ELSEVIER

# Forecasting the comovements of spot interest rates

## Miguel A. Ferreira[*]

*ISCTE Business School, Complexo INDEG/ISCTE,
Av. Prof. Anibal Bettencourt, 1600-189 Lisboa,
Portugal*

## Abstract

Time-varying covariance models are compared in the French and German interest rate markets of the pre-euro period. A bivariate, asymmetric dynamic covariance model with level effect best characterizes the in-sample variance–covariance dynamics. Model comparison using economic loss functions raises some doubts with alternative models performing similarly. Out-of-sample results show that the variance–covariance matrix is best forecasted using a VECH model with level effect but no volatility spillover, not entirely confirming the in-sample evidence. Simple models using exponentially-weighted moving averages of past observations perform similarly to the best bivariate model. Thus, some features required to obtain a good in-sample fit do not have additional out-of-sample forecasting power due to overfitting.
© 2005 Elsevier Ltd. All rights reserved.

*JEL classification:* C52; C53; G12; E43

*Keywords:* Interest rates; Covariance models; GARCH; Forecasting

* Tel.: +351 21 7958607; fax: +351 21 7958605.
  *E-mail address:* miguel.ferreira@iscte.pt

## 1. Introduction

The short-term interest rate is one of the most important and fundamental characteristics of financial markets. In particular, forecasts of interest rate volatility are important for a wide variety of economic agents. Portfolio managers require volatility forecasts to make decisions about asset allocation and portfolio selection. Risk managers use them to compute measures of portfolio market risk such as Value at Risk. Options traders require them to price and hedge options. Regulators also need them to supervise and monitor financial institutions behavior.

There are basically two approaches to forecast volatility. The market-based approach uses option market data to infer implied volatility, which is interpreted to be the market's best estimate of future volatility conditional on the option pricing model. The model-based approach forecasts volatility from time series information, generating what are usually called historical forecasts. The empirical evidence in Day and Lewis (1992) and Amin and Ng (1997)[1] suggests that implied volatility is useful in forecasting future volatility. However, there is also some consensus that implied volatility is an upward biased forecast of future volatility and that historical volatility forecasts contain some additional information.

Most of the work on the model-based approach has been focused on univariate time-varying variance models. Relatively little work has been done on multivariate time-varying covariance models. For examples of univariate models, see Glosten et al. (1993) and West and Cho (1995). Also, there has been little work on comparing the forecasting performance of time-varying covariance models with implied covariance from option prices.[2] Campa and Chang (1998) find that implied correlations from foreign exchange options tend to be more accurate forecasts of realized correlation than ones based on historical prices. However, implied correlations do not fully incorporate the information in historical data because bivariate GARCH models of correlation are able to improve forecast accuracy.

The purpose of this paper is to examine the variance—covariance process of the short-term spot riskless interest rate (short rate) in France and Germany, using the pre-euro period from 1981 to 1997 with weekly frequency. We take a model-based approach by considering multivariate GARCH models estimated from historical interest rates for both countries and examine the forecasting ability of alternative covariance models of the spot interest rate volatility.

The examination of French and German interest rates during the pre-euro era provides insights that are unavailable by studying US interest rates. First, our data include the period of the 1992—1993 exchange rate turmoil in the Exchange Rate Mechanism (ERM) of the European Monetary System (EMS) and the distinct

---

[1] Day and Lewis (1992) and Amin and Ng (1997) present empirical evidence comparing implied and historical volatility forecasts in the stock market and interest rate market, respectively.

[2] This comparison turns out to be of difficult implementation because the set of options that contains information on assets correlation is restricted. The existing work in this area uses foreign currency options. An implied correlation is inferred from the implied variances of options on three currency-pairs because the foreign currencies trade directly against each other (cross-rate).

implications of this turmoil on the French and German interest rate markets.[3] Thus, we provide new insights about the best way to model interest rate volatility and cross-correlations during periods of high uncertainty. Second, though France and Germany had many real economic factors in common as members of the European Union (EU), their central banks differed sharply in terms of credibility to control inflation. This allows us to learn the importance of differing levels of central bank credibility for interest rate volatility. Finally, in contrast with the US case,[4] the French and German central banks tended to target monetary aggregates in conducting monetary policy rather than interest rate levels over our sample period. Thus, we can provide insights to the effects of such monetary policy orientations.

We study the relation between French and German interest rates along several lines. First, there are reasons to believe that our data contain volatility spillover effects because both countries belong to the EU and to the ERM. Furthermore, the German central bank, given its credibility in lowering inflation, was taken as a benchmark for the monetary policy of the remaining European central banks. Thus, we expect to find evidence of volatility spillover effects from the German interest rate process to the French interest rate process.

Second, we compare the in-sample and out-of-sample forecasting performance of the alternative variance—covariance structures and their economic implications in terms of portfolio selection, hedging and risk management. For comparison we also consider simple historical forecasts such as the equally- and exponentially-weighted sample variance, covariance and correlation. Previous studies across a wide variety of markets compare the out-of-sample forecasting ability of several GARCH-type models and sample variances and find that there is no clear winner using a MSPE loss function. Results in Figlewski (1997) show that GARCH models fitted to daily data are effective in forecasting the volatility in stock markets both for short and long horizons, but they are much less useful in making out-of-sample forecasts in other markets beyond short horizons.[5] In particular, using 3-month US Treasury bill yields, sample variances generate more accurate forecasts than GARCH model. Also, Ederington and Guan (1999) find that sample variance dominates GARCH models in the 3-month Eurodollar market.

Third, Ferreira (2000a) studies the forecasting performance of alternative models of the spot interest rate variance using univariate GARCH models for France and Germany. The results show that modeling volatility as a function of the interest rate innovations (news effect) is equally important as modeling volatility as a function of

---

[3] The ERM was characterized by semi-fixed exchange rates. The exchange rates could float within a band around a fixed central parity. The band was 2.25% before August 1, 1993 when there was a band widening to 6% in consequence of the pressure to a depreciation of the French franc and other weak currencies against the Deutschemark. The French central bank increased short-term interest rate levels (and volatility) sharply to avoid currency devaluation, but without success.

[4] With the exception of the Federal Reserve monetary targeting experiment of October 1979 through September 1982.

[5] This sensitivity of results to the forecast horizon can also be found in West and Cho (1995) with the GARCH model having a slight edge over historical volatility for short horizons across several foreign exchange markets, but this advantage disappearing for longer horizons.

the interest rate (level effect). However, the in-sample performance of the models is not always confirmed by the forecasting power and efficiency tests. The out-of-sample tests support that models with only the news effect have similar forecasting power and efficiency to mixed level and news effects models. Moreover, sample variance forecasts using exponentially declining weights also have out-of-sample forecasting power and efficiency similar to the best models of the spot interest rate volatility. The tight links between the two economies suggest that the forecasting exercise should be done in a bivariate framework instead to increase the forecasting power.

Fourth, we consider the class of multivariate GARCH models proposed by Kroner and Ng (1998), namely the asymmetric general dynamic covariance (ADC) models, that can capture asymmetric variance effects as well as asymmetric covariance effects. The asymmetric variance effect consists of negative innovations to the asset returns process leading to a higher subsequent variance than positive innovations. There is extensive evidence on this effect across assets in a univariate framework, see for example Glosten et al. (1993). In a multivariate framework, only Kroner and Ng (1998) have extended the models to incorporate asymmetric effects in the variance–covariance specification. They introduce the possibility of an asymmetric covariance effect, i.e., negative and positive innovations might not have different impact on the subsequent covariance across assets and markets. If the asymmetric volatility effect is caused by an increase in the information flow following shocks, then the covariance can also be affected because there will be a change in the relative rate of information across assets and markets. Kroner and Ng (1998) apply the model in the context of US stock market portfolios. This paper is the first to apply the approach to interest rate markets.

Finally, we extend the class of GARCH multivariate models of Kroner and Ng (1998) to accommodate a level effect in the interest rate volatility process, following the approach of Brenner et al. (1996). A level effect in the spot interest rate volatility is first documented in Chan et al. (1992). They find that the level effect is the most important feature of the models and that the volatility sensitivity to the rate level is in excess of unity. Brenner et al. (1996), combining a univariate GARCH model with a level effect, find that the volatility sensitivity to the interest rate level is smaller than that found in Chan et al. (1992) and that innovations to the interest rate process are also an important feature of the volatility dynamics. Nevertheless, the level effect remains a significant and important effect. The same results hold for the French and German short rates as shown in Ferreira (2000b).

This paper's in-sample results show that a bivariate, asymmetric general dynamic covariance model with level effect best characterizes French and Germany interest rate variance–covariance dynamics during our sample period. Moreover, our results confirm the notion that Germany was the leader in terms of monetary policy in Europe and that France was the follower because we find evidence of volatility spillovers from Germany to France, but not the reverse. With respect to the correlation between interest rates, the evidence shows that the correlation was higher (lower) when German (French) rates decreased (increased), which can be interpreted as France taking advantage of the German central bank credibility in terms of controlling inflation.

However, the in-sample model performance under economic loss functions (hedge ratio, portfolio selection and Value at Risk) raises some doubts. The models perform similarly, and in certain cases, including additional effects, do not improve model performance. Also, in-sample model forecasting power shows that there is no significant improvement in forecasting accuracy in terms of variance relative to the univariate models in Ferreira (2000b). In contrast, including exchange rate information improves the forecasting accuracy for variance, specially for Germany, despite not improving the forecasting power for covariance.

Furthermore, the best performing models in-sample are not, in general, the best out-of-sample. In fact, the interest rate variance–covariance matrix is best forecast out-of-sample using a VECH model with level effect; including volatility spillover effects is not important. In addition, simple models using either equally- or exponentially-weighted moving averages of past observations perform similarly to the best bivariate model in terms of both variance and covariance. Overall, the out-of-sample results show evidence that some of the features required to obtain a good in-sample fit do not have additional out-of-sample forecasting power due to overfitting.

The organization of the paper is as follows. Section 2 describes the issue of forecasting variances and covariances and presents the alternative time-varying covariance specifications. In Section 3, time-varying covariance models are applied to our data set of French and German short-term spot interest rates. Parameter estimates and second moments estimates are analyzed. Section 4 compares the models in terms of out-of-sample forecasting power for variances, covariance and correlation. Section 5 completes the analysis by studying the economic implications of our results for hedging, portfolio selection and risk management. Section 6 studies the link between the interest rate and foreign exchange markets volatility and Section 7 concludes.

## 2. Time-varying covariance models

### 2.1. Modeling and forecasting variance and covariance

We are interested in forecasting both variances and covariance between French and German short rates. The empirical model for the short rate $r_{it}$ is given by, for $i$, $j = 1, 2$,

$$r_{it} - r_{it-1} = \mu_i + \kappa_i r_{it-1} + \varepsilon_{it}, \quad E[\varepsilon_{it}|\mathcal{F}_{t-1}] = 0, \quad E[\varepsilon_{it}\varepsilon_{jt}|\mathcal{F}_{t-1}] \equiv h_{ijt}, \tag{1}$$

where the conditional mean function for the interest rate is given by a first-order autoregressive process, $\mathcal{F}_{t-1}$ is the information set up to time $t-1$, $h_{iit}$ is the conditional variance at time $t$ and $h_{ijt}$ is the conditional covariance at time $t$. The conditional variance–covariance matrix $H_t$ is $2 \times 2$ with elements given by $h_{ijt}$.

Multivariate GARCH models are among the most widely used time-varying covariance models. In the basic bivariate GARCH model, the components of the conditional variance–covariance matrix $H_t$ vary through time as functions of the

lagged values of $H_t$ and of the squared innovations, $\varepsilon_t$ is $2 \times 1$ vector of innovations. We model the French and German short rates using bivariate GARCH(1,1) structures because this process is able to successfully represent most financial time series. We implement several bivariate GARCH specifications that have been suggested in the literature as we will see below.

### 2.2. VECH model

The VECH model of Bollerslev et al. (1988) is a parsimonious version of the most general bivariate GARCH model, and is given by, for $i, j = 1, 2$,

$$h_{ijt} = \omega_{ij} + \beta_{ij} h_{ijt-1} + \alpha_{ij} \varepsilon_{it-1} \varepsilon_{jt-1}, \tag{2}$$

where $\omega_{ij}$, $\beta_{ij}$ and $\alpha_{ij}$ are parameters. The VECH model is simply an ARMA process for $\varepsilon_{it} \varepsilon_{jt}$. The covariance is estimated using geometrically declining weighted averages of past cross-products of innovations to the interest rate process. The VECH model has nine parameters in the bivariate case. An implementation problem is that the model may not yield a positive definite covariance matrix. The VECH model does not allow for volatility spillover effects, i.e., the conditional variance of a given variable is not a function of other variable shocks and past variance.

### 2.3. Constant correlation (CCORR) model

In the CCORR model of Bollerslev (1990), the conditional covariance is proportional to the product of the conditional standard deviations. Consequently, the conditional correlation is constant across time. The model is described by the following equations, for $i, j = 1, 2$,

$$h_{iit} = \omega_{ii} + \beta_{ii} h_{iit-1} + \alpha_{ii} \varepsilon_{it-1}^2, \quad h_{ijt} = \rho_{ij} \sqrt{h_{iit-1}} \sqrt{h_{jjt-1}}, \quad \text{for all } i \neq j. \tag{3}$$

The conditional covariance matrix of the CCORR model is positive definite if and only if the correlation matrix is positive definite. The number of parameters is seven in the bivariate case. Like in the VECH model there is no volatility spillover effects across series.

### 2.4. BEKK model

The BEKK model of Engle and Kroner (1995) addresses the problem of the estimated conditional covariance matrix being positive definite. The model is given by:

$$H_t = \Omega + B' H_{t-1} B + A' \varepsilon_{t-1} \varepsilon_{t-1}' A, \tag{4}$$

where $\Omega$, $A$, and $B$ are $2 \times 2$ matrices, with $\Omega$ symmetric. In the BEKK model, the conditional covariance matrix is determined by the outer product matrices of the

vector of past interest rate innovations. As long as $\Omega$ is positive definite, the conditional covariance matrix is also positive definite because the other terms in Eq. (4) are expressed in quadratic form. While this model overcomes the positive definiteness problem, it has 11 parameters in the bivariate case − more than the VECH model.

## 2.5. Asymmetric general dynamic covariance (ADC) model

The ADC model proposed by Kroner and Ng (1998) nests many existing multivariate GARCH models. The conditional covariance specification is given by:

$$H_t = D_t R D_t + \Phi \circ \Theta_t, \tag{5}$$

where $\circ$ is the Hadamard product operator (element-by-element matrix multiplication), $D_t$ is a diagonal matrix with elements $\sqrt{\theta_{iit}}$, $R = [\rho_{ij}]$ is the correlation matrix, $\Phi = [\phi_{ij}]$ is a matrix of parameters with $\phi_{ii} = 0$ for all $i$, $\Theta_t = [\theta_{ijt}]$,

$$\theta_{ijt} = \omega_{ij} + b_i' H_{t-1} b_j + a_i' \varepsilon_{t-1} \varepsilon_{t-1}' a_j + g_{it-1}' \eta_{t-1} \eta_{t-1}' g_j, \quad \text{for all } i,j \tag{6}$$

$a_i$, $b_i$, $i = 1, 2$, are $2 \times 1$ vectors of parameters, $\omega_{ij}$ are scalars with $\Omega \equiv [\omega_{ij}]$ being a diagonal positive definite $2 \times 2$ matrix, and $\eta_t = [\eta_{1t}, \eta_{2t}]'$ with $\eta_{it} = \min[\varepsilon_{it}, 0]$.

The ADC model combines the CCORR and BEKK models. The first term in Eq. (5) is like the CCORR model but with variance functions given by the BEKK model. The second term in Eq. (5) has zero diagonal elements with the off-diagonal elements given by the BEKK model covariance functions scaled by the $\phi_{ij}$ parameters. This model incorporates asymmetric effects in both variances and covariances, thus extending the approach of Glosten et al. (1993) to a multivariate framework. While the existence of asymmetric volatility effects is well documented, there is not much work on asymmetric effects in covariances.

A symmetric version of the ADC model, named GDC model, can be easily obtained by dropping the last term in Eq. (6). The GDC model nests the other multivariate GARCH models discussed above.[6] The VECH model assumes that $\rho_{ij} = 0$ for all $i \neq j$.[7] The BEKK model has the restrictions $\rho_{ij} = 0$ for all $i \neq j$ and $\phi_{ij} = 1$ for all $i \neq j$. The CCORR model considers $\phi_{ij} = 0$ for all $i \neq j$. Similarly, the ADC model can be reduced to the asymmetric versions of the other multivariate GARCH specifications by imposing the appropriate restrictions,[8] which are obvious extensions of the symmetric models by adding an additional term to the variance−covariance functional form.

---

[6] See proposition 1 in Kroner and Ng (1998).
[7] The GDC model does not nest exactly the VECH model, but a restricted version of it with restrictions $\beta_{ij} = \beta_{ii}\beta_{jj}$ and $\alpha_{ij} = \alpha_{ii}\alpha_{jj}$.
[8] See proposition 2 in Kroner and Ng (1998).

## 2.6. ADC-Levels (ADC-L) model

One of the most important features of the short-term interest rate volatility is the level effect, i.e., volatility is a function of the interest rate level. Chan et al. (1992) document this effect using a conditional variance specification where the interest rate level solely drives the evolution of the volatility. They find a volatility sensitivity to the interest rate level very high in excess of the unity. Subsequent work by Brenner et al. (1996) confirms that the level effect exists, but finds it to be considerably smaller (and consistent with a square-root process) when both level and conditional heteroskedasticity effects are allowed in a GARCH framework. Here, we consider a multivariate version of the Brenner et al. (1996) model, which combines a multivariate GARCH model with level effect as well as asymmetric volatility effect.

The ADC model of Kroner and Ng (1998) can be extended to incorporate level effect in the variance−covariance function in the following way,

$$\theta_{ijt} = \omega_{ij}^* + b_i' H_{t-1}^* b_j + a_i' \varepsilon_{t-1}^* \varepsilon_{t-1}^{*'} a_j + g_i' \eta_{t-1}^* \eta_{t-1}^{*'} g_j, \quad \text{for all } i,j \tag{7}$$

where $\omega_{ii}^* = \omega_{ii}/r_{it-1}^{2\gamma_i}$, the diagonal elements of the matrix $H_{t-1}^*$ are multiplied by the factor $(r_{it-1}/r_{it-2})^{2\gamma_i}$ for all $i$, the elements of the vector $\varepsilon_{t-1}^*$ are divided by $r_{it-1}^{\gamma_i}$ for all $i$, and the elements of the vector $\eta_{t-1}^*$ are divided by $r_{it-1}^{\gamma_i}$ for all $i$. The ADC-Levels model nests the asymmetric versions of the other multivariate GARCH models combined with level effect. The exact form of the conditional variance in the asymmetric VECH model with level effect (A-VECH-Levels model, A-VECH-L) is given by:

$$h_{iit} = \left[ \omega_{ii} + \beta_{ii} h_{iit-1} r_{it-2}^{-2\gamma_i} + \alpha_{ii} \varepsilon_{it-1}^2 + \gamma_{ii} \eta_{it-1}^2 \right] r_{it-1}^{2\gamma_i}. \tag{8}$$

The CCORR model with both level and asymmetric effects (A-CCORR-Levels, A-CCORR-L) will also have a variance function given by Eq. (8). The symmetric versions of these models are obtained by dropping the last term in the Eq. (7) and referred below as, respectively, GDC-Levels (GDC-L), VECH-Levels (VECH-L) and CCORR-Levels (CCORR-L).

# 3. Models estimation

## 3.1. Data description

Our data consist of French and German interbank one-month interest rates (middle rate) obtained from Datastream (codes FFRMM1M and GERMDRM). The interest rates are given in percentage and annualized form. We interpret these rates as proxies for the instantaneous riskless interest rate. The frequency is weekly (every Tuesday). The data cover the period from January 1, 1981 to December 31, 1997, providing 887 observations in total (886 interest rate changes). Table 1 presents summary statistics for the short rate changes and squared changes. The French short

Table 1
Summary statistics on short rate

| Variable | $\Delta r_t$ | | $(\Delta r_t)^2$ | |
|---|---|---|---|---|
| | France | Germany | France | Germany |
| Mean | −0.008 | −0.007 | 0.167 | 0.044 |
| Standard deviation | 0.408 | 0.211 | 0.992 | 0.377 |
| Skewness | 2.624 | 4.532 | 13.32 | 21.21 |
| Excess kurtosis | 33.92 | 70.25 | 224.8 | 517.9 |
| Minimum | −2.750 | −1.450 | 0.000 | 0.000 |
| 25th Percentile | −0.093 | −0.050 | 0.001 | 0.000 |
| Median | 0.000 | 0.000 | 0.004 | 0.003 |
| 75th Percentile | 0.063 | 0.025 | 0.035 | 0.010 |
| Maximum | 4.500 | 3.125 | 20.25 | 9.766 |
| ADF | −15.36 | −16.42 | −12.52 | −13.44 |
| ACF(1) | −0.102 | 0.125 | 0.368 | 0.354 |
| ACF(2) | 0.074 | −0.079 | 0.163 | 0.068 |
| ACF(3) | −0.053 | −0.072 | 0.107 | 0.109 |
| ACF(10) | −0.004 | 0.079 | 0.005 | 0.076 |
| ACF(30) | −0.107 | 0.068 | 0.075 | 0.011 |
| ACF(50) | −0.021 | 0.012 | −0.009 | 0.004 |
| $Q(10)$ | 33.43 [0.000] | 49.53 [0.000] | 152.6 [0.000] | 132.6 [0.000] |
| $Q(30)$ | 106.8 [0.000] | 78.55 [0.000] | 337.3 [0.000] | 133.9 [0.000] |
| $Q(50)$ | 162.9 [0.000] | 144.8 [0.000] | 422.7 [0.000] | 147.7 [0.000] |

This table reports summary statistics on the interest rates changes $\Delta r_t \equiv r_t - r_{t-1}$ and squared changes $(\Delta r_t)^2$. ADF denotes the Augmented Dickey–Fuller unit root test statistic which has a 5% critical value of −2.86. ACF($k$) denotes the value of the autocorrelation function of order $k$. $Q(k)$ is the Ljung–Box test statistics of order $k$ with $p$-values in brackets.

rate was 2.8 percentage points higher on average than the German, and about twice as volatile.

As in most of the literature, we focus on nominal interest rates because defining real interest rates is subject to serious measurement errors. The use of weekly frequency minimizes the discretization bias resulting from estimating a continuous-time process using discrete periods between observations, but avoids the potential microstructure bias that could be caused by daily frequency. On the other hand, the use of a one-month rate maturity proxy for the instantaneous interest rate is unlikely to create a significant proxy bias as shown in Chapman et al. (1999). Finally, we use interbank interest rates instead of Treasury bill yields because of the low liquidity of the Treasury bill secondary market in both Germany and France, in sharp contrast with the high liquidity of the interbank money market. The concern regarding the existence of a default premium in the interbank money market should not be severe since these interest rates result from trades among high-grade banks that are considered almost default-free.

### 3.2. Parameter estimates

Following standard practices, we estimate the models by maximizing the likelihood function under the assumption of independent and identically distributed

(i.i.d.) innovations. In addition, we assume that the conditional density function of the innovations is given by the normal distribution. Conditional normality of the innovations is often difficult to justify in many empirical applications that use leptokurtic financial data. However, the maximum likelihood estimator based on the normal density has a valid quasi-maximum likelihood (QML) interpretation as shown by Bollerslev and Wooldridge (1992).

Table 2 presents the parameter estimates for the GDC, GDC-Levels and ADC-Levels models. We focus on these because they nest the other models as we have seen above. The estimates provide guidance on two distinct questions. First, which is the best multivariate GARCH specification, i.e., VECH, CCORR, BEKK or GDC model? Second, do we need to extend these multivariate GARCH models, which only include news effects, to also include level and asymmetric effects in the variance−covariance specification in order to model the short rate?

The first question is addressed by comparing the results for the standard multivariate GARCH models with the GDC model. In particular, we checked whether the GDC model could be reduced to one of the other restricted models. For this we tested the restrictions on the parameters $\phi_{12}$ and $\rho_{12}$. The null hypothesis $\rho_{12} = 0$ was rejected at the 5% level. The null hypothesis $\phi_{12} = 1$ was rejected at the 5% level but we could not reject that $\phi_{12} = 0$. These results indicated that the estimated GDC model was statistically different from the VECH and BEKK models. However, one of the restrictions consistent with the CCORR model ($\phi_{12} = 0$) was not rejected. To further investigate the hypothesis of constant correlation, we conducted a log-likelihood ratio test of the CCORR model against the alternative of the GDC model. We clearly rejected the CCORR model because the test statistic was 25.6 (*p*-value = 0.002).

Next, we extend the GDC model to accommodate level effect in the variance−covariance process, i.e., there is a positive relation between the interest rate volatility and the interest rate level. The level effect ($\gamma$ parameters) was statistically significant at the 5% level and the results for both countries were consistent with a square-root level effect, although France had lower volatility sensitivity than Germany. Also, the level effect results were consistent across the different multivariate models and likelihood ratio tests overwhelmingly rejected the models without level effect against the alternative of level effect. Another issue related to the introduction of the level effect was that the GDC specification became similar to the other multivariate specifications. In fact, we could not reject the null hypothesis consistent with the VECH-Levels and CCORR-Levels specifications, i.e., $\rho_{12} = 0$ and $\phi_{12} = 0$. Moreover, likelihood ratio tests did not reject these models against the alternative of the GDC-Levels model. Nevertheless, we still rejected the null $\phi_{12} = 1$ which was evidence against the BEKK-Levels model. The level effect evidence is mainly consistent with the univariate results for both countries in Ferreira (2000b), in particular the magnitude of the level effect is similar to the one found using a univariate GARCH-Levels model combining level and news effects.

Another common feature of the volatility process is the existence of asymmetric effects in the variance−covariance dynamics, meaning that positive and negative

Table 2
Model parameter estimates of the short rate

| | GDC | | GDC-L | | ADC-L | |
|---|---|---|---|---|---|---|
| | Estimate | Std error | Estimate | Std error | Estimate | Std error |
| $\mu_1(\times 100)$ | 0.1140 | 1.254 | 0.3502 | 1.190 | 1.2431 | 0.938 |
| $\mu_2(\times 100)$ | −0.6297 | 0.953 | −0.0502 | 0.976 | −0.6817 | 0.835 |
| $\kappa_1(\times 100)$ | −0.0234 | 0.255 | −0.1220 | 0.210 | −0.3353* | 0.155 |
| $\kappa_2(\times 100)$ | 0.1151 | 0.237 | −0.0270 | 0.207 | 0.0936 | 0.153 |
| $\omega_{11}$ | 0.0293** | 0.011 | −0.0225* | 0.009 | −0.0329** | 0.012 |
| $\omega_{12}$ | −0.0362** | 0.008 | −0.0231 | 0.039 | −0.0339 | 0.028 |
| $\omega_{22}$ | 0.0000 | 0.010 | −0.0105 | 0.014 | 0.0064 | 0.013 |
| $a_{11}$ | 0.8223** | 0.113 | 0.2732 | 0.163 | 0.4025* | 0.188 |
| $a_{12}$ | −0.0067 | 0.013 | −0.0014 | 0.005 | −0.0002 | 0.005 |
| $a_{21}$ | 0.1492 | 0.107 | 0.0363 | 0.036 | 0.0123 | 0.019 |
| $a_{22}$ | 0.7393** | 0.258 | 0.1503 | 0.077 | 0.0998 | 0.077 |
| $b_{11}$ | 0.7580** | 0.049 | −0.7288** | 0.085 | −0.6430** | 0.075 |
| $b_{12}$ | −0.0082 | 0.011 | −0.0050 | 0.005 | −0.0028 | 0.008 |
| $b_{21}$ | 0.1573* | 0.074 | 0.1992 | 0.132 | 0.2986* | 0.130 |
| $b_{22}$ | 0.8273** | 0.061 | 0.8133** | 0.041 | 0.8102** | 0.037 |
| $\eta_{11}$ | | | | | 0.1079 | 0.110 |
| $\eta_{12}$ | | | | | 0.0097 | 0.009 |
| $\eta_{21}$ | | | | | −0.1432 | 0.132 |
| $\eta_{22}$ | | | | | −0.0980 | 0.066 |
| $\rho_{12}$ | 0.1950** | 0.061 | 0.1045 | 0.061 | 0.0484 | 0.109 |
| $\phi_{12}$ | −0.2350 | 0.201 | 0.1638 | 0.172 | 0.3057 | 0.234 |
| $\gamma_1$ | | | 0.4922* | 0.247 | 0.3325 | 0.191 |
| $\gamma_2$ | | | 0.7915* | 0.380 | 0.9160* | 0.434 |
| $\log L$ | 2109.4 | | 2163.4 | | 2182.1 | |

This table reports the maximum likelihood estimates of following the model, $i = 1$ refers to the French short rate and $i = 2$ refers to the German short rate,

$$r_{it} - r_{it-1} = \mu_i + \kappa_i r_{it-1} + \varepsilon_{it}, \quad E[\varepsilon_{it}|\mathcal{F}_{t-1}] = 0, \quad E[\varepsilon_{it}\varepsilon_{jt}|\mathcal{F}_{t-1}] \equiv h_{ijt},$$

$$H_t = D_t R D_t + \Phi \circ \Theta_t,$$

where $D_t$ is a diagonal matrix with elements $\sqrt{\theta_{iit}}$, $R = [\rho_{ij}]$ is the correlation matrix, $\Phi = [\phi_{ij}]$ is a matrix of parameters with $\phi_{ii} = 0$ for all $i$, $\Theta_t = [\theta_{ijt}]$,

$$\theta_{ijt} = \omega_{ij}^* + b_i' H_{t-1} b_j + a_i' \varepsilon_{t-1}^* \varepsilon_{t-1}^{*\prime} a_j + g_{it-1}' \eta_{t-1}^* \eta_{t-1}^{*\prime} g_j, \quad \text{for all } i, j.$$

Std error denotes Bollerslev and Wooldridge (1992) robust standard errors. Log $L$ denotes the value of the log-likelihood function. **Denotes significance at the 1% level, *at the 5% level.

innovations have different impact in next period variance–covariance. The results for the ADC-Levels showed no evidence of asymmetric effects, as the associated coefficients were not statistically significant at the 5% level. The evidence regarding the asymmetric volatility effects is consistent with the univariate model's estimates in Ferreira (2000b).

### 3.3. News impact surfaces

The four standard multivariate GARCH models yield different variance and covariance estimates. To further investigate the differences between the models we use the news impact surfaces proposed by Kroner and Ng (1998), a multivariate extension of the news impact curves of Engle and Ng (1993). The news impact surfaces plot the conditional variance–covariance of each interest rate series against the last period's shocks of both series, holding the past conditional variance–covariance constant at their unconditional sample mean values.

We found significant differences between the news impact surfaces for the French variance obtained from the four standard multivariate GARCH models. The VECH and CCORR models restrict a country variance to be a function of only its own squared shocks to the interest rate process and, consequently, it does not depend on the other country interest rate shocks (the surface is flat along any line parallel to the other country shock axis). The GDC and BEKK news impact surfaces suggest that the French variance was impacted by German shocks as well as its own shocks. Conversely, the standard multivariate GARCH models yielded similar news impact surfaces for German variance. In addition, the GDC and BEKK models' news impact surfaces were consistent with French shocks having no significant impact on German variance. With respect to the covariance between French and German short rates, the VECH model news impact surface had a saddle shape. The covariance was small or even negative when past shocks to the French and German short rates were both large and of the same sign, because the coefficient in the cross-product form of the innovations was negative. This conclusion was also consistent with the GDC news impact surface.

Fig. 1 presents the news impact surfaces for the ADC-Levels model, which allows for both level and asymmetric effects. The news impact surface indicates that French variance was affected both by its own news and German news. While French news had a symmetric effect on French variance, on the other hand, the volatility spillover from Germany to France was asymmetric with German negative innovations having a larger positive impact on French variance than positive innovations. Thus, the French variance was specially affected by unexpected decreases in the German interest rates and consequently by increases in bond prices. Panel 2 of Fig. 1 reveals that German variance was unaffected by French shocks, but responded to its own news. Panel 3 shows evidence that the covariance was higher following a negative shock to the German short rate, while it was smaller following a French positive shock.

We can conclude that the alternative multivariate models imply different dynamic behaviors of both variance and covariance, in particular they are distinct in the way they incorporate past shocks. The empirical results confirm that Germany was the leader in terms of monetary policy and that France was the follower, as evidenced by our finding of volatility spillover effects from Germany to France, but not the reverse. With respect to the correlation between interest rates, the evidence indicates that the correlation was higher when German rates decreased, which could be interpreted as France taking advantage of the German central bank credibility in
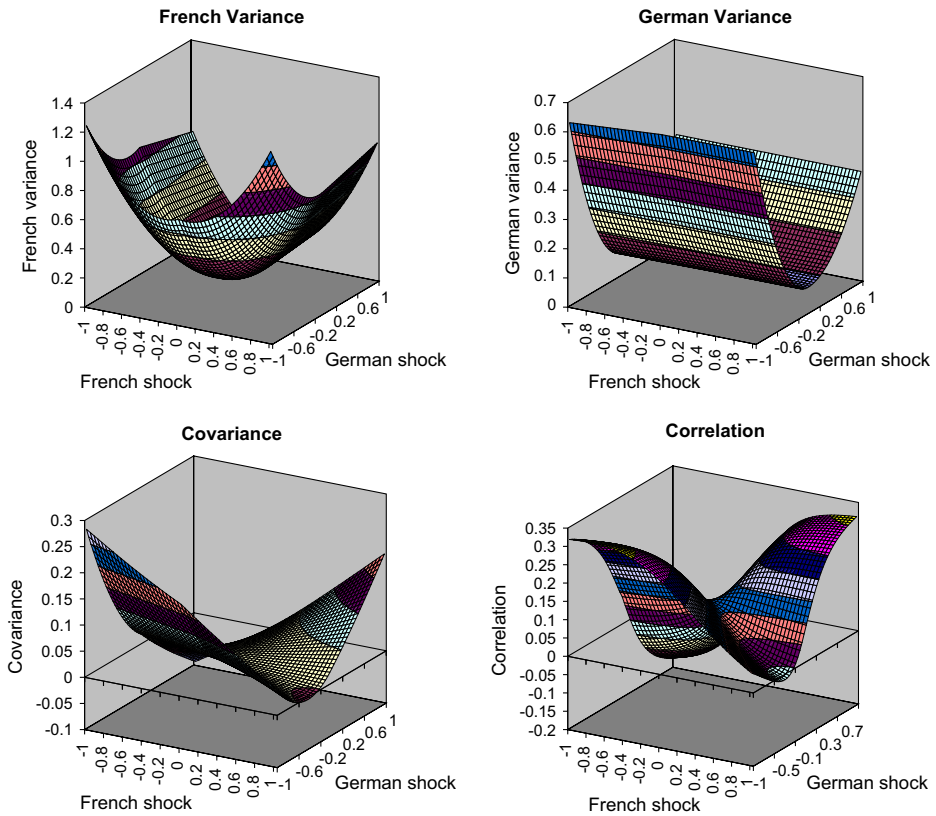
Fig. 1. The figures give the news impact surfaces for French variance, German variance, covariance and correlation between the French and German short rates under the ADC-Levels model. The news impact surface represents the variance–covariance as a function of the lagged innovations with all other terms of the variance–covariance equation fixed at their sample means.

terms of controlling inflation. On the other hand, correlation was smaller after an interest rate increase in France, which again sustains the low credibility of the French monetary authorities and the high credibility of the German ones.

## 3.4. Alternative measures of in-sample fit

The models forecasting performance evaluation require an observable proxy for the realized variance and covariance. Andersen and Bollerslev (1998) show that squared daily changes are a noisy estimate of daily realized variance and consequently this may affect the inference regarding the forecast accuracy of the models. They suggest using data sampled more frequently to obtain better estimates of the volatility dynamics, in particular using intraday data to estimate daily realized variance. Thus, we consider the variance of a week's daily changes multiplied by the

number of trading days in the week as our proxy for realized variance instead of just taking the squared weekly change. This proxy applies the idea in Andersen and Bollerslev (1998) to weekly data, and is a common procedure in studies of volatility (e.g., Schwert, 1989). We also apply the same idea to obtain realized covariance, meaning that the proxy is the covariance between a week's daily changes of both countries multiplied by the number of trading days in the week.[9] This proxy should be more accurate than simply taking the product of weekly changes of the two countries.

Another model evaluation criterion is forecasting power. Using the in-sample model parameters, we computed one-step ahead variance and covariance forecasts, $\hat{h}_{ijt}^2(1)$, for each week $t$ in our sample and for each model. Table 3 reports the root mean squared prediction error (RMSPE) given by the comparison of the forecasts with our measures of realized volatility.

The standard multivariate GARCH models for France had similar predictive power, with the exception of the GDC model, which had a slightly better performance. When including both the level and asymmetric effect in the GDC model there was an improvement in forecasting accuracy. This model (ADC-Levels) resulted in the lowest RMSPE among all models. The German variance prediction errors indicate that including a level effect in addition to a news effect considerably decreases the forecasting accuracy. The multivariate GARCH models for Germany had similar forecasting power, with the exception of the BEKK model, which generated the lowest RMSPE. In comparison with the univariate GARCH models for France and Germany in Ferreira (2000a), there was not a significant improvement in forecasting accuracy using multivariate models.

In terms of covariance forecasting, the most accurate models were the VECH-type models closely followed by the CCORR-type models. The level and asymmetric effects did not have significant additional forecasting power for the covariance. Finally, the GDC-Levels and CCORR-Levels models were the most accurate in terms of correlation forecasting. While including the level effect had a positive but small impact on forecasting accuracy, on the other hand, the asymmetric volatility effect decreased the forecasting power of the models for correlation.

## 4. Out-of-sample forecasting criteria

### 4.1. Methodology

Out-of-sample forecast accuracy provides an additional and more useful comparison of the alternative forecasts; in particular, it is a way to assess the possibility of overfitting due to a large number of parameters. Furthermore, the

---

[9] For both realized variance and covariance we assume a mean rate change of zero because the sample mean is a noisy predictor of the mean. By assuming a mean rate change of zero, the variance proxy is just the average of a week's daily squared changes (cross-product of changes for covariance) multiplied by the number of days in the week.

Table 3
In-sample RMSPE

|         | French variance | German variance | Covariance | Correlation |
|---------|-----------------|-----------------|------------|-------------|
| VECH    | 1.5522          | 0.3492          | 0.0677     | 0.4552      |
| CCORR   | 1.5596          | 0.3586          | 0.0698     | 0.4252      |
| BEKK    | 1.5679          | **0.2625**      | 0.1249     | 0.4341      |
| GDC     | **1.5110**      | 0.3577          | 0.0715     | 0.4326      |
| VECH-L  | 1.6249          | 0.6689          | **0.0675** | 0.4555      |
| A-VECH-L| 1.8299          | 0.5455          | **0.0675** | 0.4479      |
| CCORR-L | 1.6456          | 0.7186          | 0.0762     | 0.4248      |
| A-CCORR-L| 1.8341         | 0.5708          | 0.0775     | 0.4253      |
| GDC-L   | 1.5707          | 0.8344          | 0.1003     | **0.4242**  |
| ADC-L   | **1.4618**      | 0.7464          | 0.1230     | 0.4282      |

This table reports in-sample root mean squared prediction errors for the interest rates second moments at the one-week horizon. The proxy for the ex post variance is the mean of the squared daily changes of $r$ multiplied by the number of trading days in the week. The proxy for the ex post covariance is the mean of the product of the daily changes of $r$ in France and Germany multiplied by the number of trading days in the week. The model's parameters used to generate the forecasts are estimated using the 1981—1997 sample period. The minimum RMSPE forecast in each column is in bold font and the second smallest is underlined.

ability to produce useful out-of-sample forecasts is the real test of a model because it uses the same information set available to agents at each point in time. Lopez (1999) and Ederington and Guan (1999) find evidence that in-sample results are not confirmed out-of-sample for several assets and markets.

We use the fixed scheme to estimate the model's parameter vector $\theta$. In contrast with the recursive and rolling schemes,[10] which require reestimation of $\theta$ in each period, the fixed scheme estimates $\theta$ only once using data, say from 1 to $R$. The estimate is used to construct $P$ predictions with time horizon $\tau$, from period $R + \tau$ to $T + \tau$. Data realized subsequent to $R$ may be used in making these predictions. We only consider a one-week forecast time horizon,[11] $\tau = 1$. The fixed scheme is convenient because it requires less computational time, especially when estimation of the parameters is time consuming, which is the case with multivariate GARCH models. However, the fixed scheme does not use all information available to agents, consequently it may yield suboptimal forecasts when data are not stationary. However, the evidence in Ferreira (2000a) showed no significant difference between the fixed and the alternative sample schemes in terms of prediction errors and accuracy. We begin our out-of-sample forecasts at approximately the midpoint of the sample (January 2, 1990). Thus, the sample used to estimate model parameters includes 468 observations from January 6, 1981 to December 26, 1989. In our notation, $R = 468$, $P = 418$, and $T = 885$.

Since specific economic loss functions are seldom available, volatility forecast evaluation is typically conducted using a statistical loss function. We use the mean

---

[10] See West and McCracken (1998) for a more detailed description of the alternative sample schemes to estimate the parameter vector for out-of-sample analysis.

[11] The results may vary with the forecast horizon, for example, see West and Cho (1995).

squared prediction error (MSPE).[12] The use of a statistical loss function requires that we have a proxy for the realized volatility as we have explained in detail in Section 3. The prediction error is computed as the realized covariance (variance) minus the one-step ahead covariance (variance) forecast, $h_{ijt+1} - \hat{h}_{ijt}(1)$, with $h_{ijt+1}$ being the realized covariance between $\varepsilon_{it+1}$ and $\varepsilon_{jt+1}$ (when $i = j$, $h_{iit+1}$ is the variance of $\varepsilon_{it+1}$) and $\hat{h}_{ijt}(1)$ being the one-step ahead forecast of the covariance (variance when $i = j$) between $\varepsilon_{it+1}$ and $\varepsilon_{jt+1}$ formulated at time $t$, according to a model, estimated using past data $\varepsilon_{is}$, $s \leq t$.

In addition, to compare the models from a slightly different perspective, we conduct efficiency regression tests by estimating the Mincer−Zarnowitz regression by OLS for each model:

$$h_{ijt+1} = b_0 + b_1 \hat{h}_{ijt}(1) + \epsilon_{t+1}. \tag{9}$$

If the forecasts are unbiased and indeed $E_t(h_{ijt+1}) = \hat{h}_{ijt}(1)$, we would expect that $b_0 = 0$ and $b_1 = 1$, and also that $\epsilon_{t+1}$ is serially uncorrelated.

The inference about out-of-sample forecasts and forecast errors using MSPE and efficiency regression based on the conventional test statistics is not always valid, as shown in West and McCracken (1998). The reason is that the uncertainty caused by the use of an estimate of the parameter vector $\hat{\theta}$ instead of the (unknown) true $\theta$ is not necessarily taken into account by the conventional test statistic. The sampling scheme used to estimate $\theta$ determines the correction required. Under the fixed scheme, for efficiency regression, we need to rescale the least squares variance−covariance by a function ($\lambda$) of the number of prediction ($P$) and sample size used to estimate parameter vector ($R$).[13] For MSPE we use the conventional standard errors given by a Newey and West (1987) heteroskedasticity and autocorrelation robust covariance−variance matrix. We compare the forecasting accuracy of the alternative forecasts using the methodology proposed in Diebold and Mariano (1995).

## 4.2. Empirical results

Tables 4 and 5 present out-of-sample RMSPEs and efficiency regression results, respectively, for the forecast dates from January 2, 1990 through December 30, 1997. The out-of-sample results for France confirm the in-sample results that most multivariate GARCH models had similar forecasting power, which is shown by the Diebold and Mariano (1995) test results. In addition, the best model was the A-VECH-Levels, closely followed by the CCORR-Levels and A-CCORR models, which contrasted with the in-sample results where the best model was the ADC-Levels model. Models not including volatility spillover, VECH and CCORR models, outperformed the remaining models including volatility spillover. Thus, for France it

---

[12] For example, see West and Cho (1995).

[13] For the fixed scheme the conventional standard errors are too small and consequently we have too many rejections. The correct standard errors can be obtained by multiplying the conventional ones by $\sqrt{\lambda} = \sqrt{1 + P/R} = 1.3759$.

Table 4
Out-of-sample RMSPE

| | French variance | | German variance | | Covariance | | Correlation | |
|---|---|---|---|---|---|---|---|---|
| | RMSPE | DM test | RMSPE | DM test | RMSPE | DM test | RMSPE | DM test |
| VECH | 1.8268 (0.665) | 1.128 [0.259] | 0.0923 (0.019) | 1.998 [0.046] | 0.0455 (0.010) | 1.366 [0.172] | 0.4314 (0.015) | 1.829 [0.067] |
| CCORR | 1.8050 (0.547) | 1.172 [0.241] | 0.0817 (0.011) | 3.352 [0.000] | 0.0379 (0.008) | 2.307 [0.021] | 0.4247 (0.015) | 1.209 [0.227] |
| BEKK | 1.9135 (0.617) | 0.935 [0.350] | 0.2021 (0.019) | 5.060 [0.000] | 0.1213 (0.014) | 4.015 [0.000] | 0.4845 (0.018) | 4.215 [0.000] |
| GDC | 1.8285 (0.597) | 0.871 [0.384] | 0.0953 (0.011) | 4.089 [0.000] | 0.0647 (0.009) | 3.073 [0.002] | 0.4597 (0.016) | 4.374 [0.000] |
| VECH-L | 1.6044 (0.556) | 0.857 [0.391] | 0.0849 (0.011) | 3.403 [0.000] | 0.0400 (0.009) | 1.139 [0.255] | 0.4278 (0.015) | 1.407 [0.160] |
| A-VECH-L | <u>1.5568</u> (0.548) | 0.282 [0.778] | 0.1082 (0.025) | 1.860 [0.063] | 0.0411 (0.008) | 1.841 [0.066] | 0.5087 (0.019) | 5.025 [0.000] |
| CCORR-L | 1.5598 (0.565) | 0.511 [0.609] | 0.0804 (0.007) | 7.331 [0.000] | <u>0.0354</u> (0.009) | 1.849 [0.064] | **0.4210** (0.014) | |
| A-CCORR-L | 1.5598 (0.548) | 0.317 [0.751] | 0.1115 (0.026) | 1.851 [0.064] | **0.0352** (0.009) | | <u>0.4212</u> (0.014) | 0.479 [0.632] |
| GDC-L | 1.7674 (0.567) | 0.891 [0.373] | 0.1952 (0.033) | 2.769 [0.006] | 0.1011(0.027) | 1.730 [0.084] | 0.4236 (0.015) | 0.703 [0.482] |
| ADC-L | 1.8155 (0.574) | 0.940 [0.347] | 0.2186 (0.056) | 1.859 [0.063] | 0.1830 (0.054) | 1.638 [0.101] | 0.4398 (0.016) | 2.718 [0.007] |
| Equal | 1.9820 (0.762) | 1.012 [0.311] | <u>0.0445</u> (0.013) | 0.380 [0.704] | 0.0382 (0.010) | 1.068 [0.285] | 0.4241 (0.015) | 0.874 [0.382] |
| Exponential | **1.5323** (0.600) | | **0.0442** (0.013) | | 0.0542 (0.020) | 1.028 [0.304] | 0.4260 (0.015) | 1.571 [0.116] |

This table reports out-of-sample root mean squared prediction errors for the interest rates second moments at the one-week horizon. The proxy for the ex post variance is the mean of the squared daily changes of $r$ multiplied by the number of trading days in the week. The proxy for the ex post covariance is the mean of the product of the daily changes of $r$ in France and Germany multiplied by the number of trading days in the week. The model's parameters are estimated using the fixed scheme and the 1981−1989 period. The minimum RMSPE forecast in each column is in bold font and the second smallest is underlined. The standard errors are corrected for the uncertainty about the model's parameter vector as described in West and McCracken (1998). DM denotes the Diebold and Mariano (1995) test statistic on the null hypothesis of no difference in the accuracy of the forecast and the minimizing forecast with $p$-values in brackets.

Table 5
Out-of-sample efficiency regressions

| | $b_0$ | $b_1$ | $R^2$ | $\chi^2(2)$ | $b_0$ | $b_1$ | $R^2$ | $\chi^2(2)$ |
|---|---|---|---|---|---|---|---|---|
| | French variance | | | | German variance | | | |
| VECH | 0.0364 (0.145) | 0.9429 (0.695) | 0.175 | 0.0914 [0.955] | 0.0103 (0.004) | 0.0006 (0.014) | 0.051 | 11 149 [0.000] |
| CCORR | 0.0304 (0.147) | 0.9363 (0.674) | 0.195 | 0.054 [0.974] | 0.0084 (0.003) | 0.0407 (0.029) | 0.054 | 1152 [0.000] |
| BEKK | 0.0293 (0.150) | 2.1052 (1.580) | 0.132 | 1.843 [0.398] | 0.0108 (0.003) | −0.0058 (0.005) | 0.052 | 143 272 [0.000] |
| GDC | −0.0724 (0.175) | 2.8382 (1.671) | 0.295 | 2.207 [0.332] | 0.0098 (0.004) | 0.0089 (0.017) | 0.051 | 5801 [0.000] |
| VECH-L | 0.0571 (0.108) | 1.1487 (0.513) | 0.371 | 0.686 [0.710] | 0.0077 (0.003) | 0.0460 (0.035) | 0.055 | 791.5 [0.000] |
| A-VECH-L | 0.0892 (0.094) | 0.7769 (0.239) | 0.435 | 1.460 [0.482] | 0.0095 (0.003) | 0.0141 (0.017) | 0.052 | 3292 [0.000] |
| CCORR-L | 0.0727 (0.099) | 1.1551 (0.419) | 0.408 | 0.973 [0.615] | 0.0066 (0.002) | 0.0647 (0.033) | 0.056 | 805.7 [0.000] |
| A-CCORR-L | 0.0885 (0.095) | 0.7757 (0.242) | 0.433 | 1.424 [0.491] | 0.0095 (0.003) | 0.0135 (0.017) | 0.052 | 3541 [0.000] |
| GDC-L | −0.0195 (0.160) | 1.3933 (0.924) | 0.247 | 0.350 [0.840] | 0.0104 (0.003) | −0.0006 (0.005) | 0.051 | 126 764 [0.000] |
| ADC-L | 0.0114 (0.154) | 1.2634 (0.929) | 0.194 | 0.359 [0.836] | 0.0102 (0.003) | 0.0010 (0.003) | 0.051 | 215 653 [0.000] |
| Equal | 0.1610 (0.098) | 0.5723 (0.214) | 0.019 | 8.423 [0.015] | 0.0046 (0.002) | 0.4667 (0.252) | 0.021 | 4.798 [0.091] |
| Exponential | −0.0204 (0.064) | 1.1177 (0.176) | 0.411 | 0.579 [0.749] | 0.0022 (0.001) | 0.5716 (0.219) | 0.028 | 3.819 [0.148] |
| | Covariance | | | | Correlation | | | |
| VECH | 0.0066 (0.006) | −0.5400 (0.674) | 0.061 | 12.36 [0.001] | 0.0700 (0.058) | 0.1787 (0.315) | 0.052 | 11.44 [0.003] |
| CCORR | −0.0046 (0.004) | 0.7719 (0.594) | 0.069 | 12.36 [0.002] | | | | |
| BEKK | 0.0038 (0.003) | −0.0551 (0.067) | 0.030 | 4846 [0.000] | 0.0906 (0.030) | 0.0820 (0.113) | 0.053 | 65.69 [0.000] |
| GDC | 0.0031 (0.003) | −0.1274 (0.106) | 0.030 | 148.7 [0.000] | 0.1064 (0.034) | −0.0835 (0.167) | 0.051 | 42.33 [0.000] |
| VECH-L | 0.0057 (0.004) | −1.8954 (1.787) | 0.069 | 2.786 [0.248] | 0.0846 (0.036) | 0.1850 (0.351) | 0.052 | 6.760 [0.034] |
| A-VECH-L | 0.0033 (0.003) | −0.1626 (0.177) | 0.009 | 44.59 [0.000] | 0.0880 (0.031) | 0.0893 (0.095) | 0.054 | 93.85 [0.000] |
| CCORR-L | −0.0057 (0.004) | 1.5193 (0.959) | 0.178 | 3.785 [0.151] | | | | |
| A-CCORR-L | −0.0054 (0.004) | 1.5572 (0.929) | 0.190 | 3.419 [0.181] | | | | |
| GDC-L | −0.0014 (0.002) | 0.1838 (0.044) | 0.307 | 353.9 [0.000] | 0.0584 (0.070) | 0.3056 (0.526) | 0.052 | 2.626 [0.269] |
| ADC-L | 0.0002 (0.002) | 0.1026 (0.027) | 0.288 | 1066 [0.000] | 0.0860 (0.041) | 0.0824 (0.233) | 0.051 | 16.57 [0.000] |
| Equal | 0.0004 (0.002) | 0.7098 (0.537) | 0.014 | 0.368 [0.832] | 0.0679 (0.039) | 0.2619 (0.281) | 0.052 | 7.715 [0.021] |
| Exponential | 0.0050 (0.001) | −0.8592 (0.075) | 0.274 | 749.6 [0.000] | 0.0973 (0.034) | −0.0065 (0.305) | 0.050 | 11.02 [0.004] |

This table shows results of the out-of-sample efficiency regression $h_{ijt+1}^2 = b_0 + b_1 \widehat{h}_{ijt}^2(1) + \epsilon_t$ for the interest rates second moments at the one-week horizon. The proxy for the ex post variance is the mean of the squared daily changes of $r$ multiplied by the number of trading days in the week. The proxy for the ex post covariance is the mean of the product of the daily changes of $r$ in France and Germany multiplied by the number of trading days in the week. The model's parameters are estimated using the fixed scheme and the 1981−1989 period. Heteroskedasticity and autocorrelation consistent asymptotic Newey and West (1987) standard errors for $b_0$ and $b_1$ are in parentheses. The standard errors are corrected for the uncertainty about the model's parameter vector as shown in West and McCracken (1998). $\chi^2(2)$ is the test of $H_0: b_0 = 0$, $b_1 = 1$, with asymptotic $p$-values in brackets.

was more important to include level and asymmetric effects than volatility spillover because the latter did not have additional out-of-sample prediction power. In summary, the best univariate model for the French variance was a GARCH with level and asymmetric effects (GJR-Levels model)[14] and the best multivariate covariance model was the VECH model with also level and asymmetric effects (A-VECH-Levels model). This conclusion is not consistent with the in-sample results, which pointed toward a more general model like the ADC-Levels model.

The out-of-sample results for Germany were different from the ones for France along several lines. First, the out-of-sample prediction errors were smaller than the corresponding within-sample prediction errors. Second, the prediction errors were significantly smaller than the ones for France. Third, the model's forecasts were biased estimates of realized volatility with all models overpredicting realized volatility. Fourth, the models allowing for an asymmetric volatility effect performed poorly compared to symmetric volatility models. This confirmed that there was no significant asymmetric volatility effect for Germany.

The best models for Germany were the CCORR-Levels and VECH-Levels models with both models including level effect as well as news effect. The VECH and CCORR specifications are more restrictive than the GDC specification; in particular, they do not allow volatility spillovers. The out-of-sample results for Germany also showed that the only feature that distinguished the models in terms of forecasting accuracy was the existence of volatility spillovers. Similarly to France, the models not including the volatility spillover, the VECH and CCORR models, outperformed the models including volatility spillover. Thus, for Germany it was more important to include level effect than volatility spillover, and the latter did not have additional out-of-sample prediction power. This was especially true when volatility spillovers were combined with level effect, which led to severe overfitting (see the results for the GDC-Levels and ADC-Levels models). In summary, the best univariate variance model for Germany was a GARCH with level effect (GARCH-Levels model) and the best multivariate covariance models were the VECH and CCORR models, also with level effect (VECH-Levels and CCORR-Levels models, respectively). However, the news effect is more important than the level effect for out-of-sample forecasting in both a univariate and multivariate framework. These conclusions are not consistent with the in-sample results, which supported a more general model.

Table 4 also shows evidence on the out-of-sample accuracy for forecasting covariance and correlation. In contrast with the in-sample evidence, the results indicate that a model with constant correlation outperformed the alternative models and that models without volatility spillover (VECH and CCORR) had better forecasting accuracy for covariance. Adding the level effect to the multivariate GARCH models slightly improved the accuracy. In contrast, including the asymmetric effect did not sharpen forecasts. The CCORR-Levels, A-CCORR-Levels and VECH-Levels models were the only models that generated unbiased

---

[14] This result is in Ferreira (2000a) using the fixed sample scheme and under a squared loss function.

forecasts. The out-of-sample results for the correlation are consistent with the results for covariance.

We also compare the model's out-of-sample results to the equally-weighted and exponentially-weighted sample variances, covariance and correlation forecasts using information up to the forecasting week. The equally-weighted sample covariance uses 26 weeks of observations up to the forecasting week (minimum RMSPE among 26, 52 and all observations up to forecast date) and is given by $\widehat{h}_{ijt} = (1/26) \sum_{s=0}^{25} \varepsilon_{it-s} \varepsilon_{jt-s}$. The exponentially-weighted sample covariance assigns more weight to recent observations than to older ones and is given by $\widehat{h}_{ijt} = \left(\sum_{s=0}^{t-1} w^s\right)^{-1} \left(\sum_{s=0}^{t-1} w^s \varepsilon_{it-s} \varepsilon_{jt-s}\right)$ where $w$ is the weight factor (decay rate) with $0 < w \leq 1$.

The estimated mean weight factor[15] for the sample covariance was 0.9508 (mean lag is 20.3 weeks), which is similar to the one commonly used in the JP Morgan's RiskMetrics (0.94 for daily frequency and 0.97 for monthly frequency). For the correlation, the average of the estimated weight factor was 0.9828, which is consistent with a higher degree of persistence in the correlation dynamics. The average weight factor for the variances across our sample was 0.764 and 0.942 for France and Germany, respectively. Our weight factor for France is quite distinct from the one commonly used in the JP Morgan's RiskMetrics. Also, the French weight factor is consistent with the lower degree of persistence in the French volatility process found in the model's estimates of the spot interest rate volatility (see Ferreira, 2000b). The exponentially-weighted sample variance presented similar forecasting power to the best model for France, but for Germany outperformed the best models. In fact, in the case of the German variance, we only did not reject the null hypothesis at the 5% level of similar forecasting accuracy relative to the exponentially-weighted sample variance, using the Diebold and Mariano (1995) test statistic, for three of the multivariate covariance models (A-VECH-L, A-CCORR-L and ADC-L). In terms of forecasting covariance and correlation, the equally-weighted estimate was better than the exponential, but presented slightly worse forecasting power than the best covariance models.

Overall, the out-of-sample results show evidence that some of the features required to obtain good in-sample performance do not have additional out-of-sample forecasting power. They overfit the data. In particular, there is no significant improvement in forecasting power when we allow for volatility spillover and time-varying correlation. In addition, there is no significant improvement in accuracy for the German variance of including level and asymmetric effects.

---

[15] For each forecast, we estimate $w$ by minimizing the RMSPE for predetermined data in the 52 weeks prior to the forecasting week. For each week $t$, given a trial value for $w$, we compute the volatility forecast for each of the 52 weeks up to date $t$, i.e., week $t - 51$, $t - 50$,...,$t$, with all available data prior to the beginning of each of those weeks, i.e., week $t - 52$, $t - 51$,...,$t - 1$, respectively. This produces 52 forecasts and forecast errors from which RMSPE is computed. Next, we search over values for $w$ to find the value that minimizes the RMSPE criterion. This $w$ is then used on week $t$ to calculate an exponentially-weighted volatility forecast for week $t + 1$.

## 5. Economic criteria

Obtaining a correct estimate for the time-varying variance–covariance matrix is important for a wide variety of applications. Here, we apply our variance–covariance matrix estimates to three types of economic applications in portfolio selection, hedging and risk management.

First, we consider the problem of calculating the optimal portfolio allocation between the French ($i = 1$) and German ($i = 2$) interest rate sensitive claims. The problem consists of estimating the risk-minimizing portfolio weight, which is given by

$$x_{1t} = \frac{h_{22t} - h_{12t}}{h_{11t} - 2h_{12t} + h_{22t}}, \tag{10}$$

where $x_{1t}$ denotes the French portfolio weight at time $t$ ($x_{2t} = 1 - x_{1t}$ denotes the German portfolio weight). Panel A of Table 6 presents summary statistics of the in-sample estimated optimal French portfolio weights for the alternative covariance models. The average optimal weights were similar across models, with the averages

Table 6
In-sample optimal portfolio holdings and hedge ratio

| | Mean | Std dev | Minimum | Maximum | 25th Percentile | Median | 75th Percentile |
|---|---|---|---|---|---|---|---|
| *Panel A: Optimal portfolio holdings* | | | | | | | |
| VECH | 0.3789 | 0.2682 | −0.0013 | 0.9721 | 0.1370 | 0.3482 | 0.5826 |
| CCORR | 0.3786 | 0.2755 | −0.0050 | 0.9946 | 0.1306 | 0.3553 | 0.5774 |
| BEKK | 0.3801 | 0.2890 | −0.5646 | 2.1522 | 0.1487 | 0.3806 | 0.5638 |
| GDC | 0.3794 | 0.2618 | −0.0363 | 0.9443 | 0.1400 | 0.3777 | 0.5871 |
| VECH-L | 0.3624 | 0.2357 | 0.0022 | 0.9814 | 0.1621 | 0.3526 | 0.5282 |
| A-VECH-L | 0.3525 | 0.2293 | −0.0247 | 0.9650 | 0.1715 | 0.3425 | 0.4752 |
| CCORR-L | 0.3601 | 0.2452 | −0.0048 | 1.0000 | 0.1587 | 0.3617 | 0.5186 |
| A-CCORR-L | 0.3591 | 0.2385 | −0.0050 | 0.9922 | 0.1650 | 0.3645 | 0.4959 |
| GDC-L | 0.3651 | 0.2363 | −0.0108 | 0.9644 | 0.1736 | 0.3654 | 0.5246 |
| ADC-L | 0.3653 | 0.2094 | −0.0123 | 1.0230 | 0.1988 | 0.3890 | 0.5032 |
| *Panel B: Optimal risk-minimizing hedge ratio* | | | | | | | |
| VECH | 0.0613 | 0.0588 | −0.1685 | 0.3569 | 0.0156 | 0.0440 | 0.1025 |
| CCORR | 0.0955 | 0.1158 | 0.0094 | 0.7321 | 0.0376 | 0.0542 | 0.099 |
| BEKK | 0.0361 | 0.3183 | −2.9734 | 2.2942 | −0.0490 | 0.0415 | 0.1227 |
| GDC | 0.1060 | 0.1578 | −0.0061 | 1.4095 | 0.0338 | 0.0556 | 0.1027 |
| VECH-L | 0.0483 | 0.0629 | −0.1341 | 0.2819 | 0.0103 | 0.0291 | 0.0718 |
| A-VECH-L | 0.0656 | 0.0577 | −0.0003 | 0.3178 | 0.0189 | 0.0439 | 0.1057 |
| CCORR-L | 0.0886 | 0.1007 | 0.0081 | 0.8740 | 0.0401 | 0.0539 | 0.0888 |
| A-CCORR-L | 0.0889 | 0.1027 | 0.0110 | 0.9368 | 0.0429 | 0.0555 | 0.0872 |
| GDC-L | 0.0753 | 0.0726 | −0.0173 | 0.8332 | 0.0412 | 0.0518 | 0.0776 |
| ADC-L | 0.0670 | 0.0547 | −0.2556 | 0.7137 | 0.0463 | 0.0588 | 0.0760 |

Panel A reports statistics on the in-sample estimated series of optimal portfolio holdings for France which is given by $x_{1t} = (h_{22t} - h_{12t})/(h_{11t} - 2h_{12t} + h_{22t})$. Panel B gives statistics on the in-sample estimated series of risk-minimizing hedge ratio which is given by $\beta_t = h_{12t}/(h_{22t}S_t)$.

ranging from 0.3525 (A-VECH-Levels model) to 0.3801 (BEKK model). However, there were significant differences across models in terms of optimal weight dispersion, which were most evident in the standard deviation and maximum– minimum range. Moreover, when we included level and asymmetric effects in the variance–covariance specification, there was a tendency to obtain a smaller average portfolio weight and a reduction in the standard deviation of the estimated weight.

Second, we consider the problem of estimating a dynamic risk-minimizing hedge ratio for use in a cross-country interest rate spread strategy.[16] To minimize the risk of a portfolio that is long 1 French franc (FRF) in the French money market, an investor should short $\beta$ Deutschemark (DEM) in the German money market,

$$\beta_t = \frac{1}{S_t}\frac{h_{12t}}{h_{22t}}, \tag{11}$$

where $S_t$ denotes the DEM/FRF spot foreign exchange rate at date $t$. Panel B of Table 6 shows summary statistics of the in-sample estimated hedge ratios for the alternative covariance models. The average hedge ratios were quite different across models, with the averages ranging from 0.0361 (BEKK model) to 0.1060 (GDC model). The estimated hedge ratio series standard deviation and maximum– minimum range also differed significantly across models. The estimated hedge ratio was also significantly changed by the introduction of a level effect which seemed to lower the average hedge ratio. Conversely, the asymmetric volatility effects had no significant impact on the estimated hedge ratio. Thus, we should pay attention to model selection for the purpose of estimating hedge ratios.

The third application consists of evaluating the variance–covariance matrix estimates for calculating Value at Risk (VaR). Many different methods are used to estimate VaR, including historical simulation, Monte Carlo, nonparametric quantiles regressions, and delta-normal (RiskMetrics). It is not our goal to see which is the best method but only to compare our models in terms of generating accurate VaR estimates. Thus, we estimate VaR using the delta-normal method, which assumes conditional normality of returns. For a portfolio with $x_1$ invested in the asset 1 (French interest rate) and $x_2 = 1 - x_1$ invested in asset 2 (German interest rate), the VaR estimate at time $t$ under the conditional normality assumption with $(1 - \alpha)\%$ confidence level and a one-week time horizon, is given by

$$\widehat{\text{VaR}}_t = -z_\alpha\sqrt{x_1^2\widehat{h}_{11t} + x_2^2\widehat{h}_{22t} + 2x_1x_2\widehat{h}_{12t}}, \tag{12}$$

where $z_\alpha$ is the standardized normal cumulative probability distribution. For testing purposes we define an indicator variable named Hit in the following way,

$$\text{Hit}_t = I(x_1\Delta r_{1t} + x_2\Delta r_{2t} < \text{VaR}_t) \tag{13}$$

---

[16] This speculation strategy tries to make a profit from an expectation about the evolution of the interest rate spread between two countries, in this case the interest rate spread between France and Germany.

where $I(\cdot)$ is an indicator variable. If the model calculates VaR correctly, the series of the Hit variable should be unpredictable based on the past data, uncorrelated over time and have expected value equal to the desired significance level. We use the Dynamic Quantile (DQ) test proposed by Engle and Manganelli (1999), which is a Wald test of the hypothesis that all coefficients as well as the intercept are zero in a regression of the Hit variable on its past values (we use four lags) and on current VaR. A good model should produce a sequence of unbiased and uncorrelated Hit variable, so that the explanatory power of this regression should be zero.

The in-sample statistics on the estimated VaR and Hit variable are presented in Table 7. The model selection had some impact in terms of estimated VaR. The average VaR ranged from 0.3356 (BEKK) to 0.3638 (CCORR model) at the 5% level. The level and asymmetric effects tended to decrease the average VaR estimate and increase the standard deviation. The Hit rates (average of the Hit variable) were reasonably close to the target 5% level, although they tended to be slightly smaller on average, which indicated that the models were overestimating VaR. Conversely, at the 1% level, the Hit rates deviated more from the target and tended to be too large on average. This indicated that the short rate changes were too fat-tailed to be consistent with conditional normality. Looking at the autocorrelation of the Hit variable (*p*-value of the DQ test), we could see that the best performing models in relative terms were the GDC-type models followed by the CCORR models. The level and asymmetric effects did not improve the model performance in terms of DQ test. Nevertheless, we did not reject the null hypothesis that the VaR estimates had the desirable properties at the 5% level for any model, except for the BEKK model. In contrast, at the 1% level, only four models (VECH, CCORR, GDC and ADC-Levels) generated accurate VaR estimates. The out-of-sample VaR results[17] were qualitatively similar to the in-sample results. The most important differences were the worse performance of the CCORR and GDC models without the level effect and the better performance of the models with level effect, in particular at the 1% level.

## 6. Exchange rate volatility information

Our sample period includes the turmoil in the ERM in 1992–1993, when the French franc was under very strong speculative attacks in the foreign exchange market. During this period of high exchange rate volatility, the French central bank often used monetary policy as an instrument to fight against speculators. Thus, there are reasons to believe that the DEM/FRF spot exchange rate contains useful information about the variance–covariance of interest rates and that there would be additional forecasting power when extending the multivariate GARCH models to include the spot foreign exchange rate.

The data consist of DEM/FRF spot exchange rate, $S_t$, obtained from the Federal Reserve Board of Governors. The sample period and frequency coincide with our

---

[17] These results are in an Appendix that is available on request.

Table 7
In-sample Value at Risk

| | VaR (1%) | VaR (5%) | Hit rate (1%) | Hit rate (5%) | DQ test (1%) | DQ test (5%) |
|---|---|---|---|---|---|---|
| VECH | 0.5033 (0.018) | 0.3560 (0.012) | 0.0214 | 0.0434 | 9.674 [0.141] | 8.444 [0.209] |
| CCORR | 0.5145 (0.018) | 0.3638 (0.013) | 0.0225 | 0.0451 | 10.26 [0.115] | 6.261 [0.395] |
| BEKK | 0.4746 (0.012) | 0.3356 (0.017) | 0.0259 | 0.0552 | 26.88 [0.000] | 18.66 [0.005] |
| GDC | 0.4997 (0.018) | 0.3534 (0.012) | 0.0214 | 0.0462 | 9.657 [0.141] | 5.966 [0.428] |
| VECH-L | 0.4975 (0.020) | 0.3519 (0.014) | 0.0259 | 0.0474 | 15.62 [0.017] | 11.51 [0.075] |
| A-VECH-L | 0.4966 (0.021) | 0.3512 (0.015) | 0.0293 | 0.0474 | 15.23 [0.019] | 9.431 [0.152] |
| CCORR-L | 0.5129 (0.021) | 0.3627 (0.015) | 0.0271 | 0.0462 | 15.97 [0.015] | 7.457 [0.282] |
| A-CCORR-L | 0.5111 (0.022) | 0.3614 (0.016) | 0.0271 | 0.0462 | 14.86 [0.022] | 10.18 [0.119] |
| GDC-L | 0.5074 (0.021) | 0.3588 (0.015) | 0.0271 | 0.0451 | 16.08 [0.014] | 7.585 [0.271] |
| ADC-L | 0.4911 (0.020) | 0.3473 (0.014) | 0.0259 | 0.0485 | 11.47 [0.076] | 6.093 [0.414] |

VaR denotes Value at Risk and it is computed under the normality assumption at the 1% and 5% significance level and a one-week horizon. The table presents the average VaR estimate and associated standard error in parentheses. The Hit rate is the average of the indicator variable that counts the number of weeks in which our loss is greater than the VaR estimate. DQ is the Dynamic Quantile test of Engle and Manganelli (1999) with asymptotic *p*-values in brackets.

interest rate data. The logarithmic rate of return over our sample period presented an unconditional average of 0.0417% with a standard deviation of 0.4793%. We extend our bivariate model of interest rates to include the DEM/FRF rate of return as an additional explanatory variable. We also allow for asymmetric volatility effects in the exchange rate variance–covariance specification but we do not consider a level effect, which is specific to the interest rate series. The model for the logarithmic rate of return is given by,

$$\ln S_t - \ln S_{t-1} = \mu_3 + \varepsilon_{3t}, \quad E[\varepsilon_{3t}|\mathcal{F}_{-1}] = 0, \quad E[\varepsilon_{3t}\varepsilon_{jt}|\mathcal{F}_{t-1}] \equiv h_{3jt}, \text{ for } j = 1, 2, 3. \tag{14}$$

Table 8 presents the in-sample forecasting power of the multivariate models including interest rates and exchange rate data. For most models there was an improvement in forecasting accuracy in terms of variance, more significant for Germany than France. This might be the result of the exchange rate having no significant additional information for the French variance because most of the dynamics in the French interest rates was due to the exchange rate against the Deutschemark. Conversely, the exchange rate had a significant additional forecasting power for German variance, indicating that news to the exchange rate market might not have a direct impact on German interest rates. This is indirect evidence that the German exchange rate policy was not completely determined by the monetary policy (and vice versa) and that the French monetary policy was completely determined by exchange rate behavior. Thus, there is no evidence that the exchange rate contained additional information about future covariance between French and German interest rates. Out-of-sample results not shown here present a similar pattern.

Table 8
In-sample RMSPE with exchange rate information

|  | French variance | German variance | Covariance | Correlation |
|---|---|---|---|---|
| VECH | 1.5549 | 0.3422 | 0.0686 | 0.4544 |
| CCORR | 1.5601 | **0.1914** | 0.0686 | **0.4238** |
| BEKK | **1.4085** | 0.2510 | 0.1953 | 0.4999 |
| GDC | 1.4928 | 0.3894 | 0.0708 | 0.4334 |
| VECH-L | 1.6404 | 0.6891 | **0.0676** | 0.4460 |
| A-VECH-L | 1.6201 | 1.5112 | 0.0912 | 0.4613 |
| CCORR-L | 1.7086 | 0.6746 | 0.0767 | 0.4251 |
| A-CCORR-L | 1.9354 | 0.5470 | 0.0781 | 0.4257 |
| GDC-L | 1.5878 | 0.3127 | 0.0717 | 0.4283 |
| ADC-L | 1.4509 | 0.3022 | 0.1230 | 0.4298 |

This table reports in-sample root mean squared prediction errors for the interest rates second moments at the one-week horizon. The proxy for the ex post variance is the mean of the squared daily changes of $r$ multiplied by the number of trading days in the week. The proxy for the ex post covariance is the mean of the product of the daily changes of $r$ in France and Germany multiplied by the number of trading days in the week. The minimum RMSPE forecast in each column is in bold font and the second smallest is underlined. The multivariate covariance models' parameters used to generate the forecasts are estimated using the 1981–1997 sample period and including as an additional dependent variable the logarithmic weekly exchange rate returns given by the cross-rate DEM/FRF. The model for the exchange rate is the following:

$$\ln S_t - \ln S_{t-1} = \mu_3 + \varepsilon_{3t}, \quad E[\varepsilon_{3t}|\mathcal{F}_{t-1}] = 0, \quad E[\varepsilon_{3t}\varepsilon_{jt}|\mathcal{F}_{t-1}] \equiv h_{3jt}, \quad \text{for } j = 1, 2, 3$$

where $S_t$ denotes the DEM/FRF spot exchange rate on week $t$. The conditional variance–covariance equation involving the exchange rate is defined similarly to the interest rates equations, with the exception that we impose the restriction that there is no level effect for the exchange rate, $\gamma_3 = 0$.

## 7. Conclusion

Multivariate models allow us to estimate the whole variance–covariance matrix. Few studies have examined both variance and covariance dynamics, in particular the existence of asymmetric effect and volatility spillover and their economic implications. This paper adds evidence on this issue by comparing alternative multivariate GARCH models and also the asymmetric general dynamic covariance (ADC) specification proposed by Kroner and Ng (1998) that nests the most common models. We extend the ADC model to accommodate level effect in the variance–covariance specification, which is a well-known feature of the interest rate volatility dynamics.

We apply the dynamic variance–covariance matrix model to weekly short-term spot interest rates in France and Germany during the pre-euro period to study their comovements. The models' estimates show that the GDC model is different from the alternative specifications, although when including level and asymmetric effects it is not statistically different from the VECH and constant correlation (CCORR) specifications. Our results confirm the usual statement that Germany was the leader in terms of monetary police in Europe. That is, French variance is affected by its own news and German news, but German variance is unaffected by French shocks and

only responds to its own news. In addition, the covariance between French and German interest rates is higher following a negative shock to the German interest rate, while it is smaller following a positive shock to the French interest rate process.

The best performing models in-sample are not, in general, the best out-of-sample and including volatility spillover effects does not increase the out-of-sample forecasting power. The best out-of-sample model for France is the A-VECH-Levels model including level and asymmetric effects as well as news effect, but no volatility spillover. For the German variance, the best models are the CCORR-Levels and VECH-Levels, which are also the best models for forecasting the covariance between French and German interest rates. The model selection is of particular importance for calculating hedge ratios and Value at Risk. Including exchange rate information improves the forecasting accuracy for variance, especially for Germany, but does not improve the power for forecasting covariance.

## References

Amin, K., Ng, V., 1997. Inferring future volatility from the information in implied volatility in Eurodollar options: a new approach. Review of Financial Studies 10, 333–367.

Andersen, T., Bollerslev, T., 1998. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885–905.

Bollerslev, T., 1990. Modelling the coherence in short-run nominal exchange rates: a Multi-variate Generalized ARCH Model. Review of Economics and Statistics 72, 498–505.

Bollerslev, T., Wooldridge, J.M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models in time-varying covariances. Econometric Reviews 11, 143–172.

Bollerslev, T., Engle, R., Wooldridge, J., 1988. A capital asset pricing model with time-varying covariances. Journal of Political Economy 96, 116–131.

Brenner, R.J., Harjes, R., Kroner, K., 1996. Another look at models of the short-term interest rate. Journal of Financial and Quantitative Analysis 31, 85–107.

Campa, J., Chang, K., 1998. The forecasting ability of correlations implied in foreign exchange options. Journal of International Money and Finance 17, 855–880.

Chan, K., Karolyi, A., Longstaff, F., Sanders, A., 1992. An empirical comparison of alternative models of short term interest rates. Journal of Finance 47, 1209–1227.

Chapman, D., Long, J., Pearson, N., 1999. Using proxies for the short rate: when are three months like an instant? Review of Financial Studies 12, 763–806.

Day, T., Lewis, C., 1992. Stock market volatility and the information content of stock index options. Journal of Econometrics 52, 267–287.

Diebold, F., Mariano, R., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics 13, 253–263.

Ederington, L., Guan, W., 1999. Forecasting volatility. Working paper, University of Oklahoma.

Engle, R., Kroner, K., 1995. Multivariate Simultaneous Generalized ARCH. Econometric Theory 11, 122–150.

Engle, R., Manganelli, S., 1999. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. Working paper, University of California, San Diego.

Engle, R., Ng, V., 1993. Measuring and testing the impact of news on volatility. Journal of Finance 48, 987–1007.

Ferreira, M., 2000a. Forecasting spot interest rate volatility. Working paper, University of Wisconsin-Madison.

Ferreira, M., 2000b. Testing models of the spot interest rate volatility. Working paper, University of Wisconsin-Madison.

Figlewski, S., 1997. Forecasting volatility. Financial Markets, Institutions and Instruments 6.

Glosten, L., Jagannathan, R., Runkle, D., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48, 1779—1801.

Kroner, K., Ng, V., 1998. Modeling asymmetric comovements of asset returns. Review of Financial Studies 11, 817—844.

Lopez, J., 1999. Evaluating the predictive accuracy of volatility models. Working paper, Federal Reserve Bank of San Francisco.

Newey, W., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and auto-correlation consistent covariance matrix. Econometrica 55, 703—708.

Schwert, G., 1989. Why does stock market volatility change over time? Journal of Finance 44, 1115—1153.

West, K., Cho, D., 1995. The predictive ability of several models of exchange rate volatility. Journal of Econometrics 69, 367—391.

West, K., McCracken, M., 1998. Regression based tests of predictive ability. International Economic Review 39, 817—840.