



ELSEVIER

Journal of Empirical Finance 1 (1994) 279–311

**Journal of
EMPIRICAL
FINANCE**

Neglected common factors in exchange rate volatility¹

Ronald Mahieu*, Peter Schotman

Limburg Institute of Financial Economics, University of Limburg, 6200 MD Maastricht, Netherlands

Final version accepted August 1993

Abstract

The paper proposes a new multivariate model for exchange rate volatility in a system of bilateral exchange rates, using a factor structure that is invariant with respect to the numeraire currency. In a complete system of exchange rates one of the common factors is always related to the numeraire currency. Time variation in the volatility is modelled using a stochastic variance approach. The interpretation of the factors provides a new way of estimating risk premia in the foreign exchange market. Empirical results show considerable volatility spillovers among the four major currencies. Risk premia show a major sign reversal for the dollar risk premium around 1978.

Key words: Stochastic volatility, Exchange rates; Factor models; Risk premia

JEL codes: C32; G15

1. Introduction

The empirical literature on floating exchange rates has largely concentrated on the behavior of the dollar against the major other currencies like the German mark, Japanese yen, and British pound. The cross rates have attracted much less

¹The first version of this paper was written while both authors were visitors at the Woodrow Wilson School of Public and International Affairs at Princeton University. Mahieu acknowledges financial support from NWO and SWOL. We would like to thank Richard Baillie (the editor) and Geert Almekinders, John Campbell, Herman van Dijk, John Geweke, Robert Hodrick, Franz Palm, Neil Shephard, and seminar participants at ESEM (Brussels 1992), Princeton University, Econometric Society Winter Meeting (London, 1993), Federal Reserve Bank of Minneapolis, CEPR Meetings on International Finance (Maastricht, 1993) for helpful comments. All errors are of course our own.

*Corresponding author.

attention. One of the stylized facts from the literature is that the time series of the logarithm of the dollar exchange rates is close to a random walk, meaning that almost all changes in the dollar can be interpreted as unpredictable news. Another stylized fact is the conditional heteroskedasticity in all time series of exchange rate changes.² In this paper we develop a new model for exchange rate volatility that simultaneously describes the volatility in all possible bilateral exchange rates between the major currencies and does not depend on a particular numeraire currency.

In the empirical model it is assumed that the change in the logarithm of any bilateral exchange rate is the difference of two country specific news terms. This model implies that the first differences of the exchange rates are positively correlated. For example, if the dollar rises or falls with respect to both the mark, the yen, as well as the pound, it is likely that there has been some important news about the U.S. economy. In general, the higher the correlation between the exchange rates of the dollar/yen and dollar/mark, the larger is the U.S. news component in daily or weekly exchange rate changes. The separate components are identifiable by exploiting the triangular identity which states that the difference between the log of the dollar/mark and the dollar/yen exchange rate yields the log of the yen/mark exchange rate. This enables us to perform a variance decomposition of exchange rate changes.

Univariate models of conditional heteroskedasticity are abundant, but relatively few studies use a multivariate framework. The main obstacle here is the large number of parameters involved in an unrestricted model for the time variation in volatility. For estimation a large number of usually ad hoc restrictions have to be imposed.³ There is still a quest for a convenient parameterization of multivariate volatility models, which can meet the empirical success of Bollerslev's (1986) univariate GARCH(1,1) model. Our decomposition of exchange rates in country specific news components provides a parsimonious parameterization of multivariate volatility dynamics.

The decomposition results are in principle applicable to various functional forms and stochastic specifications, e.g. Bollerslev's (1986) GARCH model or Nelson's (1991) EGARCH framework. However, in the empirical part of the paper we follow the ideas of the stochastic variance model of Harvey, Ruiz, and Shephard (1994). Their model also aims at a parsimonious parameterization, and also involves a factor structure, which makes their approach closely related to ours. One advantage of the stochastic volatility model is its flexibility in specifying the dynamics and in dealing with fat-tailed distributions. The requirement of numeraire invariance imposes some further structure on their model.

²See the surveys by Bollerslev, Chou and Kroner (1992) and by Nijman and Palm (1993) for references on this extensive literature

³See section 3 for details and references.

One motivation for the factor model that we adopt in the paper comes from the covariance structure that Domowitz and Hakkio (1985) used to derive their model of risk premia in the foreign exchange market. They showed that risk premia depend on the difference between the conditional variances of the country's money supplies, which are the only stochastic elements in their model. In our approach we make a factor structure assumption to identify these news components individually.

The paper is organized as follows. In section 2 we develop the variance decomposition and report empirical results for weekly exchange rate changes for the period 1973–1991 and several subperiods. Section 3 presents the generalization of the variance decomposition to models with time varying volatility. Section 4 reports empirical results for this model. In section 5 we look at the asset pricing implications. Finally, section 6 concludes.

2. The factor structure of exchange rate news

Consider a system of $n + 1$ currencies ($i = 0, \dots, n$), and express bilateral exchange rates with respect to the common numeraire currency 0. We assume that exchange rate changes are almost unpredictable and due to news. News in each country has two parts: a component related to worldwide shocks and a country specific component. This setup leads to the following model of exchange rate movements:

$$s_{i0} = u_i - u_0, \quad i = 1, \dots, n, \quad (1)$$

$$u_i = \sum_{k=1}^M \beta_{ik} v_k + e_i, \quad i = 0, \dots, n, \quad (2)$$

where s_{i0} is the change in the logarithm of the bilateral exchange rate of currency i in units of currency 0; u_i is the news originating from country i ; v_k is a worldwide common factor of news; β_{ik} is the sensitivity of news in country i with respect to worldwide shock v_k ; and e_i is the idiosyncratic news component. It is assumed that all factors have zero mean and are mutually uncorrelated. The variances of the common factors are normalized to one, while the variances of the country specific factors are $E(e_i^2) = \lambda_i$. The difference between Eq. (2) and a standard linear factor model is that the outputs u_i are only observed through the exchange rates s_{i0} . Combining (1) and (2) gives

$$s_{i0} = \sum_{k=1}^M (\beta_{ik} - \beta_{0k}) v_k + e_i - e_0 = \sum_{k=0}^M \gamma_{ik} v_k + e_i, \quad (3)$$

where $\gamma_{i0} = -1$, $v_0 = e_0$, and $\gamma_{ik} = (\beta_{ik} - \beta_{0k})$ for $k \geq 1$. Representation (3) explains the term 'neglected' in the title of the paper. The first common factor in exchange rates is the numeraire specific news e_0 . The common factors v_1, \dots, v_M

only affect exchange rates with loadings $(\beta_{ik} - \beta_{0k})$. Even if the common factors are important in the total currency news u_i , they might not have much effect on the exchange rate changes, when β_{ik} and β_{jk} are approximately equal.

To determine the number of common factors one would need a large number of currencies. In the empirical analysis we will limit ourselves to the four major international currencies. By concentrating on these four currencies we implicitly assume that $M = 4$. To estimate the model on just four currencies we will have to make some assumptions on the structure of the factor loadings β_{ik} . In particular we assume that (2) has the form

$$\begin{bmatrix} u_{\$} \\ u_{¥} \\ u_{M} \\ u_{£} \\ u_{\text{rest}} \end{bmatrix} = \begin{bmatrix} \beta_{01} & 0 & 0 & 0 \\ 0 & \beta_{12} & 0 & 0 \\ 0 & 0 & \beta_{23} & 0 \\ 0 & 0 & 0 & \beta_{34} \\ & & \mathbf{B}^* & \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} + \begin{bmatrix} e_{\$} \\ e_{¥} \\ e_{M} \\ e_{£} \\ e_{\text{rest}} \end{bmatrix}, \quad (4)$$

where the subscripts \$, ¥, M, £, denote the dollar, yen, mark and pound; \mathbf{u}_{rest} is the $(n - 3)$ vector of news of all other currencies, and \mathbf{e}_{rest} the specific news; the matrix of factor loadings is assumed to be diagonal in the major currencies, but \mathbf{B}^* is unrestricted. This entails just 2 overidentifying restrictions on a general four factor model. Under these assumptions the covariance matrix of the four major currencies with the dollar as numeraire takes the form

$$\Sigma = \begin{bmatrix} \lambda_0 + \lambda_1 & \lambda_0 & \lambda_0 \\ \lambda_0 & \lambda_0 + \lambda_2 & \lambda_0 \\ \lambda_0 & \lambda_0 & \lambda_0 + \lambda_3 \end{bmatrix}, \quad (5)$$

where we have redefined λ_1 as $\lambda_1 + \beta_{1,1+1}^2$. This model is equivalent to a zero factor model for just four currencies, and preserves the two overidentifying restrictions. The model implies that all covariances are equal and positive. For our weekly dataset of the four major currencies (dollar, yen, mark and pound) the observed sample covariances for the period 1973–1991 with the dollar as the numeraire are

$$\hat{\Sigma} = \begin{bmatrix} 2.09 & 1.31 & 1.06 \\ 1.31 & 2.23 & 1.49 \\ 1.06 & 1.49 & 2.10 \end{bmatrix}, \quad (6)$$

At first sight this covariance matrix is remarkably close to that implied by the zero factor model (5) with all λ_i equal, so that the model merits a closer statistical investigation.

The theoretical literature on foreign exchange risk premia has mostly considered the case $M = 0$. In Domowitz and Hakkio (1985) the e_i are uncorrelated unexpected shocks to each country's money supply. In a more general setting Hodrick (1989) develops an equilibrium model of the exchange rate which can also be written as the factor model with $M = 0$, and where e_i is a linear combination of news of each country's money growth, real output growth, government expenditures, and volatility innovations. In Hodrick's model each of the currency specific factors are assumed (conditionally) heteroskedastic.

We will estimate the variance λ_i for several subsamples of weekly exchange rate data. Since normality is always strongly rejected for exchange rate changes (see, e.g., Boothe and Glassman (1987)), we adopt a moment estimator that does not rely on normality. Let $y(t)$ be the $\frac{1}{2}n(n + 1)$ vector containing the squares of all possible bilateral exchange rate changes between $n + 1$ currencies, with $y_k(t) = s_{ij}^2(t)$ as the k^{th} element of $y(t)$. Let \mathbf{Z} be the $(\frac{1}{2}n(n + 1) \times (n + 1))$ matrix, with rows containing all permutations of two ones and $(n - 1)$ zeros in $(n + 1)$ positions. The ones in the k^{th} row of \mathbf{Z} are in the positions i and j and correspond to the squared bilateral exchange rate change $y_k(t) = s_{ij}^2(t)$. Since the variance of every bilateral exchange rate s_{ij} is modelled as the sum of λ_i and λ_j , we can formulate the linear model

$$y(t) = \mathbf{Z}\lambda + v(t), \tag{7}$$

where $v(t)$ is a vector of errors with mean zero. Omitting the time indices, the model for $n = 3$ is written as

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix}, \tag{8}$$

The parameter vector λ can be estimated consistently by OLS, pooling the time series and the cross section of six bilateral exchange rate changes. To construct an efficient estimator, we can use the initial OLS estimates to form the $(\frac{1}{2}n(n + 1) \times \frac{1}{2}n(n + 1))$ weighting matrix $\tilde{\mathbf{D}}$ with typical elements $\tilde{d}_{ij,km}$ as

$$\tilde{d}_{ij,km} = \frac{1}{T} \sum_{t=1}^T (s_{ij}^2(t) - \tilde{\lambda}_i - \tilde{\lambda}_j)(s_{km}^2(t) - \tilde{\lambda}_k - \tilde{\lambda}_m), \tag{9}$$

Applying SUR we find an efficient moment estimator for λ as

$$\hat{\lambda} = (\mathbf{Z}'\tilde{\mathbf{D}}^{-1}\mathbf{Z})^{-1}(\mathbf{Z}'\tilde{\mathbf{D}}^{-1}\bar{y}), \tag{10}$$

where \bar{y} is the (6×1) vector containing the second moments of the six bilateral exchange rates. The weighting implied by this estimator is inversely related to the fourth moment of the exchange rate changes. In computing the estimates we maintain the assumption that the variances exist. Standard errors are computed by the usual formula, $\mathbf{V}(\hat{\lambda}) = (\mathbf{Z}'\hat{\mathbf{D}}^{-1}\mathbf{Z})^{-1}/T$.

Given estimates of all the individual variances we can estimate time series of the country specific news. The observed n exchange rates impose n exact conditions on the $n + 1$ individual uncorrelated news components. Formally, we want the conditional mean $E[e(t)|s_0(t)]$, where $e(t)$ is the $n + 1$ news vector at time t , and $s_0(t)$ is the n -vector of exchange rate changes with numeraire currency 0. The GLS estimator for this conditional mean is given by

$$\hat{e}(t) = \Lambda P'(P\Lambda P')^{-1} s_0(t), \quad (11)$$

where $P = (-1|\mathbf{I})$ is a $(n \times (n + 1))$ matrix, and \mathbf{I} the n -dimensional identity matrix, and Λ is an $((n + 1) \times (n + 1))$ diagonal matrix with the variances of the specific factors on its diagonal. The specific elements of $\hat{e}(t)$ can also be written in the more explicit form

$$\hat{e}_i(t) = \left(\sum_{\substack{j=0 \\ j \neq i}}^n \lambda_j^{-1} s_{ij}(t) \right) \left/ \left(\sum_{j=0}^n \lambda_j^{-1} \right) \right. \quad (12)$$

The time series $\hat{e}_i(t)$ can be interpreted as the changes in the effective exchange rate of currency i .

Our dataset consists of the bilateral exchange rates among the dollar, yen, mark and pound for the period January 1973 to June 1991 and several sub-samples. The data are weekly Wednesday closing prices at the London market taken from DATASTREAM.⁴ In the empirical analysis we take all exchange rate changes in deviation of their sample mean.

The variance decomposition results are summarized in Tables 1 and 2. Table 1 reports the estimates of the variances of the currency specific news components. The variances of the specific currencies are significantly different from each other and are also different between subperiods. Over the full 1973–1991 sample the U.S. variance has contributed most to the volatility of the exchange rate system. The ranking of the variances is: $\lambda_{US} > \lambda_{JP} > \lambda_{UK} > \lambda_{GE}$, although the absolute differences between the currency specific variances are not very large. The dominance of the U.S. variance is especially due to the later part of the sample period. Over the last five years (87–91) the U.S. variance is four

⁴If the market was closed on some Wednesday, we choose the Tuesday closing price. If Tuesday was a holiday too, we took the Thursday price.

Table 1
 Variance decomposition of exchange rate news. $\text{Var}(s_{i,t}) = \lambda_i + \lambda_j$

Period	λ_S	λ_Y	λ_M	λ_i	EQUAL	FIT
Jan 73–Jun 91 (Full sample)	1.15 (0.10)	0.93 (0.09)	0.67 (0.07)	0.70 (0.08)	18.9	43.6
Jan 73–Dec 76	0.45 (0.11)	0.75 (0.19)	1.08 (0.23)	0.74 (0.19)	8.17	7.05
Jan 77–Dec 80	0.76 (0.12)	1.40 (0.21)	0.54 (0.11)	0.54 (0.11)	17.6	6.66
Jan 81–Sep 85	1.17 (0.13)	0.69 (0.14)	0.62 (0.13)	0.84 (0.18)	10.9	26.2
Sep 85–Feb 87 (Plaza–Louvre)	1.32 (0.59)	0.43 (0.37)	0.73 (0.20)	0.89 (0.32)	1.84	12.4
Feb 87–Jun 91	1.77 (0.20)	0.94 (0.14)	0.40 (0.07)	0.44 (0.09)	50.3	5.86

Notes: λ_i (i = dollar (\$), yen (¥), mark (M), pound (£)) denotes the exchange rate variance due to country i . The estimates are from the second round of the moment estimator (10). Standard errors in parentheses. The column EQUAL is a Wald test for equality of all 4 λ_i 's, asymptotically distributed as $\chi^2(3)$. The column FIT is a Wald test for the overidentifying moment restrictions implied by the factor structure, asymptotically distributed as $\chi^2(2)$. Exchange rate changes are measured in percentage per week and are corrected with the sample mean.

times as big as the volatility originating in either the UK or Germany, while in the first years of the floating exchange rate period (73–76) most of the variance was due to events in West Germany. For the period 77–80 Japan had by far the largest variance. The fourth moments in Table 2 generally imply excess kurtosis and rejection of normality, as is common for exchange rate changes. The period between September 1985 and February 1987 (Plaza – Louvre) has been the most volatile episode of the last 20 years. The estimates are considerably above their full sample averages. At the same time the three dollar exchange rates and the yen/pound rate have much higher fourth moments than in other subperiods (see Table 2).

Returning to the results in Table 1 we find that the two overidentifying conditions implied by the factor structure (5) are strongly rejected for the full sample and also for the two subperiods between 1981 and February 1987. The test statistics do, however, not take into account any further heteroskedasticity within the subperiods. As is also suggested by Table 2 the model appears to fit well in the seventies and again after 1987.

Table 3 reports a set of diagnostics of the extracted news components using the full sample parameter estimates. The non-normality is not confined to one

Table 2
Second and fourth moments

	¥/\$	DM/\$	¥/\$	DM/¥	£/¥	£/DM
Jan 73–Jun 91 (Full sample)						
Sample	2.09	2.23	2.10	1.70	2.07	1.35
Estimated	2.08	1.83	1.86	1.60	1.63	1.37
Fourth	33.4	17.9	20.5	11.3	22.5	9.06
Jan 73–Dec 76						
Sample	1.39	1.88	1.17	1.81	1.70	1.86
Estimated	1.20	1.53	1.19	1.83	1.49	1.82
Fourth	95.6	21.5	8.87	20.6	30.1	18.0
Jan 77–Dec 80						
Sample	2.31	1.44	1.21	1.96	2.29	1.13
Estimated	2.16	1.29	1.30	1.94	1.94	1.07
Fourth	14.2	6.55	5.13	10.2	16.1	4.38
Jan 81–Sep 85						
Sample	1.86	2.61	2.99	1.54	2.41	1.41
Estimated	1.86	1.78	2.00	1.32	1.53	1.46
Fourth	10.9	16.5	38.3	8.14	23.4	9.31
Sep 85–Feb 87 (Plaza–Louvre)						
Sample	2.92	3.19	2.72	1.87	3.00	1.74
Estimated	1.75	2.04	2.21	1.15	1.32	1.62
Fourth	33.3	45.6	47.7	11.7	44.0	7.18
Feb 87–Jun 91						
Sample	2.48	2.44	2.52	1.50	1.54	0.81
Estimated	2.71	2.18	2.22	1.34	1.38	0.84
Fourth	16.2	12.5	15.5	6.14	9.71	2.34

Notes: 'Sample' is the uncentered second moment of exchange rate changes; 'Estimated' are the fitted second moments using the optimal moment estimator; 'Fourth' is the centered fourth moment required in the weighting matrix \mathbf{D} , i.e. $E\{(X^2 - E\{X^2\})^2\}$. Exchange rate changes are measured in percentage per week and are corrected with the sample mean.

particular currency, but appears in all four series. There is hardly any autocorrelation in the news series, except for a slightly significant Ljung–Box statistic for the Japanese news series. In contrast, there is strong evidence of ARCH in all four components, indicating that the ARCH behaviour is not special to the U.S. dollar. More interesting is the finding that there are strong heteroskedasticity spillovers. The Granger causality tests indicate that the Japanese and British squared news components are predictable by the other countries' lagged squared news.

Table 3
Diagnostics of news components

	e_{US}	e_{JP}	e_{GE}	e_{UK}
Skewness	0.52	− 0.88	− 0.61	0.64
Kurtosis	2.36	4.24	2.67	3.33
Normality	268.4*	849.1*	347.6*	511.8
Autocorrelation	0.057	0.031	0.100	0.066
Ljung-Box (30)	36.04	37.74	51.65*	40.30
ARCH (1)	4.66*	7.32*	8.69*	17.10*
ARCH (10)	47.52*	43.85*	86.09*	37.71*
Cross ARCH	2.68	37.15*	0.59	20.79*
Causality	0.90	16.62*	0.20	8.77*

Notes: 'Skewness' is the scaled third moment of the news series, $\Sigma e^3/(\Sigma e^2)^{3/2}$; 'Kurtosis' is the excess kurtosis ($\Sigma e^4/(\Sigma e^2)^2 - 3$); 'normality' is the Jarque-Bera test for normality, distributed as $\chi^2(2)$; 'Autocorrelation' is the first order autocorrelation; 'Ljung-Box' is a test for autocorrelation up to order 30, distributed as $\chi^2(30)$; ARCH(1) is the LM test for first order ARCH, distributed as $\chi^2(1)$; ARCH(10) is the LM test for 10th order ARCH ($\chi^2(10)$); 'Cross ARCH' is TR^2 of the regression of $\hat{e}_i^2(t)$ on a constant and all four squared news series with one lag ($\chi^2(4)$); 'Causality' is the F-statistic for the significance of the cross squared news series in the last cross ARCH regression. An asterisk (*) denotes significance at the 5% level.

3. Multivariate time varying volatility

3.1. Specification problems

A multitude of specifications exist for modelling exchange rate volatility in univariate models. Most of these models are a variant of the ARCH model developed by Engle (1982). For exchange rates the empirical evidence favours a specification with fat-tailed errors even after correcting for the conditional heteroskedasticity.⁵ The main problem in specifying a multivariate model is the number of parameters that is of the order n^4 when the univariate models are straightforwardly generalized to the multivariate framework. The factor structure investigated in the previous section will be used to specify a new, parsimonious parameterization of volatility dynamics. It turns out that such a specification is most easily achieved in the stochastic volatility framework proposed by Taylor (1986), and Harvey, Ruiz and Shephard (1994).

Most of the multivariate ARCH models that have been developed introduce ad hoc restrictions on the number of parameters, and are mostly inadmissible in the application to exchange rates. An example is the constant conditional

⁵See Engle and Bollerslev (1986), Baillie and Bollerslev (1989) and Hsieh (1989).

correlation assumption of Baillie and Bollerslev (1990). Consider the bivariate case with two exchange rates (e.g. dollar/yen and dollar/mark):

$$\Sigma_0(t) = \begin{pmatrix} \sigma_1^2(t) & \rho\sigma_1(t)\sigma_2(t) \\ \rho\sigma_1(t)\sigma_2(t) & \sigma_2^2(t) \end{pmatrix}, \quad (13)$$

with σ_1^2 and σ_2^2 the variance of the dollar/yen and dollar/mark rate respectively. By changing the numeraire from the dollar to the yen we obtain the transformed 'yen' covariance matrix

$$\Sigma_1(t) = \begin{pmatrix} \sigma_1^2(t) & \sigma_1^2(t) - \rho\sigma_1(t)\sigma_2(t) \\ \sigma_1^2(t) - \rho\sigma_1(t)\sigma_2(t) & \sigma_1^2(t) + \sigma_2^2(t) - 2\rho\sigma_1(t)\sigma_2(t) \end{pmatrix}, \quad (14)$$

which no longer has the constant conditional correlation property. The same problem arises in the diagonal ARCH model of Bollerslev, Engle and Wooldridge (1988). The factor model of Diebold and Nerlove (1989) is also currency specific. Their model reads

$$s_{i0}(t) = \gamma_i v(t) + e_i(t), \quad (15)$$

where γ_i ($i = 1, \dots, n$) are factor loadings on the single factor $v(t)$, and all exchange rates are expressed with the dollar as numeraire. Diebold and Nerlove (1989) assume that all time varying volatility is due to the common factor, i.e. the dollar numeraire effect. The diagnostics in Table 3 indicated however, that all currency factors exhibit conditional heteroskedasticity. We would therefore still need a multivariate ARCH model for $v(t)$ and all $e_i(t)$.

The FACTOR-ARCH model of Engle, Ng and Rothschild (1990) is also not directly applicable to a system of exchange rates. In our zero factor exchange rate model (7) we have $s_0(t) = \mathbf{P}e(t)$. The exchange rate model obviously has constant factor loadings, since the matrix $\mathbf{P} = (-\mathbf{1}|\mathbf{I})$ is completely known. But since all currency specific factors can be (and probably are) heteroskedastic, the number of factors ($n + 1$) is larger than the number of elements (n) in the vector time series $s_0(t)$. Specification of the GARCH structure for the factor variances is not trivial in that case, especially since the diagnostics in Table 3 also indicated the cross effects from the volatility of one factor to the volatility of all other factors.

We therefore opted to apply the multivariate stochastic volatility model used by Harvey, Ruiz and Shephard (1994, HRS), and respecify it to fit into the covariance structure of section 2. Since the properties of the stochastic volatility (SV) models are not as well developed as the properties of ARCH models, subsection 3.2 below provides a brief overview of the specification and estimation of a univariate SV model. The subsection also discusses the relation between SV and ARCH. After this digression we return to the multivariate setting in subsection 3.3.

3.2. Univariate stochastic volatility models

The univariate SV model can be written as (see Taylor (1986) and HRS)

$$s(t) = \varepsilon(t)\sigma(t) = \varepsilon(t) \exp\left\{\frac{1}{2}h(t)\right\}, \tag{16}$$

where $\varepsilon(t)$ is i.i.d with mean zero and unit variance, and where $h(t)$ is the log of the variance of some bilateral exchange rate change $s(t)$, which in the univariate case is assumed to be generated by an AR(1) process,

$$h(t) - \mu = \rho(h(t - 1) - \mu) + \eta(t), \eta(t) \sim NID(0, \omega^2), \tag{17}$$

with μ the unconditional mean of $h(t)$. Squaring (16) and taking logarithms gives

$$w(t) \equiv \ln\{s(t)^2\} = h(t) + \ln\{\varepsilon(t)^2\} = h(t) + \alpha + \zeta(t), \tag{18}$$

where $E[\zeta(t)] = 0$. HRS assume that $\varepsilon(t)$ is Gaussian, for which case they note that $\alpha = -1.27$ and $E[\zeta(t)^2] = \pi^2/2$. For $\varepsilon(t)$ and $\eta(t)$ bivariate normal with some unknown correlation, HRS show that the transformed error term $\zeta(t)$ is always uncorrelated with $\eta(t)$. The system (17) and (18) defines a standard state space model, apart from the (possible) non-normality of $\zeta(t)$. HRS suggest to ignore this non-normality and estimate the system by quasi maximum likelihood (QMLE). The steady state Kalman filter recursions then provide an expression for the conditional (log)-variance of $s(t)$,

$$\hat{h}(t + 1|t) = \mu(1 - \rho) + \frac{2\rho P^2}{2P^2 + \pi^2} w(t) + \frac{2\rho\pi^2}{2P^2 + \pi^2} \hat{h}(t|t - 1) \tag{19}$$

where P^2 is the solution to the Riccati equation $P^2 = \rho^2(2\pi^{-2} + P^{-2}) + \omega^2$. Using the properties of the log-normal distribution the conditional variance of the exchange rate innovation $s(t)$ is given by

$$\hat{\sigma}(t + 1|t)^2 = \exp\left(\frac{1}{2}\hat{h}(t + 1|t) + \frac{1}{8}P^2\right) \tag{20}$$

which establishes the relation between the SV process (17), (18) and an equivalent exponential ARCH process.⁶ Estimation of the two processes differs, however. The QMLE of the SV model involves maximization of the objective function

$$F_1 = -\frac{T}{2} \ln \psi^2 - \frac{1}{2\psi^2} \sum_{t=1}^T (w(t) - \hat{h}(t|t - 1))^2, \tag{21}$$

⁶See Andersen (1992) for a general discussion on the relation between ARCH and stochastic volatility.

where $\psi^2 \equiv P^2 + \frac{1}{2}\pi^2$ is the steady state innovation variance of $w(t) = \ln\{s(t)^2\}$. In contrast, an ARCH model would be estimated by maximizing the objective function

$$F_2 = -\frac{1}{2} \sum_{t=1}^T \left(\ln \hat{\sigma}(t|t-1)^2 + \frac{s(t)^2}{\hat{\sigma}(t|t-1)^2} \right) \quad (22)$$

The difference between the two criterion functions stems from the assumption on which innovation is taken to be Gaussian. With ARCH, $s(t)$ is assumed conditionally normal with zero mean and time-varying conditional variance, while for the QMLE of the SV model the logarithm of the squared exchange rate $w(t)$ is assumed Gaussian with time varying mean and constant variance. Otherwise the two models are equivalent. The differences in interpretation between the models are discussed in Andersen (1992).

Normality of $\varepsilon(t)$ plays an important role in the model, since it determines the mean and variance of $\ln\{\varepsilon(t)^2\}$, and also leads to a very skewed distribution of $\xi(t)$, see Fig. 1. Ruiz (1992) compares the QMLE with a method of moments estimator and concludes that the QMLE has better relative asymptotic efficiency. Jacquier, Polson and Rossi (1993) compare both these estimators with the exact maximum likelihood estimator, i.e. using a log chi-squared density for $\xi(t)$. They conclude that the exact maximum likelihood estimator is far superior to the QMLE, both with respect to bias as well as variance.

These results depend heavily on the normality of $\varepsilon(t)$, and the implied skewness of $\xi(t)$. But the skewness of $\xi(t)$ is an empirical matter, just like the normality of $\varepsilon(t)$. In the empirical ARCH literature Engle and Bollerslev (1986) find that the standardized exchange rate innovations are still leptokurtic, which leads them, and Baillie and Bollerslev (1991), to consider the Student-t distribution as an alternative. If $\varepsilon(t)$ is fat-tailed the implied distribution of $\ln(\varepsilon(t)^2)$ becomes less skewed than the log chi-squared. In the extreme case that $\varepsilon(t)$ has a Cauchy distribution, the implied density of $\xi(t)$ is the symmetric function

$$p(\xi) = \frac{e^{\xi/2}}{\pi(1 + e^{\xi})}. \quad (23)$$

In this case the constant term α in (18) equals zero, while $\text{Var}[\xi(t)] = \pi^2$. The distribution $p(\xi)$ is plotted in Fig. 1b together with the normal and the log chi-squared. Given the symmetry and exponential tails of $p(\xi)$, we would expect that the QMLE performs much better in this case. Fig. 1a also shows the case that $\xi(t)$ is normally distributed, so that $\varepsilon(t)$ has a lognormal distribution. Gaussianity of $\xi(t)$ implies that very small innovations in $\varepsilon(t)$ are relatively unlikely, see the dip in the plotted density.

Finally, the constant terms α and μ are not separately identifiable as free parameters. We therefore reparameterize the system as

$$\begin{aligned} w(t) &= x(t) + \zeta + \xi(t) \\ x(t) &= \rho x(t-1) + \eta(t), \end{aligned} \quad (24)$$

Figure 1A. Density of $\varepsilon(t)$

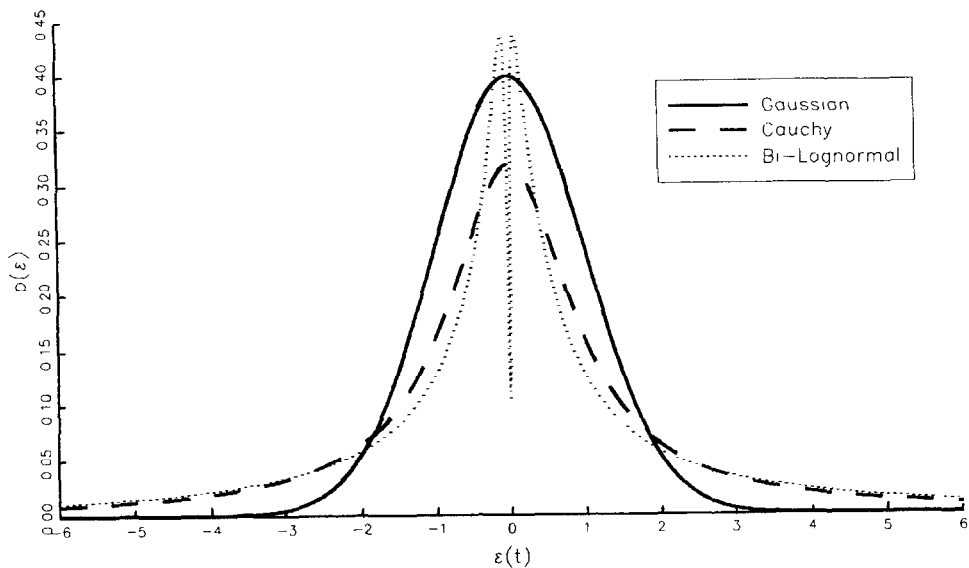


Figure 1B: Density of $\xi(t)$

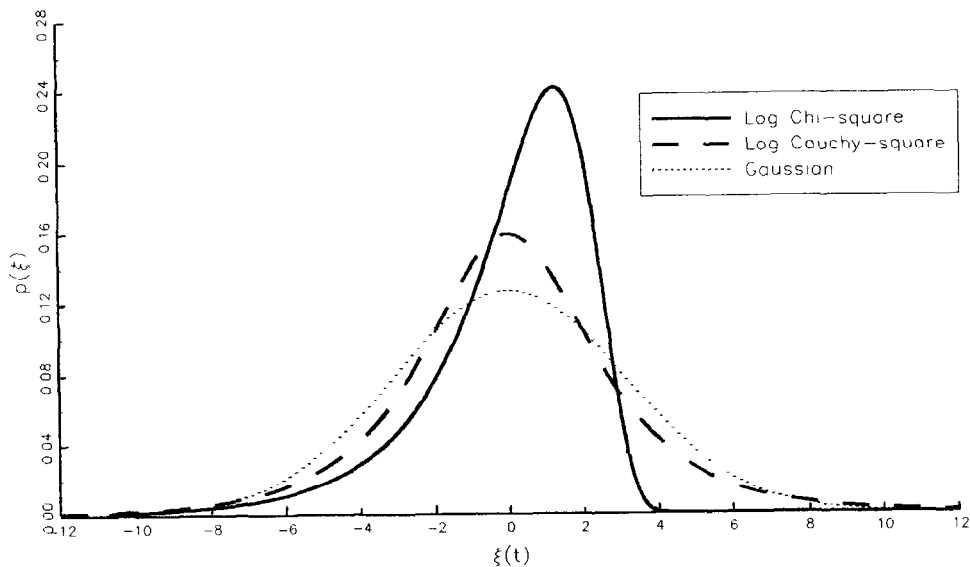


Fig. 1. Densities of exchange rate innovations (A) and of the implied errors in the measurement equation of the volatility model (B).

where $\zeta = \alpha + \mu$, and the state vector $x(t)$ is defined as $h(t) - \mu$. A possible way to identify α is through the estimated variance of $\zeta(t)$. For example, if $\phi^2 = \pi^2/2$, $\zeta(t)$ might be generated by the log chi-squared, in which case $\alpha = -1.27$. Alternatively, if $\phi^2 = \pi^2$, and the residuals $\zeta(t)$ are symmetric, we might be dealing with the Cauchy distribution for $\varepsilon(t)$, so that $\alpha = 0$. In this case the conditional variances of exchange rates do not exist, but there is still a meaningful way to describe time-varying volatility through the $h(t)$ process, i.e. $\sigma(t) = e^{h(t)/2}$ still functions as a scale parameter.

3.3. A multivariate stochastic volatility model for exchange rates

In our model $w(t)$ is a vector of length $\frac{1}{2}n(n+1)$ with typical element $w_k(t) = \ln\{y_k(t)\} = \ln\{s_{ij}^2(t)\}$, ($i = 0, \dots, n-1$; $j = i+1, \dots, n$), i.e. we use all possible bilateral exchange rates, see section 2. According to the factor model (7) the variance of any bilateral exchange rate s_{ij} is the sum of two currency specific variances. This implies that

$$\exp(h_{ij}(t)) = \lambda_i(t) + \lambda_j(t) = \exp(h_i(t)) + \exp(h_j(t)), \quad (25)$$

where $h_i(t)$ measures the specific volatility of currency i . Linearizing (25) around some \bar{h}_i and \bar{h}_j gives

$$h_{ij}(t) = \ln\{\exp(\bar{h}_i) + \exp(\bar{h}_j)\} + \frac{\exp(\bar{h}_i)(h_i(t) - \bar{h}_i) + \exp(\bar{h}_j)(h_j(t) - \bar{h}_j)}{\exp(\bar{h}_i) + \exp(\bar{h}_j)}. \quad (26)$$

If we further assume a common point of linearization $\bar{h}_i = \bar{h}_j$ the linear approximation reduces to

$$h_{ij}(t) = \ln(2) + \frac{1}{2}(h_i(t) + h_j(t)), \quad (27)$$

Eq. (27) shows that every bilateral exchange rate has two factors that define its volatility. In a system with $n+1$ currencies we thus have $n+1$ common volatility factors, which leads to the following structure for the measurement equations of the state space model

$$w(t) = \frac{1}{2} \mathbf{Z}h(t) + (\alpha + \ln(2))\mathbf{1} + \xi(t), \quad \text{Var}[\xi(t)] = \Phi, \quad (28)$$

where \mathbf{Z} has been defined in (7) and (8). To reduce the number of parameters in the estimation procedure we assumed that all exchange rates have the same type of distribution, implying that the diagonal elements of Φ are equal to ϕ^2 . Furthermore, we specified the correlation structure of $\xi(t)$ as

$$\begin{aligned} \text{corr}(\xi_{ij}, \xi_{kl}) &= r \quad \text{if } i = (k \text{ or } l), \text{ or } j = (k \text{ or } l) \\ &= 0 \quad \text{otherwise.} \end{aligned} \quad (29)$$

In this case Φ contains only two unknown parameters, and is found to be

$$\Phi = \phi^2 \begin{bmatrix} 1 & r & r & r & r & 0 \\ r & 1 & r & r & 0 & r \\ r & r & 1 & 0 & r & r \\ r & r & 0 & 1 & r & r \\ r & 0 & r & r & 1 & r \\ 0 & r & r & r & r & 1 \end{bmatrix}, \tag{30}$$

If ϑ is the correlation between two elements of $\varepsilon(t)$, then HRS derive the correlation between the corresponding elements of $\xi(t)$ as,

$$r = \frac{2}{\pi^2} \sum_{n=1}^{\infty} \frac{(n-1)! \Gamma(\frac{1}{2})}{\Gamma(n + \frac{1}{2})} g^{2n}. \tag{31}$$

The unconditional correlation between ε_{ij} and ε_{kl} is equal to 0.5 under the assumption of equal factor variances. This implies that we expect r to be approximately equal to 0.11.

The state transition equation for a multivariate process is the first order vector autoregression

$$h(t) = (\mathbf{I} - \mathbf{A})\mu + \mathbf{A}h(t - 1) + \eta(t), \tag{32}$$

where \mathbf{A} is a $((n + 1) \times (n + 1))$ matrix; μ is an $(n + 1)$ vector of constants; and where $E[\eta(t)\eta(t)'] = \Omega$. As in the univariate model, the constant terms μ and α are not separately identified. We therefore reparameterize the system as

$$\begin{aligned} w(t) &= \frac{1}{2} \mathbf{Z}(x(t) + \zeta) + \xi(t), \\ x(t) &= \mathbf{A}x(t - 1) + \eta(t), \end{aligned} \tag{33}$$

where $\zeta = \mu + \tilde{v}(\alpha + \ln(2))$, with \tilde{v} an $(n + 1)$ vector of ones, and where $x(t) = h(t) - \mu$. Without restrictions on \mathbf{A} or ζ the total number of parameters is 32 in our application with the four major currencies.

Our model differs from the factor model of HRS, who define $w(t)$ as a vector of logs of squared exchange rate changes against a common numeraire currency. Their typical element $w_j(t) = \ln\{s_{0j}^2(t)\}$, currency 0 being the common numeraire. The problem with this specification is that it is numeraire dependent. In squaring the exchange rates against a single numeraire one loses valuable information about the covariances. In order to construct a model for the whole system of exchange rates we need to augment the vector $w(t)$ to include the log of the squares of all possible bilateral exchange rate changes, as described above and in accordance with the covariance structure investigated in section 2.

Using specification (33), the Kalman filter prediction equation provides a recursive formula for the conditional log variance $x(t)$ (up to a constant). Analogous to the univariate case the implied EGARCH type specification reads

$$\begin{aligned}\hat{x}(t+1|t) &= 2\mathbf{A}\bar{\mathbf{P}}\mathbf{Z}'(\mathbf{Z}\bar{\mathbf{P}}\mathbf{Z}' + 4\Phi)^{-1}w(t) \\ &\quad + \mathbf{A}(\mathbf{I} - 2\bar{\mathbf{P}}\mathbf{Z}'(\mathbf{Z}\bar{\mathbf{P}}\mathbf{Z}' + 4\Phi)^{-1}\mathbf{Z})\hat{x}(t|t-1) \\ &= \mathbf{B}w(t) + \mathbf{C}\hat{x}(t|t-1),\end{aligned}\tag{34}$$

where $\bar{\mathbf{P}}$ is the solution to an algebraic Riccati equation. The conditional heteroskedasticity formulation also involves all squares and cross products of the log squared exchange rate changes. However, the number of free parameters in \mathbf{B} and \mathbf{C} is restricted.

4. Stochastic volatility results

The data for the empirical results in this section is $w(t) = \ln\{y(t)\} = \ln\{(\Delta\ln\{S(t)\} - \delta)^2\}$, where δ is the sample mean of $\Delta\ln\{S(t)\}$. A set of descriptive statistics for the six bilateral exchange rates is reported in Table 4. The transformed series show considerable autocorrelation; especially the autocorrelations of the dollar/yen and dollar/pound exchange rates are very persistent and still sizable after 30 lags. The differences between these two autocorrelation patterns and the other four give an indication that one time-varying factor is not enough to describe the dynamics of the currencies in our sample. Furthermore, for all exchange rates the $w_{ij}(t)$ are negatively skewed. This is an indication that the exchange rates are not as heavily fat-tailed as implied by the Cauchy distribution. The negative outliers in the data are due to many small changes in the exchange rates. The lower part of Table 4 shows parameter estimates for six univariate GARCH(1,1) models. Two exchange rates (dollar/yen and dollar/mark) turn out to be almost IGARCH, while the other four are much less persistent. The clear differences in the dynamics provides additional evidence that there are more factors than just the dollar.

The system (33) has been estimated by numerical optimization of the quasi-likelihood function obtained by assuming normality for $\zeta(t)$ and $\eta(t)$. Parameter estimates are in Table 5; diagnostics in Table 6. The estimates of ϕ and the residual characteristics provide information about the type of distribution for $\varepsilon(t)$ and $\zeta(t) = \ln\{\varepsilon(t)^2\}$. The error variance in the measurement equations is estimated very precisely and almost equal to π^2 , which is very close to the value implied by a Cauchy distribution for exchange rate innovations, and much larger than what is implied by the log chi-squared distribution. The prediction errors are less negatively skew than is implied by the log chi-squared. The results seem to indicate that a Student-t distribution with low degrees of freedom might be a good choice for $\varepsilon(t)$. The dynamic specification does not completely describe

Table 4
Summary statistics of univariate exchange rate changes

$w_{i,j}$	¥/\$	DM/\$	£/\$	DM/¥	¥/£	£/DM
Mean	-1.25	-0.79	-1.08	-1.01	-1.06	-1.43
Std. dev	2.73	2.42	2.58	2.32	2.52	2.37
Skewness	-1.14	-1.14	-1.03	-1.21	-1.09	-0.86
Kurtosis	2.04	1.60	1.20	2.31	1.48	0.93
Normality	376.77*	310.70*	226.41*	448.79*	278.39*	152.51*
Minimum	-15.11	-12.00	-14.33	-13.25	-10.49	-11.44
Maximum	4.87	4.13	4.04	3.61	4.19	3.61
Ljung-Box (30)	510.37*	127.81*	459.97*	86.00*	93.57*	59.91*
Autocorrelations						
1	0.20	0.09	0.17	0.08	0.07	0.12
2	0.15	0.12	0.19	0.06	0.10	0.10
3	0.16	0.10	0.16	0.11	0.07	0.09
4	0.15	0.07	0.20	0.03	0.08	0.07
5	0.14	0.10	0.14	0.07	0.03	0.02
10	0.11	0.04	0.15	-0.01	0.04	0.07
15	0.12	0.03	0.14	-0.02	0.05	0.03
20	0.10	0.05	0.08	0.09	0.04	-0.01
25	0.11	0.02	0.08	0.03	0.07	-0.00
30	0.09	0.03	0.06	0.00	-0.04	0.02
GARCH(1,1) $s_{i,j}(t) = \delta + v_{i,j}(t)$;						
$E_{t-1}[v_{i,j}(t)^2] = h(t) = \alpha_0 + \alpha_1 h(t-1) + \alpha_2 v_{i,j}^2(t-1)$						
δ	-0.075 (0.04)	-0.082 (0.04)	0.037 (0.04)	0.027 (0.04)	0.131 (0.05)	0.075 (0.04)
α_0	0.009 (0.01)	0.020 (0.02)	0.085 (0.04)	0.118 (0.18)	0.103 (0.10)	0.172 (0.44)
α_1	0.044 (0.01)	0.069 (0.02)	0.079 (0.02)	0.179 (0.14)	0.059 (0.04)	0.101 (0.13)
α_2	0.952 (0.01)	0.924 (0.03)	0.881 (0.02)	0.763 (0.22)	0.889 (0.07)	0.774 (0.44)

Notes: Variables are defined as $w_{i,j}(t) = \ln\{[\Delta \ln\{S_{i,j}(t)\} - \text{mean}]^2\}$. See Table 3 for definitions of statistics. Standard errors are between parentheses for GARCH(1,1) parameters.

the mark volatility, as the diagnostics for the prediction errors involving the mark imply significant residual autocorrelation.

The correlation parameter r was estimated freely but corresponds to the theoretical value ($r \approx 0.11$) derived from the formulas in HRS, supporting the approximation $\bar{h}_i = \bar{h}_j$ that we used in the linearization of (26).

The structure of the error covariance matrix Ω of the transition equations implies that the innovations of all four variance components are highly positively correlated. Given our use of weekly data, it means that an increase in the

Table 5
Parameter estimates of stochastic volatility model

	dollar	yen	mark	pound
$w(t) = \frac{1}{2} \mathbf{Z}(x(t) + \zeta) + \zeta(t) \quad \text{Var}[\zeta(t)] = \Phi$ $x(t+1) = \mathbf{A}x(t) + \eta(t) \quad \text{Var}[\eta(t)] = \Omega$				
ζ_1	-1.018 (0.47)	-1.152 (0.29)	-0.100 (0.18)	-1.384 (0.25)
$a_{\$j}$	0.820 (0.10)	-0.266 (0.14)	-0.871 (0.31)	0.109 (0.14)
$a_{¥j}$	0.122 (0.11)	0.541 (0.15)	-1.019 (0.33)	0.146 (0.17)
$a_{\text{M}j}$	-0.080 (0.05)	-0.146 (0.06)	0.496 (0.14)	0.032 (0.07)
$a_{\text{£}j}$	-0.062 (0.06)	-0.143 (0.10)	-0.449 (0.25)	0.980 (0.08)
Volatility innovation covariance matrix $\Omega = \text{Var}[\eta(t)]$				
dollar	1.183	0.957	0.510	0.570
yen	<i>0.783</i>	1.263	0.446	0.569
mark	<i>0.842</i>	<i>0.713</i>	0.310	0.361
pound	<i>0.736</i>	<i>0.711</i>	<i>0.913</i>	0.507
Unconditional variance $\bar{\Omega} = \mathbf{A}\bar{\Omega}\mathbf{A}' + \Omega$				
dollar	4.781	2.374	-0.458	1.262
yen	<i>0.627</i>	2.999	-0.339	1.010
mark	<i>-0.245</i>	<i>-0.229</i>	0.728	0.239
pound	<i>0.437</i>	<i>0.442</i>	<i>0.212</i>	1.742
Standard deviation of $\zeta(t)$: $\phi = 3.148$ (0.037)				
Correlation $\zeta_1(t)\zeta_j(t)$ $r = 0.130$ (0.012)				
Roots of system (eigenvalues of \mathbf{A}):				
	0.979	0.929	0.875	0.054

Notes: Exchange rates are expressed in percentages per week. Standard errors are in parentheses. *Italics* in the lower triangle of a covariance matrix denote correlations. All eigenvalues of \mathbf{A} are real.

volatility of one currency gets transmitted to increased volatility in all other currencies within a week. It also means that, for example, an increase in the volatility of dollar exchange rates leads to increased volatility in the mark/pound cross rate. Although the weekly innovations are highly correlated, in the long run the four variance components behave very differently (see Table 5). High volatility of the mark tended to go together with low volatility of

Table 6
Diagnostics of stochastic volatility model

$$\xi(t) = w(t) - \frac{1}{2} \mathbf{Z}(\hat{x}(t|t-1) + \zeta)$$

ξ_{ij}	¥/\$	DM/\$	£/\$	DM/¥	¥/£	£/DM
Skewness	-1.43	-0.94	-0.94	-1.44	-1.08	-1.44
Kurtosis	3.17	1.72	1.91	3.63	1.86	1.17
Normality	729.9*	261.4*	289.0*	860.2*	326.4*	448.7*
Ljung-Box(10)	8.67	20.55*	17.82	19.62*	5.40	27.77*
Ljung-Box(20)	16.38	37.25*	30.26	28.69	15.54	48.29*
ARCH(1)	0.34	5.57*	0.77	0.17	0.40	1.02

This table provides diagnostics of the Kalman filter prediction errors. See table 3 for explanatory notes

the dollar and yen. This is in close agreement with the stylized facts for different subperiods in Table 1.

The differences between the innovations covariance structure Ω and the unconditional covariance matrix are caused by the large negative estimates for some elements in the transition matrix \mathbf{A} . Despite the negative off-diagonal elements the system has three large eigenvalues (see Table 5) that are close to unity, indicating that the volatility series $h(t)$ might be integrated. Further analysis of the dynamic implications is reported in a set of causality tests in Table 7. There seems to be no lagged relations from the dollar volatility to volatility in any of the other currencies: dollar news is transmitted within a week. The yen and mark, however, strongly influence all the other currencies. Especially the large negative elements of the mark column in \mathbf{A} are noticeable (see Table 5).

The causality pattern is consistent with the results of Engle, Ito and Lin (1990), who used a dataset with four observations a day for the dollar/yen exchange rate. They can identify the separate country specific news because they have observations on the opening and closing prices of different markets. Strong volatility spillovers were found from Japan to the U.S., and vice versa. Their results seem to suggest that 'the Tokyo news has a greater impact on the volatility spillovers' and 'the volatility in the Tokyo market (...) had a great impact on the world volatility' (Engle, Ito and Lin (1990, p. 535, 538)). Both these facts are related to our finding of a strong lagged effect from yen volatility to volatility in the other currencies. Interestingly, we find the same effect for the mark, but not for the pound and dollar.

Figs. 2 and 3 contains time series plots of the series for $\hat{x}(t|t-1)$ and $\hat{x}(t|T)$, respectively. These series are obtained from the standard Kalman prediction and smoother recursions. The conditional log volatility $\hat{x}(t|t-1)$ is used in the EGARCH representation of the stochastic volatility model. It is much smoother

Table 7
Stochastic volatility tests

$$w(t) = \frac{1}{2} \mathbf{Z}(h(t) + \zeta) + \zeta(t)$$

$$x(t) = \mathbf{A}x(t-1) + \eta(t)$$

Hypothesis	Wald	df
(1) Equal unconditional variances: $\zeta_1 = \zeta_2 = \zeta_3 = \zeta_4$	2.78	3
(2) Diagonal dynamics: $a_{ij} = 0$ ($i \neq j$)	69.02*	12
(3) Causality (columnwise tests)		
dollar to other currencies: $a_{21} = a_{31} = a_{41} = 0$	1.59	3
yen to other currencies: $a_{12} = a_{32} = a_{42} = 0$	13.38*	3
mark to other currencies: $a_{13} = a_{23} = a_{43} = 0$	14.17*	3
pound to other currencies: $a_{14} = a_{24} = a_{34} = 0$	2.27	3
(4) Causality (rowwise tests)		
other currencies to dollar: $a_{12} = a_{13} = a_{14} = 0$	9.10*	3
other currencies to yen: $a_{21} = a_{23} = a_{24} = 0$	15.71*	3
other currencies to mark: $a_{31} = a_{32} = a_{34} = 0$	11.77*	3
other currencies to pound: $a_{41} = a_{42} = a_{43} = 0$	9.91*	3
(5) Diagonal variance innovations: $\omega_{ij} = 0$ ($i \neq j$)	146.5*	6

Notes: 'Wald' is the Wald test statistic for the hypothesis in the first column. The covariance matrix of the parameters is computed from the outer product of the scores of the quasi log-likelihood function. An asterisk (*) denotes significance at the 5% level using the χ^2 (df) table.

than the series $\hat{x}(t|T)$ that attempts to add an estimate of $\xi(t)$ based on all information in the sample. The dollar volatility series shows more fluctuations than the other ones, especially when compared to the relatively constant Deutsche mark (log-) volatility. The sub-period averages of $x(t)$ are consistent with the estimates in Table 1: the dollar volatility is moderate until the end of 1977 compared to the eighties; yen volatility is high in the late seventies; and mark volatility is slowly but steadily decreasing over the sample. It is hard to make any reliable inference on the volatility at a particular point in time, since the standard errors of the state vector elements $\hat{x}_i(t|t-1)$ and $\hat{x}_i(t|T)$ obtained from the steady state Kalman recursions are large, even conditional on the parameter estimates.

5. Implications for asset pricing

It is a well-known fact in the exchange rate literature that the forward rate is a biased estimator for the future spot rate. Maintaining rational expectations, the combined existence of an efficient foreign exchange market and a time-varying risk premium can account for the bias. The literature on risk premia has

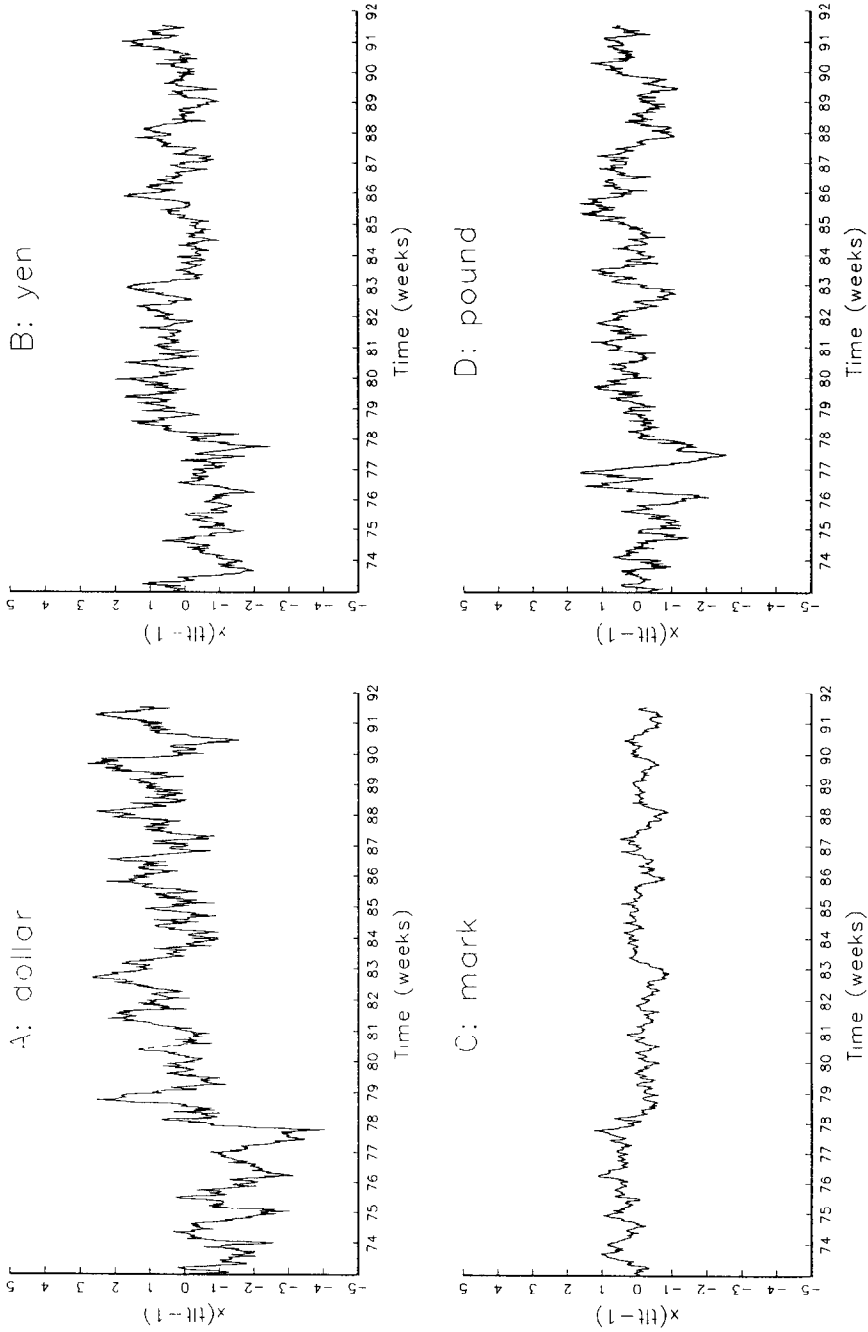


Fig. 2. Time series of estimated conditional log-variances of the four currencies.

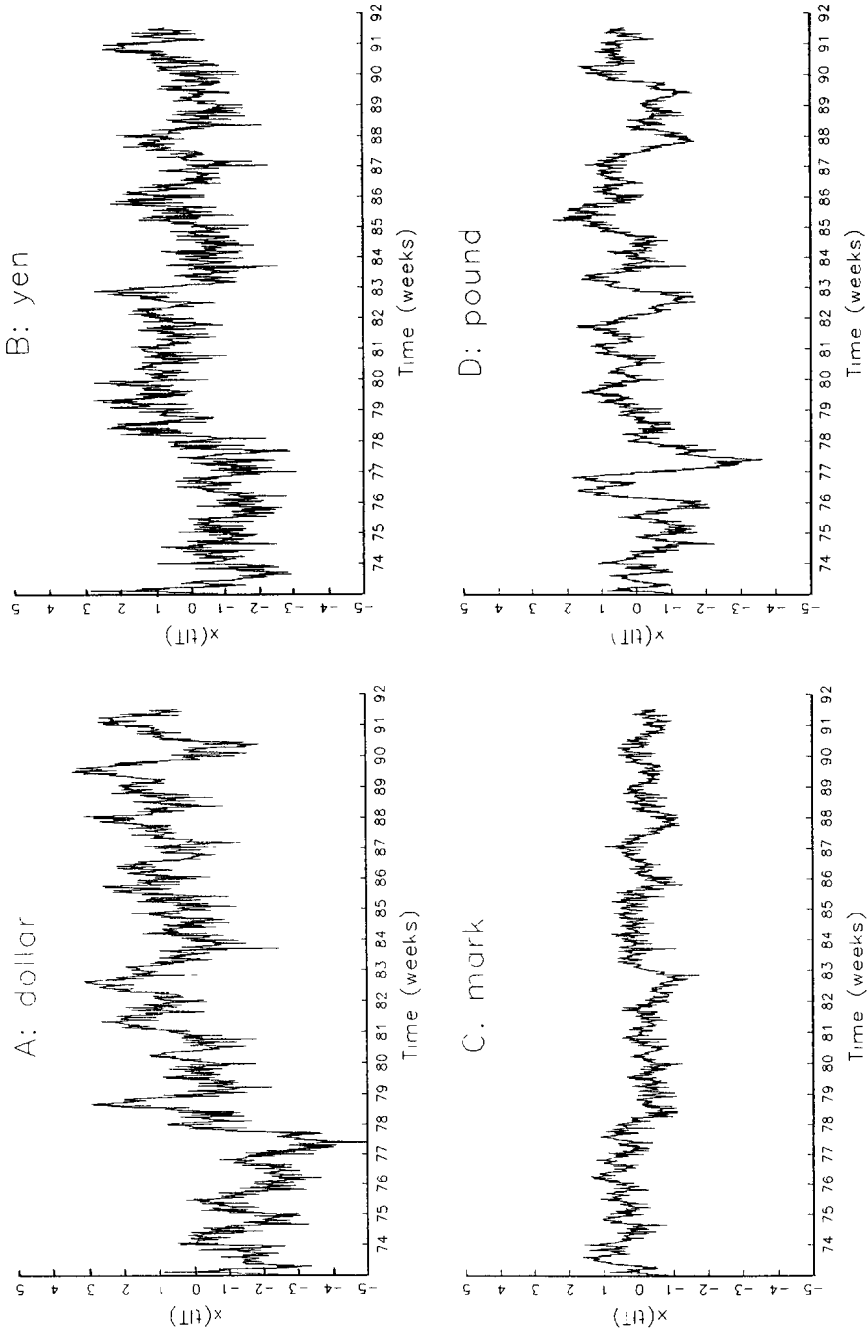


Fig. 3 Time series of 'smoothed' estimates of the (log-)volatilities of the four currencies.

concentrated on developing models to disentangle the joint hypothesis of market efficiency and constant risk premia.⁷

The simplest model in this field is the model of Domowitz and Hakkio (1985) (DH from now on). DH base their model on the asset pricing theory of Lucas (1982). In the DH model the risk premium is related to the difference between the conditional variances of the money supply in two countries. In the empirical part of their paper DH investigate if there is any evidence of a time-varying risk premium in several dollar exchange rates. DH analyze univariate cases and, due to lack of identification, assume that only the dollar news term is heteroskedastic. In a multivariate setup this simplifying assumption is unnecessary. The covariance between exchange rate changes helps to identify all three news components, and thus their variances, in a bivariate system of exchange rates. The identification is brought upon by the existence of a common US news factor in a system of dollar exchange rates. Our extension to the multivariate case differs from Baillie and Bollerslev (1990) due to the explicit use of the currency specific factors.

Extending the risk premium model of DH to a model with the factor structure of the previous sections, leads to the following expression for the one period risk premium:

$$RP_{ij}(t) = \frac{1}{2}(\lambda_i(t) - \lambda_j(t)), \quad (37)$$

where $\lambda_i(t) \equiv \exp(h_i(t|t-1))$ is the conditional variance of factor i . Our dataset consists of weekly exchange rate changes, whereas the most commonly studied risk premium is the one implicit in forward contracts with a maturity of one month. The one month risk premium is computed using the expected variance over a period of four weeks:

$$\lambda_i(t, 4) = C_{i4} \sum_{k=1}^4 \exp(g_i A^k \hat{x}(t+k|t)), \quad (38)$$

where g_i is an indicator vector with zeros in every row except in row i , which is one; and where C_{i4} is a constant arising from taking expectations of a log-normal random variable, and also related to the constant terms α and μ .⁸ The resulting expression for the risk premium of the exchange rate between countries i and j is

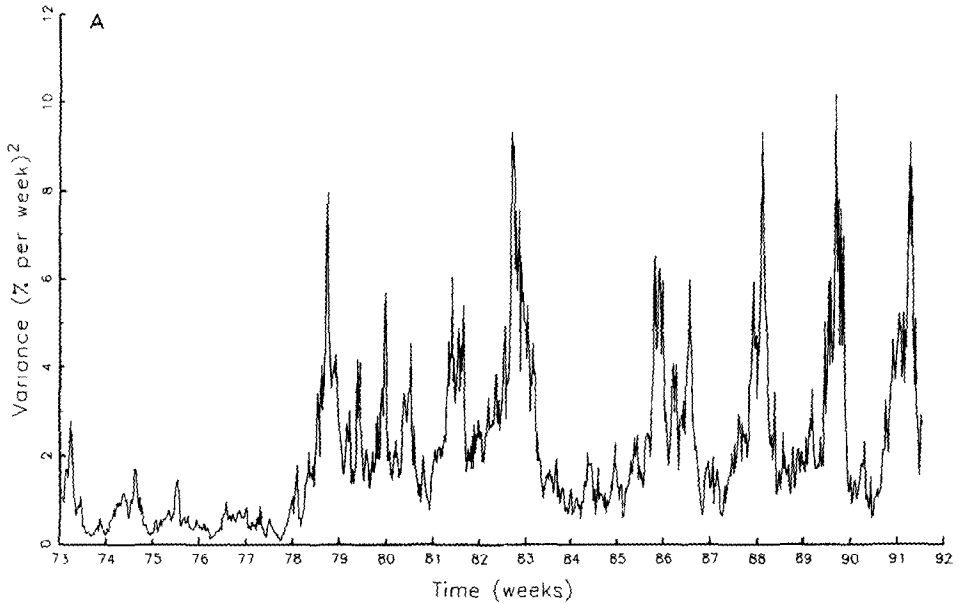
$$\widehat{RP}_{ij}(t) = \frac{1}{2}(\lambda_i(t, 4) - \lambda_j(t, 4)). \quad (39)$$

Figs. 4 through 9 show the estimated risk premia and also the conditional variances of the six bilateral exchange rates, which were the original proxy for

⁷See Hodrick (1987) and Bollerslev, Chou and Kroner (1992) for surveys, and the literature on survey data, e.g. Frankel and Froot (1987), for direct evidence on risk premia. Bekaert and Hodrick (1993) provide an extensive econometric investigation on the existence of the risk premium for the major currencies in the presence of conditional heteroskedasticity.

⁸Because α and μ are not separately identified we have estimated the scale factor implicitly by requiring that the sample average of $\lambda_i(t)$ is equal to the full sample estimation of the constant λ_i in Table 1.

Yen/Dollar Conditional Variance



Yen/Dollar Risk Premium

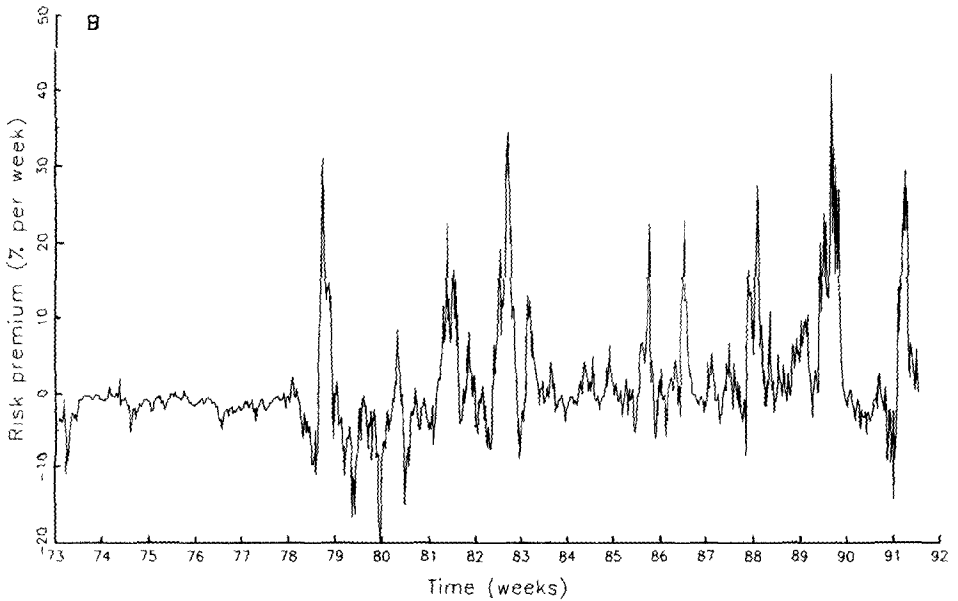


Fig. 4. Conditional variance of yen-dollar exchange rate, $\lambda_y(t, 4) + \lambda_s(t, 4)$ (A), and the risk premium $\frac{1}{2}(\lambda_s(t, 4) - \lambda_y(t, 4))$ (B).

the risk premium used by DH. Table 8 provides some summary statistics of the time series of the risk premia.

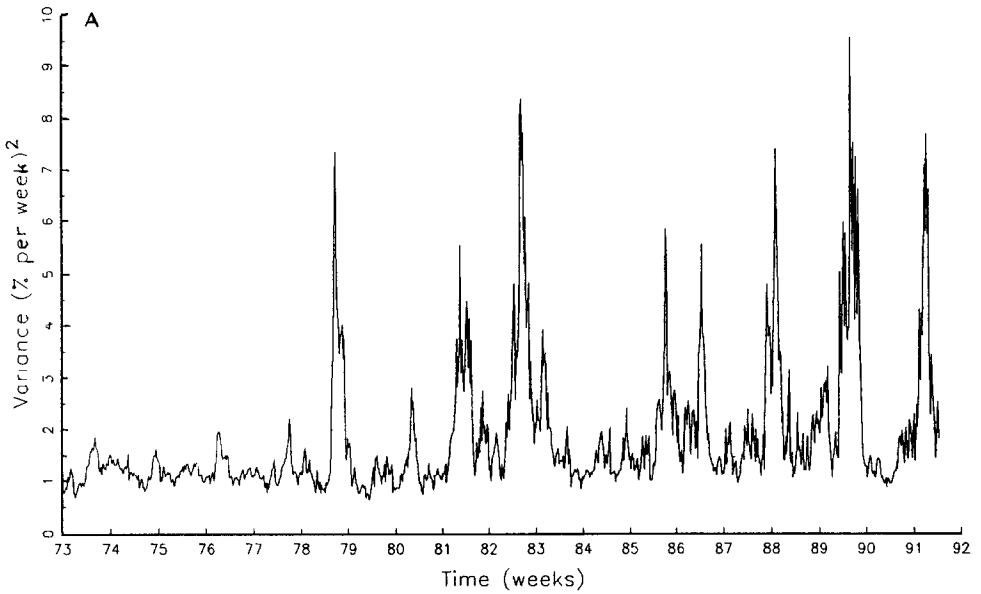
For the three dollar exchange rates the risk premium is dominated by the US news component after 1980. In the seventies the risk premia on the dollar were much smaller and fluctuated less. The sign of the dollar risk premium also differs between the seventies and eighties. For example, the pound/dollar rate (Fig. 6) shows two troughs in 1976 that correspond to peaks in the variance of the pound/dollar rate. The same phenomenon shows up in the yen/dollar rate in

Table 8
Summary statistics of risk premia

	¥/\$	DM/\$	£/\$	DM/¥	£/¥	£/DM
Jan 73–Jun 91						
Mean	1.04	3.59	3.53	2.55	2.50	–0.05
Std. dev.	7.05	8.80	8.04	5.17	4.05	2.80
Minimum	–19.88	–12.00	–10.86	–11.57	–8.79	–9.81
Maximum	42.18	47.96	47.85	22.45	20.16	11.36
Jan 73–Dec 76						
Mean	–1.60	–3.89	–1.56	–2.29	0.04	2.33
Std. dev.	1.57	2.36	2.27	3.13	2.56	2.43
Minimum	–11.07	–9.68	–10.86	–9.03	–8.79	–7.03
Maximum	2.07	2.16	2.80	11.39	9.29	7.86
Jan 77–Dec 80						
Mean	–2.34	1.12	1.67	3.45	4.00	0.55
Std. dev.	6.60	6.74	5.62	5.89	4.01	3.12
Minimum	–19.88	–12.00	–5.78	–11.57	–3.04	5.87
Maximum	31.09	37.06	36.20	22.45	20.16	11.36
Jan 81–Sep 85						
Mean	2.53	6.24	4.84	3.71	2.31	–1.40
Std. dev.	6.72	8.07	8.32	4.16	4.37	2.30
Minimum	–9.05	–3.47	–5.85	–2.67	–7.52	–9.81
Maximum	34.49	41.61	41.72	18.05	16.84	2.22
Sep 85–Feb 87 (Plaza–Louvre)						
Mean	2.31	7.98	5.54	5.66	3.23	–2.43
Std. dev.	5.85	7.17	6.63	5.12	4.62	1.64
Minimum	–6.17	–3.21	–5.78	–2.09	–4.66	–8.77
Maximum	22.87	28.56	27.08	20.81	17.99	0.19
Feb 87–Jun 91						
Mean	4.51	8.66	8.02	4.14	3.51	–0.63
Std. dev.	8.93	9.93	10.13	4.04	3.61	1.63
Minimum	–14.35	–4.14	–6.16	–2.43	–5.28	–7.07
Maximum	42.18	47.96	47.85	20.92	17.58	2.08

Notes: Units are annualized percentages of risk premium on one month forward contracts. A positive entry denotes that the numeraire in the column heading is the more risky currency.

Mark/Dollar Conditional Variance



Mark/Dollar Risk Premium

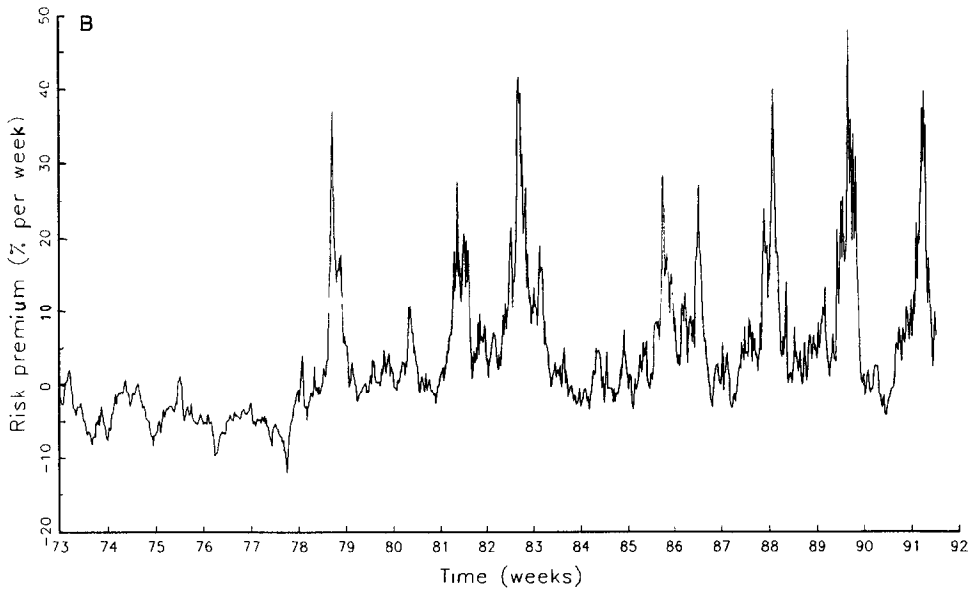
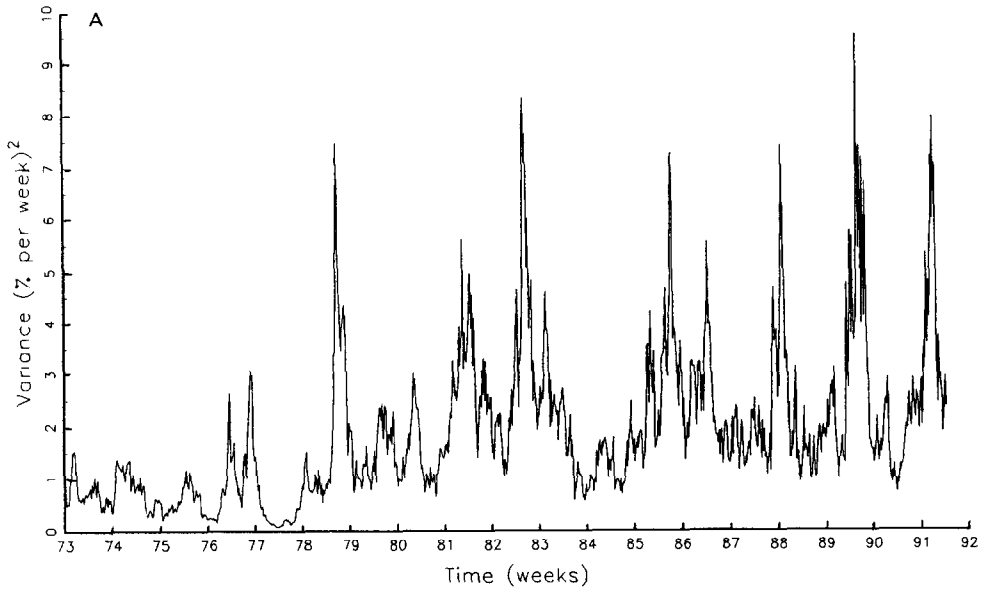


Fig. 5. Conditional variance of mark-dollar exchange rate, $\lambda_M(t, 4) + \lambda_S(t, 4)$ (A), and the risk premium $\frac{1}{2}(\lambda_S(t, 4) - \lambda_M(t, 4))$ (B).

Pound/Dollar Conditional Variance



Pound/Dollar Risk Premium

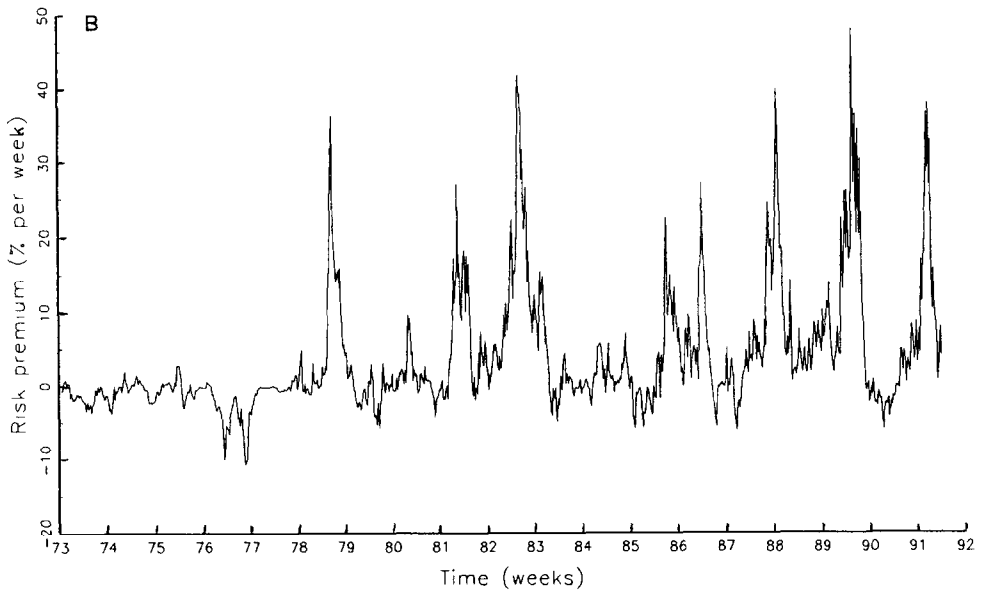
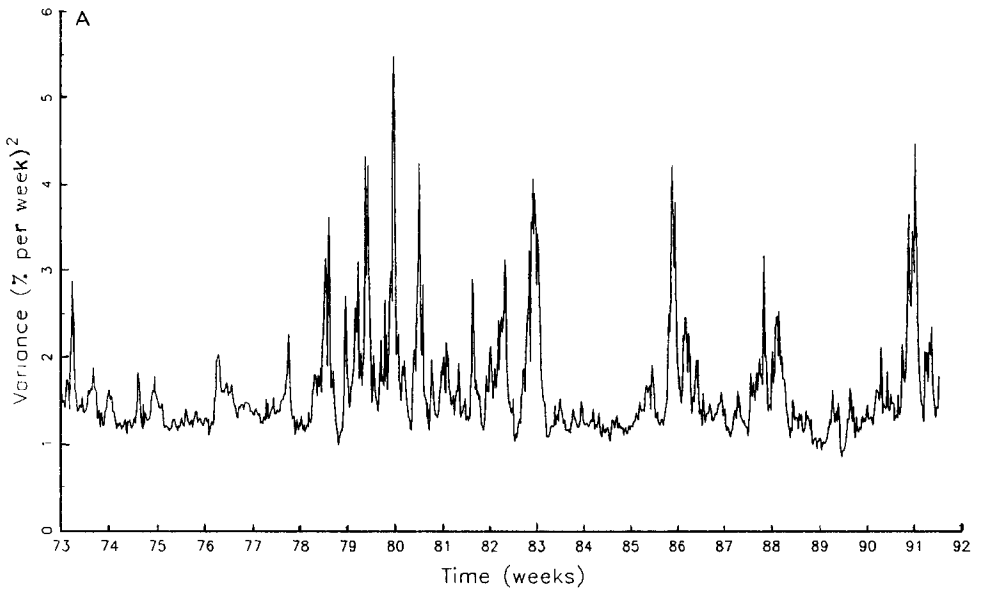


Fig. 6. Conditional variance of pound-dollar exchange rate, $\lambda_s(t, 4) + \lambda_d(t, 4)$ (A), and the risk premium $\frac{1}{2}(\lambda_s(t, 4) - \lambda_d(t, 4))$ (B).

Mark/Yen Conditional Variance



Mark/Yen Risk Premium

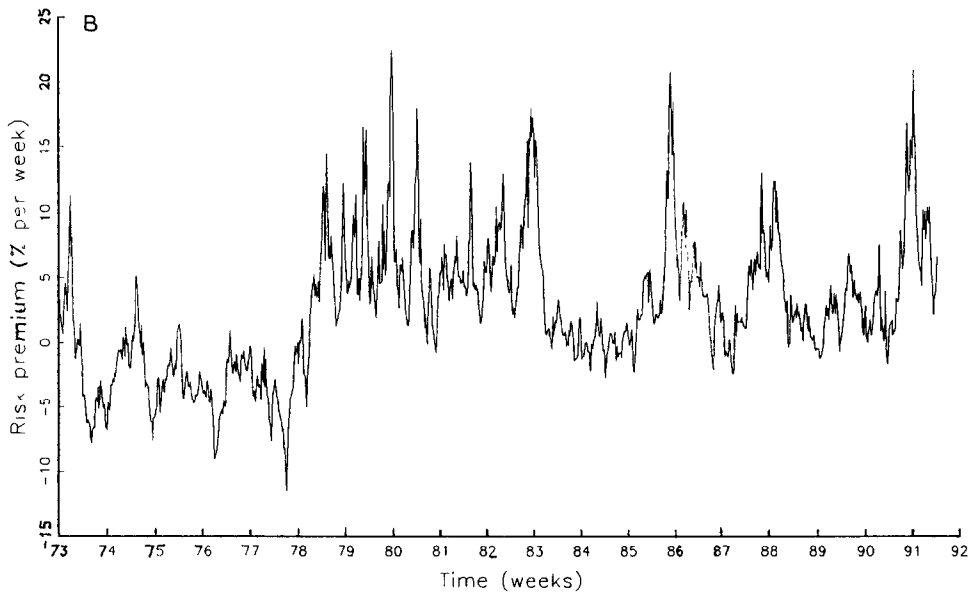
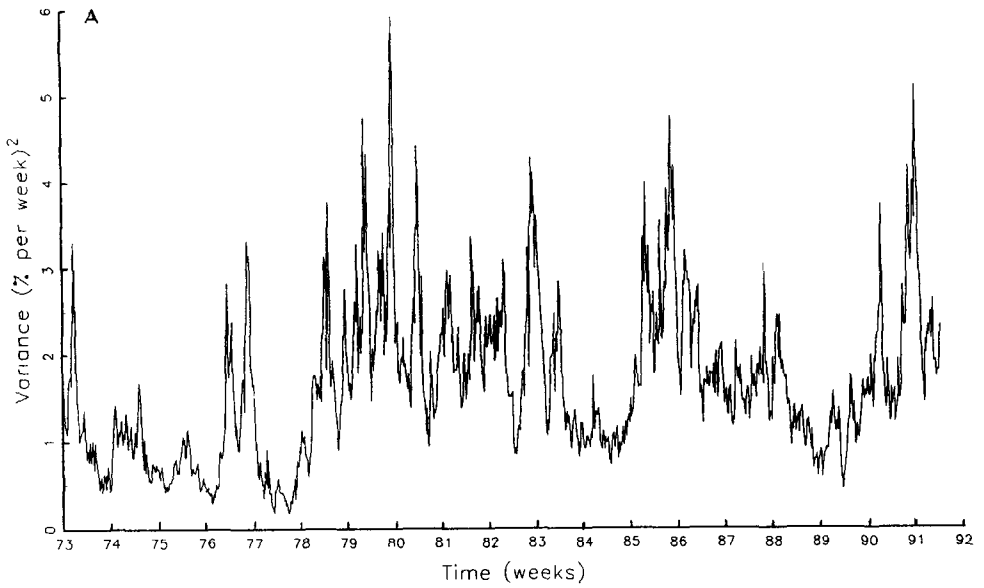


Fig. 7. Conditional variance of mark-yen exchange rate, $\lambda_M(t, 4) + \lambda_Y(t, 4)$ (A), and the risk premium $\frac{1}{2}(\lambda_Y(t, 4) - \lambda_M(t, 4))$ (B).

Pound/Yen Conditional Variance



Pound/Yen Risk Premium

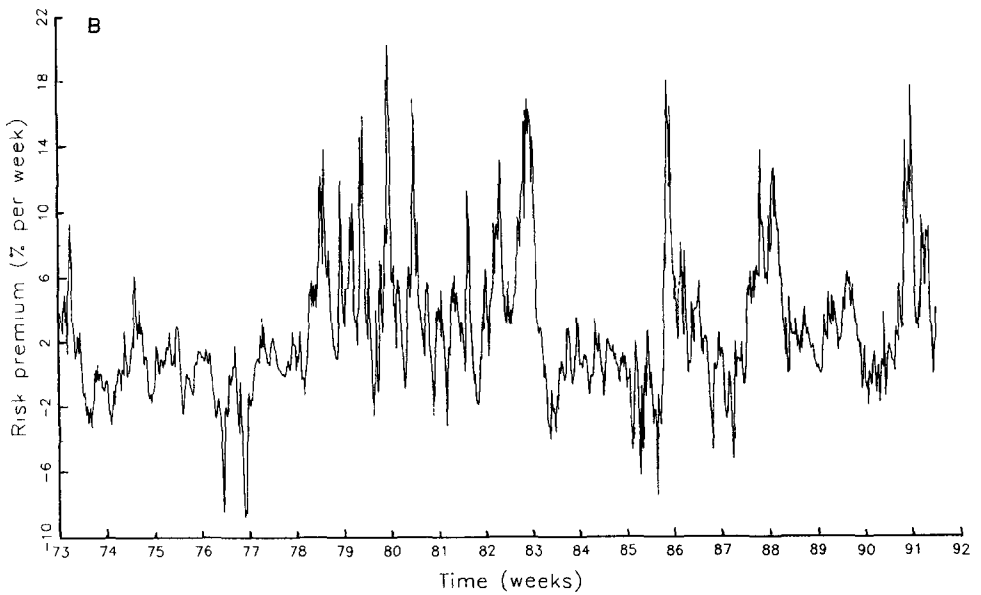
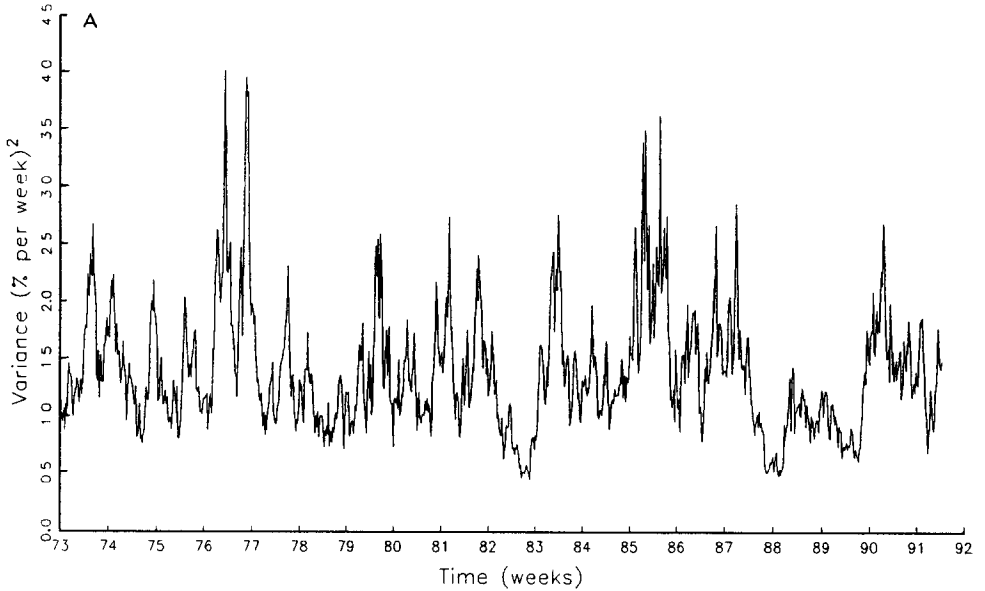


Fig. 8. Conditional variance of pound-yen exchange rate, $\lambda_v(t, 4) + \lambda_f(t, 4)$ (A), and the risk premium $\frac{1}{2}(\lambda_v(t, 4) - \lambda_f(t, 4))$ (B).

Pound/Mark Conditional Variance



Pound/Mark Risk Premium

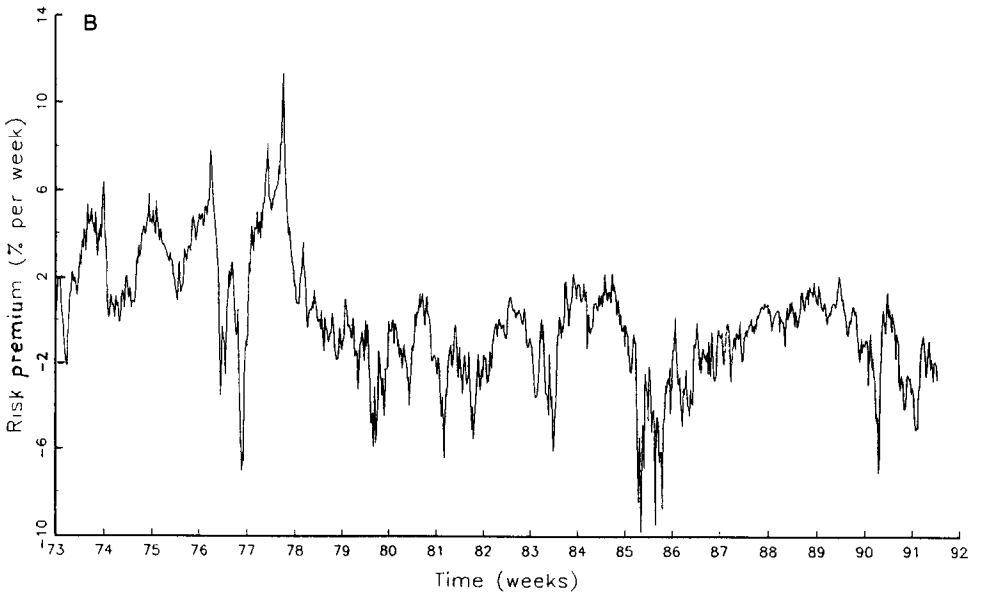


Fig. 9. Conditional variance of pound-mark exchange rate, $\lambda_M(t, 4) + \lambda_E(t, 4)$ (A), and the risk premium $\frac{1}{2}(\lambda_M(t, 4) - \lambda_E(t, 4))$ (B).

1973. In the eighties the signs are reversed: all peaks in the conditional variance of the dollar rates are also peaks in the risk premia.

The figures and Table 8 indicate that the estimated risk premia fluctuate substantially. On average the size of the risk premium is small relative to its own standard deviation for all the subperiods, and almost negligible compared to the variance of exchange rate innovations (see Tables 1 and 2 and note the difference in units). The conditional variances in Figs. 4–9 appear much more irregular than the log-volatility in Fig. 1. Taking the exponent of the conditional log variances attenuates the upward movements in the conditional variance plots during periods of high (log) volatility.

Our measures of the risk premium differ from studies that use survey data. With survey data the risk premium is directly observable as the difference between the expected future spot rate and the forward rate for the same horizon. These direct estimates are on average not very different from our results. Frankel and Froot (1987), for instance, found that the risk premium varied between 2 and 10 percent on an annual basis. Similar magnitudes are reported in Cavaglia, Verschoor and Wolff (1992). We find weekly risk premia sometimes to be 10 percent per week. The excess variability in our risk premia might be due to the fact that we investigate the time series behaviour of exchange rates alone. It is possible that the inclusion of other variables, like macroeconomic variables, would smooth our estimates.

6. Concluding remarks

All bilateral exchange rates, expressed vis-a-vis a common numeraire currency, contain at least one common factor due to the numeraire effect. We have examined empirically to what extent the movements among the four major currencies can be explained by just a set of currency specific factors, each representing the specific news in one of the currencies (dollar, yen, mark, pound). For the seventies we find that all currencies were approximately equally volatile, with some short periods of high German or high Japanese volatility. During the eighties the volatility of the dollar was dominant.

This factor structure has been used to specify a parsimonious multivariate model of time varying volatility. The resulting model is an extension of the stochastic variance model of Harvey, Ruiz and Shephard (1994). The conditional variances from the stochastic variance factor model follow approximately a highly restricted multivariate EGARCH process.

Using weekly data for the full floating exchange rate period 1973–1991, it appears that changes in dollar volatility quickly spread to changes in the volatility of other currencies, even affecting the volatility of cross exchange rates. The effects of increased yen and mark volatility take much longer to transmit to increases in the volatility of other currencies. The Deutsche mark appears to be

the most stable major currency of the last two decades. Its average volatility is below that of the other major free floating currencies, and its volatility has been relatively constant over time.

The variance decomposition of exchange rates also provides a new approach to estimate foreign exchange risk premia in a complete system of currencies. We find that risk premia fluctuate considerably over the sample period.

References

- Andersen, T.G., 1992, Volatility, Working paper, Northwestern University.
- Baillie, R.T. and T. Bollerslev, 1989, The message in daily exchange rates: a conditional variance tale, *Journal of Business and Economic Statistics*, 7, 297–305.
- Baillie, R.T. and T. Bollerslev, 1990, A multivariate generalized ARCH approach to modeling risk premia in forward foreign exchange markets, *Journal of International Money and Finance*, 9, 309–324.
- Baillie, R.T. and T. Bollerslev, 1991, Intra-day and inter-market volatility in foreign exchange rates, *Review of Economic Studies*, 58, 565–585.
- Bekaert, G. and R.J. Hodrick, 1993, On biases in the measurement of foreign exchange risk premiums, *Journal of International Money and Finance*, 12, 115–138.
- Bekaert, G., R.J. Hodrick and D. Marshall, 1993, The implications of first order risk aversion for asset market risk premiums, working paper.
- Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307–327.
- Bollerslev, T., R.Y. Chou and K.F. Kroner, 1992, ARCH modeling in finance: a review of the theory and empirical evidence, *Journal of Econometrics*, 52, 5–59.
- Bollerslev, T., R.F. Engle and J.M. Wooldridge, 1988, A capital asset pricing model with time-varying covariances, *Journal of Political Economy*, 96, 116–131.
- Boothe, P. and D. Glassman, 1987, The statistical distribution of exchange rates, *Journal of International Economics*, 22, 297–319.
- Cavaglia, S.M.F.G., W.F.C. Verschoor and C.C.P. Wolff, 1992, On the biasedness of forward foreign exchange rates: irrationality or risk premia?, unpublished, University of Limburg.
- Diebold, F.X. and M. Nerlove, 1989, The dynamics of exchange rate volatility: a multivariate latent factor ARCH model, *Journal of Applied Econometrics*, 4, 1–21.
- Domowitz, I., C.S. Hakkio, 1985, Conditional variance and the risk premium in the foreign exchange market, *Journal of International Economics*, 19, 47–66.
- Engle, R.F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987–1007.
- Engle, R.F. and T. Bollerslev, 1986, Modeling the persistence of conditional variances, *Econometric Reviews*, 5, 1–50.
- Engle, R.F., T. Ito and W.-L. Lin, 1990, Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market, *Econometrica*, 58, 525–542.
- Engle, R.F., V.K. Ng and M. Rothschild, 1990, Asset pricing with a FACTOR-ARCH covariance structure, *Journal of Econometrics*, 45, 213–237.
- Frankel, J.A. and K.A. Froot, 1987, Using survey data to test standard propositions regarding exchange rate expectations, *American Economic Review*, 77, 133–153.
- Gallant, A.R., D.A. Hsieh and G.E. Tauchen, 1991, On fitting a recalcitrant series: the pound/dollar exchange rate, 1974–1983, in: W.A. Barnett, J. Powell and G. Tauchen (eds.), *Nonparametric and semiparametric methods in econometrics and statistics*, Cambridge University Press (Cambridge).

- Harvey, A.C., E. Ruiz and N. Shephard, 1994, Multivariate stochastic variance models, *Review of Economic Studies*, 61, 247–264.
- Hodrick, R.J., 1987, The empirical evidence on the efficiency of forward and futures foreign exchange markets, Harwood (Chur).
- Hodrick, R.J., 1989, Risk, uncertainty, and exchange rates, *Journal of Monetary Economics*, 23, 433–459.
- Hsieh, D.A., 1989, Modeling heteroskedasticity in daily foreign exchange rates, *Journal of Business and Economic Statistics*, 7, 307–317.
- Jacquier, E., N.G. Polson and P.E. Rossi, 1993, Bayesian analysis of stochastic volatility models, forthcoming in *Journal of Business and Economic Statistics*.
- Lucas, R.E., 1982, Interest rates and currency prices in a two-country world, *Journal of Monetary Economics*, 10, 335–360.
- Nelson, D.B., 1991, Conditional heteroskedasticity in asset returns: a new approach, *Econometrica*, 59, 347–370.
- Nijman, T.E. and F.C. Palm, 1993, Recent developments in modeling volatility in financial data, Discussion paper, Tilburg University (Center).
- Taylor, S.J., 1986, *Modelling Financial Time Series*, Wiley (Chichester).