtures were very similar, only the latter are shown (see also Table 4). All models are estimated with the maximum likelihood routine from the software package Gauss using the algorithm of

Table 2. Coefficients of MA(r)–GARCH(p, q) Models

Distribution	μ ($\times 10^4$)	ψ_1	α_0 ($\times 10^6$)	α_1	β_1	λ	ν ($\times 10^2$)	δ^2 ($\times 10^4$)
BFr								
normal	1.277	−.363	.000	.094	.926			
	(1.43)	(8.19)	(.00)	(3.53)	(47.30)			
Bern/norm	−.423	−.365	.537	.149	.769	.023	1.264	3.608
	(.47)	(8.05)	(1.73)	(2.07)	(7.93)	(2.25)	(1.66)	(1.21)
DFl								
normal	−.011	−.245	.065	.298	.763			
	(.02)	(3.69)	(.97)	(2.25)	(8.72)			
Bern/norm	−.485	−.332	.096	.263	.651	.075	.120	.222
	(1.02)	(6.38)	(1.20)	(3.43)	(6.04)	(2.32)	(1.39)	(2.16)
FFr								
normal	2.857	.249	1.661	1.051	.527			
	(1.16)	(.39)	(.58)	(1.92)	(2.55)			
Bern/norm	1.144	−.138	4.212	.157		.036	.964	3.436
	(1.31)	(3.96)	(8.51)	(2.88)		(2.90)	(2.16)	(2.21)
DKr								
normal	3.157	−.131	.675	.150	.847			
	(2.36)	(1.70)	(.99)	(1.83)	(11.67)			
Bern/norm	.584	−.156	7.078	.151		.041	1.076	1.312
	(.53)	(3.19)	(11.11)	(3.98)		(3.09)	(3.38)	(2.76)
IPd								
normal	1.263	−.358	9.955	1.051				
	(1.10)	(4.62)	(3.71)	(2.32)				
Bern/norm	1.211	−.213	4.782	.305	.115	.034	.614	3.189
	(1.18)	(5.33)	(5.41)	(2.61)	(1.68)	(1.80)	(1.44)	(1.40)
ILi								
normal	5.142	.003	8.519	.315	.485			
	(2.03)	(.06)	(2.35)	(1.75)	(3.74)			
Bern/norm	−.772	−.144	2.897	.180	.396	.094	.674	1.380
	(.57)	(3.56)	(2.43)	(2.58)	(2.76)	(2.93)	(3.29)	(2.26)
BPd								
normal	5.073	.144	6.767	.071	.877			
	(1.04)	(2.82)	(1.53)	(2.40)	(18.48)			
Bern/norm	−5.048	.068	.000	.084	.834	.363	.285	1.771
	(1.21)	(1.33)	(.01)	(1.87)	(8.57)	(2.70)	(1.88)	(4.41)
US$								
normal	1.323		35.445	.135	.727			
	(.22)		(2.79)	(3.25)	(10.88)			
Bern/norm	−29.744		15.925	.121	.721	.500	.646	1.493
	(1.79)		(1.52)	(2.71)	(9.71)	(1.78)	(2.00)	(3.19)

NOTE: Absolute asymptotic heteroscedasticity-consistent t values are in parentheses.

Table 3. Diagnostics of MA(r)−GARCH (p, q) Models

Distribution	ln(L)	$\chi^2(19)$	$Q_y(25)$	$Q_{yy}(25)$	$Q_{\lvert y \rvert}(25)$	M3	M4
Bfr							
normal	2560.2	66.5	29.5	11.4	31.8	2.29	21.31
		[.000]	[.245]	[.990]	[.162]	(23.35)	(108.85)
Bern/norm	2654.2	20.5	35.0	21.3	27.6	−.05	.22
		[.364]	[.087]	[.677]	[.326]	(.51)	(1.13)
DFl							
normal	3010.8	53.8	32.0	14.9	26.7	.59	5.09
		[.000]	[.159]	[.943]	[.373]	(5.98)	(25.99)
Bern/norm	3063.6	13.4	32.2	15.6	18.9	.06	.10
		[.820]	[.152]	[.927]	[.804]	(.57)	(.52)
FFr							
normal	2522.6	210.0	34.6	1.9	18.9	4.36	43.77
		[.000]	[.096]	[1.00]	[.801]	(44.49)	(223.52)
Bern/norm	2802.5	38.9	24.2	22.0	27.8	.11	.33
		[.005]	[.506]	[.634]	[.317]	(1.14)	(1.69)
DKr							
normal	2534.1	110.2	28.1	21.8	33.2	1.92	9.84
		[.000]	[.302]	[.646]	[.126]	(19.58)	(50.27)
Bern/norm	2656.3	21.8	25.0	40.4	36.3	.08	.20
		[.295]	[.463]	[.027]	[.066]	(.78)	(1.00)
IPd							
normal	2534.8	139.8	47.6	4.4	27.3	2.81	22.77
		[.000]	[.004]	[1.00]	[.341]	(28.73)	(116.27)
Bern/norm	2675.4	36.5	40.0	41.0	42.2	.13	.46
		[.009]	[.029]	[.023]	[.017]	(1.31)	(2.37)
ILi							
normal	2401.3	149.1	34.2	36.6	19.2	2.35	20.31
		[.000]	[.103]	[.063]	[.789]	(23.99)	(103.71)
Bern/norm	2561.9	27.5	35.9	20.2	22.2	.13	.33
		[.093]	[.073]	[.739]	[.623]	(1.37)	(1.69)
BPd							
normal	1856.4	47.8	32.3	16.6	26.9	.39	2.01
		[.000]	[.150]	[.896]	[.363]	(3.92)	(10.06)
Bern/norm	1887.4	18.9	33.9	27.0	25.0	.07	.07
		[.465]	[.110]	[.355]	[.464]	(.66)	(.37)
US$							
normal	1727.8	29.0	25.7	22.1	23.8	.28	.45
		[.066]	[.424]	[.629]	[.530]	(2.81)	(2.32)
Bern/norm	1731.1	23.4	26.5	22.6	23.8	.08	.09
		[.220]	[.382]	[.543]	[.531]	(.83)	(.47)

NOTE: Marginal significance levels and absolute *t* values are in square brackets and parentheses, respectively.

Broyden, Fletcher, Goldfarb, and Shano (see Broyden 1965).

In the models without stochastic jumps, negative autocorrelation is not always present. For the French franc

Table 4. Schwarz Criteria and Goodness-of-Fit Measures for the MA(1)−GARCH(1, 1) Specification

Currency		1 normal	2 normals	3 normals	4 normals	Poisson/ normal
BFr	SC	−5088.8	−5256.9	−5260.2	−5247.6	−5257.9
	P	.000	.364	.521	.534	.474
DFl	SC	−5989.4	−6075.7	−6062.6	−6043.8	−6075.9
	P	.000	.820	.942	.949	.833
FFr	SC	−5013.0	−5553.5	−5556.6	−5546.8	−5553.7
	P	.000	.005	.183	.840	.025
DKr	SC	−5036.0	−5261.1	−5248.4	−5233.4	−5261.1
	P	.000	.295	.474	.577	.371
IPd	SC	−5043.8	−5299.3	−5317.6	−5299.4	−5299.7
	P	.000	.009	.780	.806	.015
ILi	SC	−4770.4	−5072.3	−5075.6	−5056.4	−5076.1
	P	.000	.093	.564	.516	.474
BPd	SC	−3681.2	−3723.6	−3713.8	−3697.4	−3724.2
	P	.000	.465	.861	.870	.683
US$	SC	−3429.8	−3417.1	−3400.0	−3381.5	−3417.7
	P	.066	.220	.308	.546	.210

NOTE: The Schwarz criterion is given by $SC = -2\ln(L) + n\ln(T)$, n = number of parameters. *P* represents the *p* value of the $\chi^2(19)$ goodness-of-fit test.

and the Italian lira, the MA parameter is even positive, although not significant, whereas we expected a negative sign from economic theory. Furthermore, the estimated GARCH parameters for the Belgian franc, the Dutch guilder, and the French franc seem to be too high since they are explosively nonstationary. The strict stationarity condition is not fulfilled for these currencies: $E[\ln(\beta_1 + \alpha_1 Z^2)] > 0$ with $Z \sim N(0, 1)$ (Nelson 1990, fig. 1). The α_1 parameter for the Irish pound is also greater than 1, violating the weak stationarity condition, but the strict stationarity condition is fulfilled for this currency. The high values for the GARCH parameters can be explained from misspecification. In these models, the increased volatility resulting from a jump is probably captured by the high values of α_1 or β_1.

When jumps are taken into account, the MA parameter becomes negative and significant for all ERM currencies and the GARCH specification is weakly stationary for all currencies with r, p, and q being at most equal to 1. The GARCH(1, 1) specification is appropriate for all series, except for the French franc and the Danish kroner, for which an ARCH(1) specification is sufficient. In none of these specifications is the intercept significant.

For the Bernoulli model, the jump intensity ranges from 2.3% (Belgian franc) to 9.4% (Italian lira). Since we have 626 observations, this means the estimated expected number of jumps lies between 14 and 59. Given the fact that there have only been 12 parity adjustments during the sample period, part of the jumps must have taken place within the band, for instance, as a result of speculative attacks.

For all currencies, the expected jump size ν is positive, which is in accordance with the positive skewness. It is significantly different from 0 (at the 5% level) for the French franc, the Danish kroner, the Italian lira, and the U.S. dollar.

To check for the appropriateness of the model (which assumes normally distributed jump sizes and innovations), a Pearson chi-squared goodness-of-fit test is performed on the residuals of the estimated models. This goodness-of-fit test compares the empirical distribution of the (standardized) residuals with the theoretical distribution. This is done by classifying the residuals in groups according to their magnitude. For iid observations, it can be shown that $\Sigma_{i=1}^{g}(n_i - En_i)^2/En_i \sim \chi^2(g - 1)$, where g is the number of groups and n_i is the number of observations in each group. For a model with estimated parameters, the distribution is actually bracketed between $\chi^2(g - 1)$ and $\chi^2(g - k - 1)$, where k is the number of estimated parameters in the likelihood function (Kendall and Stuart 1967).

A problem with this test statistic for models with a time-varying variance and jumps is that their residuals are neither identically nor independently distributed. Even if we standardize the residuals, the higher moments of the MA–GARCH–jump models will still be time-varying. This problem can be solved by reclassifying the residuals. Instead of classifying the residuals according to their value, we calculate the probability of observing a value smaller than the residual. These probabilities should be identically uniformly distributed between 0 and 1. The grouping mechanism for equally sized groups thus becomes

$$n_i = \sum_{t=1}^{T} I_{it}, \text{ where } I_{it} = 1 \quad \text{if } (i - 1)/g < F(X_t, \hat{\phi}) \leq i/g$$
$$= 0 \quad \text{otherwise}$$

($\forall\, 1 \leq i \leq g$), where $F(X_t, \hat{\phi})$ represents the value of the cumulative distribution function given the data and the estimated parameters.

In the second column of Table 3, the results of this test are shown for $g = 20$. The first row for each currency clearly shows the inappropriateness of the MA–GARCH normal model for ERM currencies. The smallest value of this $\chi^2(19)$ statistic, that for the Dutch guilder, is still 54. The rejection according to this statistic of the model for the British pound is also noticeable. Although the excess kurtosis and skewness for this currency are much less extreme than for ERM currencies (Table 1), normality is still rejected. The null hypothesis is not rejected for the U.S. dollar.

When stochastic jumps are included, the results improve tremendously. Using the goodness-of-fit test, the Bernoulli–normal model is not rejected at the 5% level for six out of eight currencies. The only models that are rejected at the 1% level are the Bernoulli–normal mixtures for the French franc and the Irish pound. The improvement of the fit for ERM currencies can also be seen from the first column of Table 3. The log-likelihood function increases by values between 53 (Dutch guilder) and 280 (French franc) points when stochastic jumps are included.

Correlation in the residuals, the squared residuals, and the absolute residuals can be detected using the Box–Pierce statistics $Q_y(25)$, $Q_{yy}(25)$, and $Q_{|y|}(25)$. Notice that these tests assume normality. For the MA–GARCH–normal models, we can compute these test statistics for the standardized residuals. For the Bernoulli–normal mixture, however, this procedure will not be of help. For these models we compute "normalized" residuals by means of the cumulative distribution function. Given the values of the cumulative distribution function for the Bernoulli–normal mixture models, which we used for the Pearson goodness-of-fit test, we can compute the normalized residuals by means of the inverse of the cumulative normal distribution function.

Under the assumption of a correctly specified model, the residuals should be independent. This can be checked by means of the Box–Pierce tests on the normalized residuals. It turns out that the MA(1) specification appropriately accounts for the correlation in the mean.

The resulting autocorrelations in the residuals are insignificant for all currencies, except the Irish pound. The time-dependence of volatility seems to be correctly modeled as well, since both the correlations in the squared residuals, and in the absolute residuals are insignificant for most currencies.

Finally, $M3$ and $M4$ measure the skewness and excess kurtosis of the residuals. Under the assumption of normality, these test statistics are asymptotically normally distributed with expectation 0 and variance $6/T$ and $24/T$, respectively. The hypothesis of normality is clearly rejected for all the MA–GARCH normal models, even for the U.S. dollar. For the MA–GARCH and Bernoulli–normal models, however, the results are much better. Only for the Irish pound is the excess kurtosis of the normalized residuals significant at the 5% level.

From these results we conclude that the normal–Bernoulli mixture performs rather well for the weekly ERM exchange-rates data, except for the French franc and the Irish pound. One possible reason for the failure of the normal–Bernoulli distribution to pass the goodness-of-fit test for these currencies might be that the number of two normal distributions included in the mixture is too low. A possible reason for adding a third normal distribution to the mixture could be that the jump sizes at realignment dates are much larger than for the jumps inside the band.

Table 4 shows the value of the Schwarz (1978) criterion and the p values associated with the goodness-of-fit test for the results of MA–GARCH models for mixtures of up to four normal distributions and for the normal–Poisson mixture. The mixtures of normals are estimated as "models with jumps" in the sense that the mean of the jump size ν_i can vary with i and the variances of the second to fourth distributions are estimated as $\text{var}_{it} = h_t^2 + \delta_i^2$. This procedure is preferred to that of specifying independent variances, since it seems reasonable to assume that the same GARCH effect is present in all variances. The weights of the separate distributions are determined by estimating k_i in $\lambda_i = \text{abs}(k_i)/(\Sigma_j \text{abs}(k_j))$, $k_1 = 1$.

The values of the Schwarz criterion (SC) hardly differ for Bernoulli and Poisson mixtures of normals. As the mixtures of two, three, or four normals are not nested in the Poisson–normal model, a standard likelihood-ratio test cannot be applied to compare the models. Tucker and Pond (1988) showed, by means of a simulation experiment, that the SC can be used to discriminate between mixtures of normals and Poisson–normal models. With the exception of the lira, the value of the SC of the Bernoulli mixture differs by at most one point from that of the Poisson–normal process. The implied values of the log-likelihood function for these models hardly differ either. Notice also that the Pearson statistic is similar for these models. This is remarkable especially for the free-float currencies, since the jump intensity λ is quite high for these currencies (Table 2).

The results in Table 4 indicate that there are no strong reasons to prefer the Poisson mixture to the Bernoulli–normal model. Since the Bernoulli–normal mixture is much easier to estimate than the Poisson–normal mixture and does not require choosing a truncation point, we prefer the first. Moreover, the economic interpretation of the Bernoulli normal mixture is more appealing than that of the normal–Poisson one. The Poisson process models large changes as a sequence of several jumps occurring within a week. Within the ERM, however, the largest weekly changes in the exchange rate originate from large devaluations, which can hardly be seen as a sum of small changes. For three currencies, the mixture of three normals performs best, according to the SC. Not surprisingly, these three currencies (plus the Italian lira for which the mixture of three was also preferred to the mixture of two normals) were exactly those that experienced large devaluations during the sample period. In the mixture of three normals, these large devaluations were modeled by a separate jump process. For the U.S. dollar, the normal distribution shows the lowest SC.

These results differ from the results of Tucker and Pond (1988) and Akgiray and Booth (1988), who concluded that, for daily U.S. dollar data, the normal–Poisson mixture always performed better than a discrete mixture of up to five normals. The differences in log-likelihood values they found are also much larger than those found in our study. This finding might be the result of including a GARCH specification in our models. The Poisson–jump specification generates a mixture of normal distributions at the price of just including three parameters. For the other models, the compound normal distribution needs three additional parameters for each normal distribution included in the mixture. Since, in the absence of a GARCH specification, part of the changing variance over time has to be taken into account by including additional normals in the mixture, the more parsimoniously parameterized Poisson specification might perform better.

Finally, we jointly estimate the MA(1)–GARCH(1, 1) model with Bernoulli–normal mixture distribution for the six ERM currencies, assuming a constant correlation matrix (e.g., see Bollerslev 1990) and identical jump probability for all six currencies. Since these currencies participate in the ERM and are expressed in terms of the D-mark, it is plausible to assume their disturbances to be correlated. When, for instance, the German mark is strong compared to the U.S. dollar, all ERM currencies are expected to be affected. Similarly, a stochastic shock leading to a jump is likely to simultaneously affect all of the currencies in the system. Therefore, we restrict the jump probability to be the same for the currencies. The size of the jump, however, is allowed to differ across them. Because the univariate models estimated previously are consistent with this multivariate model, the results for the multivariate model provide information on the appropriateness of the uni-

Table 5. The Multivariate MA(1)–Garch(1, 1)–Bernoulli/Normal Model

	BFr	DFl	FFr	DKr	IPd	ILi
ν (×10²)	.399	.107	.444	.503	.294	.709
	(1.57)	(1.37)	(1.86)	(2.51)	(1.58)	(2.77)
δ^2 (×10⁴)	1.918	0.214	2.585	1.354	1.772	1.857
	(1.25)	(2.72)	(2.02)	(2.10)	(1.42)	(2.95)
			Correlation matrix			
BFr	1	.333	.384	.341	.320	.366
		(8.49)	(9.10)	(8.41)	(6.40)	(7.37)
DFl		1	.333	.249	.219	.324
			(7.93)	(6.44)	(4.67)	(6.66)
FFr			1	.323	.332	.459
				(8.29)	(7.40)	(12.45)
DKr				1	.253	.284
					(6.19)	(7.88)
IPd					1	.231
						(5.51)
ILi						1

NOTE: Heteroscedasticity-consistent t values are in parentheses. The log-likelihood value is 16, 704.5. The common jump intensity equals .0625 (3.14).

variate models. In addition, for the composition of an efficient portfolio of ERM currencies, information on the cross-correlations is required. The resulting multivariate model has the following form: $\Delta s_{it} = \mu_i + \lambda\nu_i + \psi_i\,\varepsilon_{it-1} + \varepsilon_{it}$, with i denoting the country and the vector ε_t having the distribution $\varepsilon_t \sim (1-\lambda)N(-\lambda\nu,\ \bar{h}_t'\Omega\bar{h}_t) + \lambda N((1-\lambda)\nu,\ (\bar{h}_t+\bar{\delta})'\Omega(\bar{h}_t+\bar{\delta}))$, where ε_t and ν are 6×1 vectors with typical element ε_{it} and ν_i and where $\bar{h}_t$ and $\bar{\delta}$ are diagonal matrices with typical elements h_{it} and δ_i, respectively, with $h_{it}^2 = \alpha_{i0} + \alpha_{i1}\varepsilon_{it-1}^2 + \beta_{i1}h_{it-1}^2$. To restrict the elements of the correlation matrix Ω to be smaller than 1 in absolute value, we estimate them by estimating $\omega_{ij} = 2\arctan(\rho_{ij})/\pi$. The jump probability λ is common to all countries. The results for this model are given in Table 5. A selection of the coefficients estimates is included in this table since many parameters are essentially the same as for the univariate models, a finding that is expected to occur when both the univariate and multivariate models are correctly specified. The point estimate of λ is .0625. The estimated jump sizes differ somewhat from those of the univariate models. These differences reflect the differences between the estimates of λ_i for the univariate models. All elements of the correlation matrix range between .22 and .46 and are highly significant, suggesting that the ERM rates should be jointly modeled.

The value of the log-likelihood function is 16,704.5, which is 290.5 points larger than the sum of the values of the log-likelihood functions for the univariate models. Although the models are not nested, the increase of the log-likelihood function by 290.5 resulting from adding 10 extra parameters seems to be sufficient to conclude that the exchange rates in the ERM are highly interrelated. The joint model presented previously yields a satisfactory description of the interrelationships.

4. CONCLUSIONS

In this article we examined the time series properties of weekly exchange rates participating in the ERM of the EMS. One feature of the ERM is that the central banks of the participating countries intervene whenever a bilateral exchange rate deviates more than 2.25% from its agreed central rate. Therefore, exchange rates within the ERM should exhibit parity reversion.

There have been several parity adjustments in the last 12 years, resulting in discontinuities in the time series. Since at times of a realignment the change in the spot rate can be very large compared to the usual variations within the band, the excess kurtosis for ERM currencies is much larger than that for free-float currencies such as the U.S. dollar or the British pound. Moreover, since all realignments within the ERM in fact have been appreciations of the D-mark, all ERM rates in terms of the D-mark are positively skewed.

These features have been modeled using an MA(1)–GARCH(1, 1)–jump model. The parity reversion is captured by negative MA parameters, which turn out to be highly significant for all models. The parity adjustments are taken into account by means of a stochastic jump process. We compare the Bernoulli and Poisson specifications. For most currencies, the two give similar results. A GARCH specification accounts for the changing volatility over time. If we do not allow for stochastic jumps, the MA–GARCH specification changes dramatically. For three out of six currencies, the GARCH specification becomes explosive and for two the MA parameter has the wrong sign, although they are not significant. These findings can be explained from the influence of outliers, resulting from parity adjustments, on the model. For the U.S. dollar, jumps are not significant.

The models are checked by means of an adjusted Pearson chi-squared goodness-of-fit test. For four out of six ERM exchange rates, both the Bernoulli and the Poisson specification pass the test. For the other two, a mixture of three normal distributions, which can be interpreted as a normal distribution plus two independent jump specifications, performs best.

Finally, joint estimates of an MA(1)–GARCH(1, 1) normal–Bernoulli model with constant contemporaneous correlation matrix have been obtained. The results for the joint model were found to be very much in line with the findings for the single processes, but, as expected, the efficiency of the estimates improved.

One of the main aims of future research will be to allow for a time-varying jump intensity. Since the economic performance of the ERM countries has become more similar and, as a consequence, frequency and size of realignments have been declining, we should expect the influence of the jumps in our models to have declined as well. Estimates of the model over three subperiods did indeed indicate that the importance of the jump process has diminished, especially since 1987. This will be modeled by making the jump intensity (or size) a function of economic indicators, such as trade deficits, inflation differentials, or interest differentials. The dependence of λ on deviations of the spot rate from the parity rate has been investigated but is not significantly different from 0. The very recent turbulences in the EMS stress once again the importance of allowing the model to account for the occurrence of jumps in the form of parity realignments even though no realignment had taken place since 1987 and an estimate of the jump intensity based on the recent five years only would have been close to 0.

ACKNOWLEDGMENTS

This research was sponsored by the Economics Research Foundation, which is part of the Netherlands Organization for Scientific Research (NWO). We thank Denis de Crombrugghe, Frank de Jong, Kees Koedijk, Peter Schotman, Philip Stork, an associate editor, and two anonymous referees for many helpful comments. We are responsible for any errors.

APPENDIX: MOMENTS OF THE BERNOULLI JUMP PROCESS

The distribution of the error term of the Bernoulli jump model with normally distributed jump sizes and innovations can be expressed in the following way: $\varepsilon_t^{\text{Bern}} \equiv \Delta s_t - \mu - \psi\, \varepsilon_{t-1}^{\text{Bern}} - \lambda\nu \sim (1 - \lambda)\, N(-\lambda\nu, h_t^2) + \lambda N((1 - \lambda)\nu, h_t^2 + \delta^2)$. From this distribution we compute the first four unconditional moments, μ_1^B to μ_4^B, as linear combinations of noncentral moments of normal distributions:

$$E(\varepsilon_t^{\text{Bern}}) = E[(1 - \lambda)(-\lambda\nu) + \lambda(1 - \lambda)\nu] = 0 = \mu_1^B, \qquad \text{(A.1)}$$

$$E(\varepsilon_t^{\text{Bern}})^2 = \mathrm{E}[(1 - \lambda)\{(-\lambda\nu)^2 + h_t^2\} + \lambda\{(1 - \lambda)^2\nu^2 + h_t^2 + \delta^2\}],$$

and

$$\mu_2^B = (\lambda - \lambda^2)\nu^2 + Eh^2 + \lambda\delta^2, \qquad \text{(A.2)}$$

where Eh^2 is the unconditional expectation of h_t^2, which can be estimated by its sample mean. For the GARCH(p, q) specification, it is possible to eliminate Eh^2 and express μ_2^B only in terms of the parameters in the model. This is not done here to keep the result applicable to other variance specifications as well.

$$E(\varepsilon_t^{\text{Bern}})^3 = E[(1 - \lambda)\{(-\lambda\nu)^3 - 3\lambda\nu h_t^2\} + \lambda\{(1 - \lambda)^3\nu^3 + 3(1 - \lambda)\nu(h_t^2 + \delta^2)\}]. \qquad \text{(A.3)}$$

$$\mu_3^B = (\lambda - \lambda^2)\nu\{(1 - 2\lambda)\nu^2 + 3\delta^2\}.$$

$$M3^B \equiv \mu_3^B/(\mu_2^B)^{3/2} = (\lambda - \lambda^2)\nu\,\{(1 - 2\lambda)\nu^2 + 3\delta^2\} \div ((\lambda - \lambda^2)\nu^2 + Eh^2 + \lambda\delta^2)^{3/2}.$$

$$E(\varepsilon_t^{\text{Bern}})^4 = E[(1 - \lambda)\{(-\lambda\nu)^4 + 6(-\lambda\nu)^2h_t^2 + 3h_t^4\} + \lambda\{(1 - \lambda)^4\nu^4 + 6(1 - \lambda)^2\nu^2(h_t^2 + \delta^2) + 3(h_t^2 + \delta^2)^2\}].$$

$$\mu_4^B = (\lambda - \lambda^2)\nu^2\{(-3\lambda^2 - 3\lambda + 1)\nu^2 + 6Eh^2 + 6(1 - \lambda)\delta^2\} + 3Eh^4 + 6\lambda\delta^2Eh^2 + 3\lambda\delta^4. \qquad \text{(A.4)}$$

So $M4^B \equiv \mu_4^B/(\mu_2^B)^2 - 3 = [3V(h^2) + (\lambda - \lambda^2)\{3\delta^4 + (6 - 12\lambda)\nu^2\delta^2 + (1 - 6\lambda)\nu^4\}]/((\lambda - \lambda^2)\nu^2 + Eh^2 + \lambda\delta^2)^2$, where $V(h^2)$ is the unconditional variance of h_t^2, which can be estimated by its sample analog.

[Received February 1992. Revised January 1993.]

REFERENCES

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

Baillie, R. T., and Bollerslev, T. (1989), "The Message in Daily Exchange Rates: A Conditional-Variance Tale," *Journal of Business & Economic Statistics*, 7, 297–305.

Ball, C., and Torous, W. (1983), "A Simplified Jump Process for Common Stock Returns," *Journal of Financial and Quantitative Analysis*, 18, 53–65.

——— (1985), "On Jumps in Stock Prices and Their Impact on Call Option Pricing," *Journal of Finance*, 40, 155–173.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31, 307–327.

——— (1988), "On the Correlation Structure of the Generalized Autoregressive Conditional Heteroscedastic Process," *Journal of Time Series Analysis*, 9, 121–131.

——— (1990), "Modeling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *The Review of Economics and Statistics*, 72, 498–505.

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

Broyden, G. (1965), "A Class of Methods for Solving Nonlinear Simultaneous Equations," *Mathematics of Computation*, 19, 577–593.

Diebold, F. X. (1988), *Empirical Modeling of Exchange Rate Dynamics* (Lecture Notes in Economics and Mathematical Systems, vol. 303), New York: Springer-Verlag.

Engle, R. (1982), "Autoregressive Conditional Heteroscedasticity With Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1007.