Yes, implied volatilities are not informationally efficient: an empirical estimate using options on interest rate futures contracts

Giulio, Cifarelli

Dipartimento di Scienze Economiche Università di Firenze

February 2004

DIPARTIMENTO DI SCIENZE ECONOMICHE

n.137

Giulio Cifarelli

YES, IMPLIED VOLATILITIES ARE NOT INFORMATIONALLY EFFICIENT. AN EMPIRICAL ESTIMATE USING OPTIONS ON INTEREST RATE FUTURES CONTRACTS

February 2004

UNIVERSITA' DEGLI STUDI DI FIRENZE

Pubblicazione depositata a norma di legge

Implied volatility is assumed to quantify, under certain conditions, the market's expectation of the average volatility of the underlying until contract expiration. Accurate assessment here is of great relevance for reliable option pricing and profitable trading. It has been exhaustively investigated using (mostly) as benchmark the ex post realised volatility of the underlying. Daily observations involve sequential forecasts for overlapping time periods since the maturity time of the option contracts exceeds the sampling interval of the data. However, use of these forecasts raises some relevant econometric problems.

Two approaches have been used in a large number of statistical investigations based to some extent on the methods proposed in the literature of the 1980s to assess the unbiasedness and the efficiency of the forward exchange rate. The first (daily) estimation approach is set out by Canina and Figlewski (1993), Lamoureux and Lastrapes (1993) and Ané and Geman (1998) among others. It corrects for the statistical problems due to data overlapping and serial dependence in the time series of forecast errors using some seminal findings by Hansen (1982) and Hansen and Hodrick (1980, 1983). The high level of persistence of the relevant time series may, however, bring about additional estimation difficulties such as small sample bias or spurious regression distortions.

The second approach, typified by the works of Neuhaus (1995), Christensen and Prabhala (1998) and Bahra (1998), modifies the sampling procedure. It discards enough data to exactly match the maturity time of the contract with the sampling interval. The statistical difficulties mentioned above disappear and model estimation is greatly simplified. However, cutting the number of observations reduces the power of the statistical tests and the efficiency of the econometric estimates. Moreover, relevant but short lived phenomena of volatility turbulence may well be missed if, as is the case here, this procedure should require the use of a quarterly sampling interval.

This paper investigates the properties of implied volatility from options on two short-term interest rate futures contracts traded at the LIFFE, the Three Month Sterling (or Short Sterling) interest rate future contract and the Three Month Eurodeutschmark interest rate future contract. They are highly liquid and play a significant role in both interest rate and exchange rate risk hedging. Short-term interest rate implied volatility is an indicator of expectations on future short-term interest rate behaviour and is positively correlated with uncertainty on future monetary and exchange rate policy measures.

A panel data approach is proposed in order to avoid some of the difficulties mentioned above. Each daily volatility time series is indexed by the day i and the quarter t and each day left to contract expiration is seen as a distinct "unit" observed over the quarterly sample interval. Power considerations no longer apply since every daily observation of the sample is used in the estimation and, at the same time, the quarterly sampling procedure eliminates the overlapping data pitfalls. The highly distortive (and hard to manipulate) persistence properties of the daily volatility time series are appreciably attenuated.

This type of analysis improves upon previous work in the following aspects.

(i) The econometric investigation of the informational efficiency of implied volatilities is preceded by thorough examination of the statistical properties of the relevant time series. Standard problems of inference, due to the overlapping nature of the forecasts, may well be compounded by additional difficulties due to the long term dependence nature of the data.

(ii) The introduction of a panel data approach is justified in a number of respects. It is associated with the timing of the contract expiration cycle and is seen as a natural extension of previous empirical investigations with non overlapping data. Here too careful analysis of the time series properties of the pooled data determines the choice of the parameterization of the efficiency tests. Finally, panel data estimates are used as validation benchmarks for the results obtained with previous econometric procedures.

(iii) Investigation is extended across contracts labelled in different currencies. The relative efficiency of London traders in dealing with both a national and a foreign interest rate does not seem to be homogeneous.

Panel data estimates prove similar to those obtained using the levels of the overlapping time series and adjusting for the bias in the coefficient standard errors. An approach *à la* Canina and Figlewski (1993) thus seems to be justified. Our results differ from theirs, however, since our implied volatilities systematically outperform historical and GARCH volatility forecasts as volatility predictors.

The analysis is organised as follows. Section 1 derives the volatility time series and investigates their statistical properties; section 2 describes the financial and statistical pitfalls in analysis of the predictive power of implied volatility using both a standard and a pooled data set; section 3 discusses the empirical findings; section 4 concludes the paper.

1. Preliminary statistical analysis

End-of-day data on short-term interest rate derivatives traded in London, the Short Sterling, 3 Month Euromark futures and corresponding option contracts are provided by the LIFFE. The time period under investigation, from January 1, 1993 through December 31, 1997, encompasses periods of severe financial and exchange rate turbulence, such as the July - August 1993 French Franc crisis, the December 1994 - March 1995 Mexican crisis and the onset of the Asian crisis in the Summer of 1997. Contract expiration follows the standard March, June, September and December cycle. For the sake of homogeneity, the auxiliary Euromark contracts introduced from June 1994 onwards are disregarded. Each contract lasts at least nine months. One trading week before expiration the series switches into the next contract in order to minimise contract expiration biases. Continuous time series of futures prices, option prices and corresponding at-the-money implied volatilities quoted by the LIFFE are thus obtained.

Options are margined and, since the buyer is not required to pay a premium up-front, discounting of the option price is dispensed with. The value of a call is provided by a simplified version of Black's (1976) European style options formula

$$C_t = F_t N(d_1) - K_t N(d_2) \tag{1}$$

with $d_1 = \frac{1}{\sigma_t \sqrt{T}}[\ln(F_t / K_t) + (1/2)\sigma_t^2 T], \quad d_2 = d_1 - \sigma_t \sqrt{T}$

where F_t is the futures price, K_t is the strike price, σ_t is the volatility of the underlying, T is the option's time to maturity and $N(.)$ denotes the standard normal distribution function.[1] As usual in the case of interest futures option pricing the underlying instrument is the percentage interest rate i_t that is implied by the futures price F_t, where *100-* $i_t = F_t$, rather than the futures price itself.[2] Estimates of implied standard deviations with different time to expiration are drawn up with the help of a close screening of the maturity of the option contact.[3] Short term options, with a time to maturity (tenor) between zero (in reality 6 trading days) and three months (a 63 trading-day interval) are used to build the implied volatility time series.

Realised future volatility $\sigma_{t,T}$ is reproduced as follows

$$\sigma_{t,T} = \left[\frac{1}{T-t}\sum_{i=1}^{T-t} R_{t+i}^2\right]^{0.5} \tag{2}$$

where $R_t = \log(i_t / i_{t-1}).\sqrt{252}$, 252 is the number of trading days per year and T-t is the time left to the option expiration. We label R_t, the daily logarithmic relative implied interest rate change, the "return" of the underlying contract. Following suggestions by Jorion (1995) and Figlewski (1997), the mean of R_t is dropped from the standard error estimation. For each option contract the implied standard deviation will be matched with the sequence of future realised standard deviations of the return of the underlying contract until option expiration in various option pricing efficiency tests.

Following Lamoureux and Lastrapes (1993), Day and Lewis (1993) and Guo (1996), among others, the relative accuracy of implied volatility is assessed using as benchmark alternative volatility forecast proxies. Historical volatility $\sigma_{H,t}$ is the annualised standard deviation of the logarithmic interest rate changes of the previous 60 trading days and is computed as follows

$$\sigma_{H,t} = \left[\frac{1}{N} \sum_{i=1}^{N} R_{t+1-i}^{2} \right]^{0.5} \tag{3}$$

where N=60. (We have tested 90 or 120 trading day time spans with no significant improvement in forecasting accuracy.)

GARCH forecasts of daily variance are transformed into forecasts of the average daily standard deviation over the remaining life of the corresponding option using the approach set forth by Lamoureux and Lastrapes (1993). On day t a GARCH(1,1) model of the conditional variance of the return of the underlying is specified as

$$\begin{aligned} R_t &= \mu + \varepsilon_t \\ h_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2 \end{aligned} \tag{4}$$

By recursive substitution of the one period ahead variance forecast $\hat{h}_{t+1,t}^2$ the k period ahead prediction can be constructed for any k, where

$$\hat{h}^2_{t+k,t} = \omega + (\alpha + \beta)\hat{h}^2_{t+k-1,t} \qquad k > 1 \tag{5}$$

If *T-t* is the number of days left in the life of the option contract at time *t*, a GARCH forecast that is comparable to its implied standard deviation reads as

$$\sigma_{G,\mathrm{t}} = \left[\frac{1}{T-t} \sum_{k=1}^{T-t} \hat{h}^2_{t+k,t} \right]^{0.5} \tag{6}$$

The GARCH model is estimated at time *t* using data from the previous 126 trading days (6 months); on the following day the data set is shifted forward by one time period, using a rolling estimation procedure. The GARCH model is thus reestimated for every period of the sample and provides forecasts that are assumed to include information available to traders at the time the forecasts are derived. The estimation has been repeated using a 252 trading days window (12 months) with no significant change in the results. In the same way an AR(1) conditional mean parameterisation of R_t does not seem to alter the nature of the $\sigma_{G,t}$ time series.

Before estimating the realised, historical and GARCH volatility time series, we have to adjust implied interest rates for the effects of futures contracts expiration dates. As pointed out by Amin and Ng (1997) switching from one expiring contract to the successive one introduces a potential spurious shock to the interest rate (and brings about a spurious increase in R_t^2). Indeed, with an upward (downward) sloping yield curve, we should have a spurious increase (decrease) in interest rate the day which follows the change of contract. Those R_t estimates whose large absolute values are due to contract shifts at time *t* are replaced by $\left(R_{t+1} + R_{t-1}\right)/2$, the average of the adjoining returns. In this way we avoid the "ghost effects" that tend to bias volatility averages such as historical or realised volatilities.[4]

Table 1

Descriptive Statistics

	Short Sterling					Euromark				
	σ_t	$\sigma_{t,T}$	$\sigma_{H,t}$	$\sigma_{G,t}$	ISD_t	σ_t	$\sigma_{t,T}$	$\sigma_{H,t}$	$\sigma_{G,t}$	ISD_t
Mean	8.2157	10.835	12.371	12.233	12.215	7.9995	10.514	15.358	15.668	14.005
S.D.	9.2517	3.6385	3.4433	3.5346	4.1028	8.1829	3.1987	6.7037	6.3196	2.8601
Sk.	2.4102	0.9432	0.3487	0.2287	0.5089	2.7648	0.3253	1.0226	1.2521	-0.0491
Kurt.	10.461	4.6388	2.7988	2.4881	2.6963	18.689	3.0141	3.6565	3.7906	3.7336
J.B.	4119.5 [0.0]	325.99 [0.0]	27.500 [0.0]	24.604 [0.0]	58.890 [0.0]	14448 [0.0]	22.114 [0.0]	240.87 [0.0]	360.02 [0.0]	28.600 [0.0]
A.C. Lags										
1	0.095	0.964	0.991	0.989	0.964	0.145	0.961	0.981	0.976	0.902
2	0.123	0.933	0.981	0.983	0.935	0.081	0.925	0.962	0.986	0.843
3	0.067	0.907	0.970	0.977	0.911	0.057	0.891	0.943	0.966	0.808
4	0.062	0.866	0.961	0.970	0.890	0.084	0.860	0.936	0.958	0.772
5	0.039	0.833	0.952	0.964	0.872	0.107	0.829	0.925	0.952	0.736
10	0.016	0.706	0.910	0.933	0.804	0.072	0.701	0.890	0.928	0.582
20	0.055	0.533	0.822	0.869	0.708	0.050	0.522	0.811	0.881	0.351
30	0.036	0.433	0.738	0.803	0.622	0.037	0.344	0.726	0.836	0.225

Notes. *S.D.* : Standard Deviation; *Sk.* : Skewness; *Kurt.* : Kurtosis; *J.B.*: Jarque Bera normality test; *A.C.* : autocorrelation coefficient; σ_t : daily volatility; $\sigma_{t,T}$: future realised volatility; $\sigma_{H,t}$: historical volatility; $\sigma_{G,t}$: GARCH volatility forecast; ISD_t : implied volatility. Probability values are in square brackets.

Table 1 provides a preliminary description of the time series. The coefficients of skewness and kurtosis of daily volatilities $(R_t^2)^{0.5}$, measured here on a percentage basis, are appreciably large and do not seem to be compatible with a Gaussian distribution - a result corroborated by the significance of the Jarque Bera test statistics. Small but significant autocorrelation coefficients of various orders detect a degree of persistence. The statistics of the realised, historical, and GARCH volatilities provide figures that are typical for time series that are based on averaged overlapping data, i.e. smaller standard deviations and very high serial correlations. Indeed, the autocorrelation functions remain large, positive and significant at very long lags - a pattern which they share with the implied volatilities.

Standard unit root tests are reported in the upper half of table 2. They reject the null in all cases but that of σ_{Gt}, the GARCH volatility forecast time series. We are thus led to reject the unit root hypothesis in favour of a highly persistent stationary process. Implied, realised and alternative volatility forecasts are moving averages of short memory daily volatilities. Granger (1980) has shown that aggregation (averaging) of stationary short memory time series can lead to long memory. (For an heuristic proof, see Beran 1994, pages 15-16.) In the second half of table 2 are set forth estimates of the R/S (Hurst) statistic adjusted for short time dependence by Lo (1991). Long range dependence is systematically detected in the volatility time series of both contracts - a finding which justifies thorough investigation into the possibility of ARFIMA parameterisation *à la* Granger and Joyeux (1980).

Table 2

ADF Unit Root Test Statistics

S. Sterling						Euro-Mark					
	ϖ	n		ϖ	n		ϖ	n		ϖ	n
$\sigma_{t,T}$	-4.7448*	0	$\Delta\sigma_{t,T}$	-37.748*	0	$\sigma_{t,T}$	-4.6618*	0	$\Delta\sigma_{t,T}$	-35.989*	0
$\sigma_{H,t}$	-3.9472°	1	$\Delta\sigma_{H,t}$	-34.807*	0	$\sigma_{H,t}$	-3.6564*	0	$\Delta\sigma_{H,t}$	-35.090*	0
$\sigma_{G,t}$	-2.8634	0	$\Delta\sigma_{G,t}$	-67.146*	0	$\sigma_{G,t}$	-2.4286	2	$\Delta\sigma_{G,t}$	-56.430*	0
ISD_t	-3.9712*	1	ΔISD_t	-39.912*	0	ISD_t	-5.0996*	4	ΔISD_t	-43.443*	0

Notes. The ϖ test statistics are obtained from the following estimates: $\Delta x_t = \iota + \varpi x_{t-1} + \Sigma_{i=1,\ldots,n}\varphi_i \Delta x_{t-i} + e_t$ where n is selected using the Akaike Information Criterion *(AIC)*. *: significant at the 1 percent level; °: significant at the 5 percent level.

Adjusted R/S (Hurst) Test Statistics for Long Memory

	$\sigma_{t,T}$	$\sigma_{H,t}$	$\sigma_{G,t}$	ISD_t
S. Sterling	3.7830*	5.0518*	5.0867*	4.9388*
Euromark	3.2799*	4.2997*	4.6561*	3.4625*

Notes. The R/S statistic reads as $Q_T = (1/\sigma_T(b))[\max_{1\le k\le T}\sum_{j=1}^{k}(x_j - \bar{x}) - \min_{1\le k\le T}\sum_{j=1}^{k}(x_j - \bar{x})]$ where $\bar{x}$ is the sample mean and $\sigma_T(b)$ is the square root of the Newey West (long run) variance estimate with bandwidth b=7 (given by the integer part of $4(T/100)^{0.25}$) and sample size T =1260. *** :** the null of no long term dependence is rejected at the 1 percent level of significance.[5]

Table 3 shows the ARFIMA(*p,d,q*) estimates of the volatility time series obtained with the error decomposition procedure set forth by Beran (1995). They are selected – among alternative parameterisations – according to the BIC minimisation criterion. Parameter *d* reflects the long term behaviour, whereas *p*, *q*, and the corresponding AR and MA coefficients determine the short term correlation structure.[6] The range of *d* that is of interest in the context of stationary long memory modelling is $0 \le d < 1/2$. In that case the process is mean-reverting. It is stationary with long memory and is appropriate for long-term persistence modelling.

Table 3

ARFIMA(*p*,*d*,*q*) Parameter Estimates

	d	Φ_1	m	Ψ_1	BIC	LLF
S. Sterling						
$\sigma_{t,T}$	-0.0293 (0.0293)	0.9716 (0.0091)	0		3465.341	-1725.54
$\sigma_{H,t}$	0.0081 (0.0257)	0.9899 (0.0047)	0		1200.687	-593.21
$\sigma_{G,t}$	0.0146 (0.0316)	0.9944 (0.0035)	0	0.3994 (0.0336)	1374.449	-678.24
ISD_t	0.4256 (0.0846)	0.6632 (0.0764)	0	0.3435 (0.0347)	3639.420	-1809.01
Euromark						
$\sigma_{t,T}$	-0.0197 (0.0304)	0.9695 (0.0096)	0		3170.643	-1578.19
$\sigma_{H,t}$	0.7794 (0.0409)		1	-0.1571 (0.0515)	3972.738	-1979.24
$\sigma_{G,t}$	0.0064 (0.0289)	0.9940 (0.0360)	0	0.5825 (0.0271)	3907.988	-1943.29
ISD_t	0.4052 (0.0138)	0.8379 (0.0629)	0	0.5608 (0.0354)	4007.598	-2001.01

Notes. The estimates come from the zero mean volatility process $\Phi(L)(1-L)^{\delta}\,[(1-L)^{m}x_t - \mu] = \Psi(L)\varepsilon_t$, where μ is the mean of the x_t time series and $-1/2<\delta<1/2$. The difference parameter is computed as $d=\delta+m$, where the integer m indicates the number of times that x_t must be differenced to achieve stationarity. Estimated asymptotic standard errors in parentheses. *BIC*: Bayes Information Criterion; *LLF*: Log Likelihood Function.

The ARFIMA estimates in table 3 do not provide homogeneous results. The parameter d estimates of most realised, historical and GARCH volatilities are not significantly different from zero, suggesting that deviations from the mean be short memory. The associated autoregressive parameters, however, imply substantial shock persistence and (being close to one) may explain the rejection of the null of no long term dependence obtained with the R/S tests. The d parameters lie in the stationary long memory 0 - 0.5 range in the case of the implied volatilities of both contracts and in the covariance non stationarity 0.5 - 1 range in the case of the Euromark historical volatility.[7]

The estimates cited in this section suggest that the volatilities of both contracts are characterised by substantial shock persistence, but do not usually behave as random walks. Estimation in terms of levels might thus be appropriate. However, whenever the regressand and the regressors lie on the borderline separating mean and variance stationarity from non stationarity any a priori selection of the model parameterisation can be misleading, which probably explains the arbitrariness of this choice in most of the literature. Estimation in terms of levels can lead to a spurious regression bias unless the time series are cointegrated and, in terms of first differences, to a misspecification bias due to over-differencing. The latter may be costly; it tends to discard low frequency information and eliminate cointegration effects.

2. Analysis of the predictive power of volatility forecasts

Description of the different techniques introduced in the literature to assess the information content of implied volatility is followed by the discussion of a panel data approach that will be used to avoid the main econometric pitfalls outlined in the previous section.

2.1. Unbiasedness and efficiency

An interesting insight into the information content of the alternative volatility forecasting proxies is provided by the following regression

$$\sigma_{t,T} = a + b\sigma^F_{t,T} + u_{t,T} \tag{7}$$

where $\sigma_{t,T}$ is the realised volatility between time t and T and $\sigma^F_{t,T}$ is a volatility forecast derived at t over the period from t to $t+T$. We would obtain a slope of one and a zero intercept if $\sigma^F_{t,T}$ were to be an efficient and unbiased forecast of future volatility. If all the information available at time t is processed rationally, the disturbance term should be serially uncorrelated. This condition cannot be assessed with daily data because of the overlapping nature of the volatility time series. $\sigma^F_{t,T}$ can be quantified either with the implied standard deviation ISD_t provided by the LIFFE option pricing or with the volatility proxies $\sigma_{H,t}$ and $\sigma_{G,t}$ obtained manipulating the returns of the underlying.

If the option market is informationally efficient, implied volatility should incorporate all the available information about future volatility, while alternative forecast proxies should provide no additional information on the dependent variable. Forecast restrictions can then be tested with the help of the following encompassing regression *à la* Fair and Shiller (1990)

$$\sigma_{t,T} = a + b\, ISD_t + c\sigma_{P,t} + u_{t,T} \tag{8}$$

where $\sigma_{P,t}$ is either $\sigma_{H,t}$ or $\sigma_{G,t.}$

Under the null of implied volatility informational efficiency, the restriction is that $c = 0$.[8]

2.2. Previous estimation methodologies

Alternative approaches have been used in order to deal with the pitfalls that are due to the properties of the volatility time series. Canina and Figlewski (1993), Lamoureux and Lastrapes (1993) and Jorion (1995) among others estimate relationships analogous to equations (7) and (8) and disregard potential spurious regression biases. Since the length of the contracts – and thus the span of the forecasts – exceeds the sampling frequency of the data, the error term will have a moving average structure and, as pointed out by Hansen and Hodrick (1980), OLS will provide consistent, albeit inefficient estimates. The downward bias in standard errors is corrected using Hansen's (1982) GMM procedure, the lag truncation parameter of the Newey West kernel being selected in various ways. (We shall follow Lamoureux and Lastrapes here and use the well known Andrews (1991) bandwidth selection approach.)

Guo (1996) is concerned with the finite sample bias that affects the regression coefficients in the presence of highly persistent time series - a problem also discussed in Richardson and Smith (1991). He estimates equations (7) and (8) using the Fully Modified Least Squares estimator of Phillips and Hansen (1990), a procedure originally developed in order to correct for the finite sample bias which affects OLS cointegration estimates.[9]

The analyses of Scott (1992), Fleming (1993), and of ap Gwylim and Buckle (1999) posit a reformulation of equation (7) in terms of volatility differentials, thus avoiding any spurious regression effects that may be due to the unit root structure of the time series. (Fleming points out that tests in levels of the association of differing variables with the dependent variable tend to be biased in favour of the most highly correlated variable.) Economic interpretation of the results has to be adjusted accordingly as the equation estimates are meant now to assess the predictability of realised volatility changes.

Neuhaus (1995), Christensen and Prabhala (1998) and Bahra (1998) avoid these complex econometric problems by altering the sampling procedure. In order to obtain statistically independent errors they restrict the sampling to one observation per contract, and the forecast horizon thus coincides with the frequency of the data. This approach eliminates overlapping observations and has two major advantages; (i) it reduces the relevance of the long range dependence of the daily volatility time series as the data are sampled on a monthly (or quarterly) basis; (ii) it eliminates the complex moving average structure of the residuals of the regression estimates. Canina and Figlewski (1993) show, however, that this sampling procedure will raise the standard error of the slope coefficient and reduce the power of the estimation of equation (7). One additional weakness of the approach, is that a constant residual maturity (or tenor) has to be selected a priori. Indeed, it tends to vary over time, in the case of the fixed date maturity contracts examined in the paper, with potentially significant effects on the forecasts. Neuhaus (1995) points out that the effect of a diminishing maturity on the quality of volatility forecasts is twofold. As the time to expiration declines forecasts tend to become more accurate since there is less to forecast. At the same time, however, as the residual maturity of the option decreases, the contract tends to lose its option characteristic and the volatility forecasts play a declining role in price formation. Ap Gwilym and Buckle (1999) maintain that by construction the potential for forecast errors should rise as maturity declines; realised volatility becomes increasingly variable as the time to expiration decreases since the effect of a shock to the returns in equation (2) is averaged on a smaller sample size.

2.3. A panel data reformulation

A panel data procedure is set out in order to compensate for the loss of information due to the use of quarterly data and to eliminate the arbitrariness associated with the selection of a specific contract tenor. The daily time series are regrouped in 63 sets or "units" of 20 quarterly data corresponding to each fixed tenor non overlapping volatility observation over the January 1993 – December 1997 time period.
Equation (7) is then reformulated as

$$\sigma_{it,T} = a_i + b\sigma^F_{it,T} + u_{it,T} \tag{9}$$

for i = 1, 2,…, 63 (the differing times to contract expiration, which correspond, respectively, to 7, 8, …, 69 trading days to expiration) and t = 1, 2, …, 20 (the number of non overlapping quarter to quarter observations corresponding to each tenor). The residuals $u_{it,T}$ are assumed to be uncorrelated with the regressors. In this context the a_i coefficients summarise the effects of the change over time of the tenor of the contract. These relationships can be estimated using either the fixed effect or the random effect procedure.

If we assume that $a_i = a \quad \forall i$, i.e. that tenor variation has no effect on the interpretation of realised volatility, OLS pooled regression will provide consistent and efficient estimates of the coefficients in (9), a^t and b^t. If we assume that the unit tenor specific effect is constant over time but differs across units, OLS estimates of *b* can be obtained using a partitioned regression procedure. The following OLS regression on transformed data provides the "within units" estimator b^w

$$(\sigma_{it,T} - \overline{\sigma}_{i.,T}) = b(\sigma^F_{it,T} - \overline{\sigma}^F_{i.,T}) + (u_{it,T} - \overline{u}_{i.,T}) \tag{10}$$

where

$\overline{\sigma}_{i.,T} = \frac{1}{20}\sum_{t=1}^{20}\sigma_{it,T}$, $\overline{\sigma}^F_{i.,T} = \frac{1}{20}\sum_{t=1}^{20}\sigma^F_{it,T}$, and $\overline{u}_{i.,T} = \frac{1}{20}\sum_{t=1}^{20}u_{it,T}$.

For each group *i* an estimator of a_i corresponds to the mean residual $a_i = \overline{\sigma}_{i.,T} - b^w\overline{\sigma}^F_{i.,T}$.

Alternatively, we can run a regression in terms of the 63 group means

$$\overline{\sigma}_{i.,T} = a + b\overline{\sigma}^F_{i.,T} + \overline{u}_{i.,T} \tag{9'}$$

and compute the "between units" estimator of *b*, b^b.

It can be shown (Greene, 1993, pages 471-473) that the estimator b^t is a weighted sum of b^w and b^b.

$$b^t = \frac{S^w}{S^w + S^b} b^w + \frac{S^b}{S^w + S^b} b^b \tag{11}$$

where S^w and S^b are, respectively, the within units and the between units sums of squares that enter the OLS estimation of b^w and b^b.

A second "random effect" estimation procedure posits that the individual tenor specific effects be random variables. The tenor (and those factors which may influence the realised volatility but are not captured in the model specification) can be summarised by a random disturbance. A group specific error term α_i is thus added to the non specific error term $u_{it,T}$ and model (9) is rewritten as

$$\sigma_{it,T} = a + b\sigma^F_{it,T} + w_{it,T} \tag{12}$$

where

$w_{it,T} = \alpha_i + u_{it,T};\qquad E(\alpha_i) = E(u_{it,T}) = 0 \qquad \forall i,t$

$\text{var}(\alpha_i) = \sigma^2_\alpha;\ \text{var}(u_{it,T}) = \sigma^2_u \qquad \forall i,t$

$\text{cov}(\alpha_i, \alpha_j) = 0 \qquad \forall i \neq j;$

$\text{cov}(u_{it,T}, u_{js,T}) = 0 \qquad \forall i \neq j, \quad t \neq s$

and $\text{cov}(\sigma^F_{it,T}, \alpha_i) = 0$, $\text{cov}(\alpha_i, u_{it,T}) = 0$

Efficient estimators of model (12) are obtained using GLS. The appropriate transformation of the variables reads as

$$\tilde{\sigma}_{it,T} = \sigma_{it,T} - (1-\lambda^{0.5})\bar{\sigma}_{i.,T}$$
$$\tilde{\sigma}^F_{it,T} = \sigma^F_{it,T} - (1-\lambda^{0.5})\bar{\sigma}^F_{i.,T}$$

where $\lambda = \dfrac{\sigma_u^2}{\sigma_u^2 + z\sigma_\alpha^2}$, z being the number of observations of the ith unit. A GLS estimator is obtained running the following OLS regression

$$\tilde{\sigma}_{it,T} = \mu + b\tilde{\sigma}^F_{it,T} + \varepsilon_{it,T} \tag{13}$$

Here too the GLS estimate of *b*, b^G, is a weighted average of the between and within units estimates and is formulated as

$$b^G = \frac{S^w}{S^w + \lambda S^b} b^w + \frac{\lambda S^b}{S^w + \lambda S^b} b^b \tag{14}$$

Whenever λ is different from one, the standard pooled sample OLS estimator b^t set forth in equation (11) will prove inefficient because of an incorrect weighting of the b^w and b^b estimates. The entire variability of $\sigma_{it,T}$ is explained in terms of the variation of $\sigma^F_{it,T}$, whereas the appropriate procedure would be to attribute a fraction of it to random variations across units associated with the variation of α_i. OLS thus places excessive weight on between groups variation.

Two extreme cases are of interest. If $\lambda = 1$ (and $\sigma^2_\alpha = 0$) $b^G = b^t$ from equation (11) and the standard pooled sample fixed effect regression model applies. If $\lambda = 0$ (and either $\sigma_u^2 = 0$ or $z \to \infty$) the GLS estimator coincides with the fixed effect within units estimator since $b^G = b^w$. In the intermediate case, provided that the consistency pre-conditions are satisfied, the GLS (random effect) estimator will be more efficient than the fixed effect within units OLS counterpart. The λ estimates will therefore be used, along with the standard Wu-Hausman test for orthogonality of the random effects and the regressors, in order to assess whether a fixed effect or a random effect modelling procedure is appropriate for our panel.[10] The corresponding encompassing model reads as

$$\tilde{\sigma}_{it,T} = \mu + b\widetilde{ISD}_{it} + c\tilde{\sigma}_{P,it} + \varepsilon_{it,T} \tag{15}$$

where $\sigma_{P,it}$ is either the historical or the GARCH out of sample volatility forecast.

3. Empirical analysis: results and discussion

The purpose of this section is to compare daily overlapping and non overlapping analyses of the information content of implied volatility. As usual the tests are based on the joint hypothesis that the LIFFE option market is informationally efficient and that the selected option pricing model is correct.

3.1. Overlapping daily data estimation

The estimates of equation (7) with daily overlapping observations are set out in table 4 (rows 1 to 3 and 7 to 9). What is being tested here is not only the unbiased expectations hypothesis, but also the less stringent proposition that implied volatility forecasts contain some information about realised future volatility. The relevance of this hypothesis is then assessed using the $\sigma_{H,t}$ and $\sigma_{G,t}$ volatility forecasts. Implied volatilities seem to be biased predictors of future realised volatility as the corresponding null hypothesis ($H_o : a = 0, b = 1$) is consistently rejected. An analogous result is obtained with historical and GARCH out of sample volatility forecasts across both the Short Sterling and Euromark contracts. The positive intercept and the slope coefficient less than one are in line with previous findings in the literature; they suggest that realised volatility is usually underpredicted in the low variance periods and overpredicted in the high variance ones. The results are not homogeneous, however, as the implied volatility coefficient estimate is larger for the Short Sterling than for the Euromark contract, which seems to be more biased. Similarly, the adjusted R^2 coefficient suggests that the explanatory power of implied volatility is higher for the former than for the latter. This ranking is corroborated by the properties of the remaining volatility forecasts; Euromark regressors having small or even insignificant coefficients, with very low explanatory power. On the whole implied volatilities tend to outperform historical volatilities which, in turn, dominate the GARCH out of sample forecasts.[11]

Table 4

$$\sigma_{t,T} = a + b\,\sigma^F_{t,T} + u_{t,T} \qquad (7)$$

	$\sigma^F_{t,T}$	a	b	$\overline{R}^2$	*S.D.*	*D.W.*	*LB(5)*	*Arch(1)*	W_0
S.Sterling									
OLS [1]	ISD_t	5.571 (0.974)	0.431 (0.064)	0.235	3.181	0.099	4607.3 [0.00]	886.78 [0.00]	99.406 [0.00]
OLS [2]	$\sigma_{H,t}$	6.281 (1.509)	0.368 (0.119)	0.121	3.412	0.079	4929.0 [0.00]	851.27 [0.00]	31.545 [0.00]
OLS [3]	$\sigma_{G,t}$	6.447 (1.472)	0.357 (0.108)	0.120	3.414	0.082	4941.9 [0.00]	849.14 [0.00]	41.451 [0.00]
FD [4]	ISD_t	0.002 (0.027)	0.141 (0.258)	0.023	0.957	2.109	8.127 [0.15]	0.307 [0.58]	1105.4 [0.00]
FD [5]	$\sigma_{H,t}$	0.003 (0.027)	0.324 (0.069)	0.016	0.960	2.108	9.305 [0.10]	0.755 [0.38]	95.261 [0.00]
FD [6]	$\sigma_{G,t}$	0.001 (0.027)	0.048 (0.061)	0.0005	0.968	2.136	9.870 [0.08]	0.649 [0.421]	242.64 [0.00]
Euromark									
OLS [7]	ISD_t	5.268 (1.972)	0.375 (0.134)	0.111	3.015	0.098	4749.9 [0.00]	925.84 [0.00]	282.34 [0.00]
OLS [8]	$\sigma_{H,t}$	9.212 (0.968)	0.085 (0.049)	0.031	3.149	0.075	4999.0 [0.00]	955.90 [0.00]	443.97 [0.00]
OLS [9]	$\sigma_{G,t}$	9.443 (0.539)	0.068 (0.029)	0.017	3.171	0.072	5044.2 [0.00]	960.33 [0.00]	1732.5 [0.00]
FD [10]	ISD_t	-0.008 (0.024)	0.055 (0.019)	0.006	0.858	2.026	5.106 [0.40]	0.016 [0.90]	2352.4 [0.00]
FD [11]	$\sigma_{H,t}$	-0.008 (0.024)	0.038 (0.020)	0.002	0.859	2.028	4.361 [0.50]	0.020 [0.89]	2109.6 [0.00]
FD [12]	$\sigma_{G,t}$	-0.007 (0.024)	0.116 (0.018)	0.031	0.847	2.015	8.383 [0.14]	0.128 [0.72]	2309.4 [0.00]

Notes. *FD*: First differences OLS estimates; *LB(x)*: Ljung-Box Q statistic for xth order serial correlation; *Arch(x)*: LM test for xth order ARCH; W_0 : Wald test χ^2 statistic for the null hypothesis that $a = 0$, $b = 1$. Probabilities are in square brackets and standard errors in parentheses. The standard errors of the levels estimates are robust to heteroskedasticity.

The estimates are characterised by a high serial correlation of the residuals commonly attributed to the overlapping nature of the data and a consistent estimator of the variance covariance matrix of the residuals is obtained using the GMM procedure mentioned above. In four out of six regressions, however, the coefficients of multiple correlation are larger than the corresponding D.W. statistics. These findings, combined with the persistence of the regressors and of the regressands detected in section 1 suggest that a spurious regression bias *à la* Granger and Newbold (1974) might affect the estimation. The analysis was thus performed in terms of first differences, following the standard cure suggested in Hamilton (1994, page 562). The null that the constant term is nil and the slope of the regressor one is rejected throughout (see rows 4 to 6 and 10 to 12). No evidence is found of serial correlation and, as expected in the analysis of first differences, multiple correlation coefficients become very small, especially those of the Euromark contract estimates. Finally, the value of the slope coefficients tends to drop significantly, implied standard deviations being more affected by this reduction than the alternative volatility forecasts. The quality of the estimates thus shows a decline, bearing out the criticisms of Christensen and Prabhala (1998), as indeed does their financial relevance, changes in realised volatilities being more difficult to forecast than the corresponding levels. Unfortunately, because of the

Table 5

$$\sigma_{t,T} = a + bISD_t + c\,\sigma_{P,t} + u_{t,T} \qquad (8)$$

	$\sigma_{P,t}$	a	b	c	$\overline{R}^2$	*S.D.*	*D.W.*	*Arch(1)*	W_1
S.Sterling									
OLS [1]	$\sigma_{H,t}$	5.6920 (1.0931)	0.4466 (0.0969)	-0.0253 (0.1179)	0.2351	3.1821	0.1009	886.94 [0.00]	105.37 [0.00]
OLS [2]	$\sigma_{G,t}$	5.1959 (1.1277)	0.3944 (0.0817)	0.0668 (0.1047)	0.2374	3.1774	0.0979	884.49 [0.00]	103.64 [0.00]
FD [3]	$\sigma_{H,t}$	0.0055 (0.0267)	0.1488 (0.0256)	0.3467 (0.0685)	0.0415	0.9480	2.0845	0.3449 [0.56]	1152.7 [0.00]
FD [4]	$\sigma_{G,t}$	0.0024 (0.0271)	0.1407 (0.0259)	0.0264 (0.0606)	0.0220	0.9576	2.1115	0.3102 [0.58]	1104.8 [0.00]
Euromark									
OLS [5]	$\sigma_{H,t}$	4.0555 (1.1647)	0.3546 (0.0821)	0.0648 (0.0278)	0.1289	2.9853	0.0985	926.05 [0.00]	100.09 [0.00]
OLS [6]	$\sigma_{G,t}$	4.5338 (1.1698)	0.3650 (0.0817)	0.0554 (0.0271)	0.1227	2.9960	0.0971	918.10 [0.00]	87.526 [0.00]
FD [7]	$\sigma_{H,t}$	-0.0076 (0.0242)	0.0537 (0.0195)	0.0355 (0.0203)	0.0072	0.8573	2.0210	0.0162 [0.90]	2359.3 [0.00]
FD [8]	$\sigma_{G,t}$	-0.0069 (0.0239)	0.0506 (0.0192)	0.1147 (0.0183)	0.0349	0.8452	2.0087	0.1404 [0.71]	459.58 [0.00]

Notes. *FD*: First differences OLS estimates; *Arch(x)*: LM test for xth order ARCH; W_1 : Wald test χ^2 statistic for the null hypothesis that $a = c = 0$ and $b = 1$. Probabilities are in square brackets and standard errors in parentheses. The standard errors of the levels estimates are robust to heteroskedasticity.

borderline nature of the stationarity of most volatility time series, we are unable to discriminate a priori between the level and first difference parameterisations. Also the encompassing regression estimates of equation (8), set out in table 5, provide results that differ according to the parameterisation of the regressions. Levels estimates suggest that only the Short Sterling implied standard deviations contain relevant information on future realised volatility that cannot be duplicated by the alternative volatility proxies. (The coefficient estimates of the latter are not significantly different from zero.) First difference estimates find that historical and GARCH out of sample volatility changes have, in the case of the Short Sterling and Euromark contracts respectively, appreciable explanatory power.

3.2. Daily analysis using panel data

The daily time series are regrouped in 63 sets of 20 quarterly data corresponding to each fixed tenor non overlapping observation over the whole sample. Three tenors of 60, 40 and 20 days are selected at first - following Neuhaus (1995) – and three sets of OLS regressions of the realised volatility on the implied standard deviation are accordingly performed. The null of informational efficiency requires that *a* = 0, *b* = 1 and that the residuals be serially uncorrelated in

$$\sigma_{jt,T} = a + bISD_{jt} + u_{jt,T} \tag{16}$$

where j = 20, 40 and 60 days to expiration of the option contract.

Table 6

$$\sigma_{jt,T} = a + bISD_{jt} + u_{jt,T} \qquad (16)$$

Short Sterling

j	a	b	$\overline{R}^2$	$S.E.$	$D.W.$	$LB(5)$	W_o
20	4.622 (1.895)	0.478 (0.159)	0.297	2.803	1.626	1.363 [0.92]	7.273 [0.00]
40	4.075 (2.576)	0.577 (0.207)	0.262	2.962	2.531	22.363 [0.00]	3.218 [0.06]
60	5.167 (1.533)	0.459 (0.107)	0.476	2.118	1.940	11.703 [0.04]	23.873 [0.00]
Euromark							
20	0.225 (3.403)	0.763 (0.248)	0.309	2.614	1.041	9.674 [0.08]	35.339 [0.00]
40	9.312 (3.999)	0.088 (0.294)	0.050	2.988	1.436	2.767 [0.74]	14.469 [0.00]
60	6.226 (3.267)	0.326 (0.208)	0.071	2.777	1.960	8.812 [0.12]	27.975 [0.00]

Notes. *LB(x)*: Ljung Box Q-statistic for xth order serial correlation; W_o: Wald test χ^2 statistic for the null hypothesis that $a = 0$, $b = 1$. Probability values are in square brackets and standard errors in parentheses.

The estimates of table 6 show that the quality of fit is far from homogeneous across tenors. In the case of the Euromark contract the explanatory power of implied volatility seems to rise as the time to expiration declines, whereas in the case of the Short Sterling the opposite seems to obtain. Arbitrary choice of a tenor may thus affect the results erratically. As expected, standard errors tend to be much larger than in the corresponding estimates with overlapping data, and the coefficient estimates less accurate. With panel data analysis the information provided by all the 63 tenors entering the data set can be exploited and these pitfalls overcome.

Preliminary analysis using the truncated cross sectionally adjusted IPS unit root test of Pesaran (2003) set forth in table 7 shows that the time series entering the panel data set are always stationary. R/S (Hurst) tests were performed for each volatility time series. No evidence

Table 7

Truncated Cross Sectionally Augmented IPS Panel Unit Root Test Statistics

	$\sigma_{it,T}$	$\sigma_{H,it}$	$\sigma_{G,it}$	ISD_{it}
Short Sterling	-3.8487[a]	-6.3452[b]	-4.9409[a]	-5.3399[a]
Euromark	-3.3460[a]	-5.8299[a]	-5.9411[a]	-3.9623[b]

Notes. The tests of the unit root hypothesis are based on the t-ratios of the OLS estimation of b_i in the following Cross Sectionally Augmented DF (CADF) regressions

(a) No intercept, no trend $\Delta y_{it} = b_i y_{i,t-1} + c_i \overline{y}_{t-1} + d_i \Delta\overline{y}_t + e_{it}$

(b) Intercept only $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \overline{y}_{t-1} + d_i \Delta\overline{y}_t + e_{it}$

where, for i = 1, 2,…, N and t = 1, 2,…, T, y_{it} is an observation on the *ith* cross section unit at time t and $\overline{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$ is the cross section mean of y_{it} . The no intercept specification (a) is selected whenever, in a preliminary panel estimation of equation (b), the intercept a_i is not significantly different from zero.

The truncated version of the cross sectionally augmented IPS statistic (Im et al. 2003) reads as

$$CIPS^* = N^{-1}\sum_{i=1}^{N} t_i^*(N,T)$$

where $t_i^*(N,T)$ is the truncated t-ratio of b_i in the CADF regressions (a) and (b) above. The 5 percent critical values set forth in Pesaran (2003, tables 3a and 3b) are, respectively, –1.535 and –2.105.[12]

Average Adjusted R/S (Hurst) Test Statistics for Long Memory

	$\sigma_{it,T}$	$\sigma_{H,it}$	$\sigma_{G,it}$	ISD_{it}
S. Sterling	1.2221	1.2872	1.2105	1.2329
Euromark	1.2053	1.1484	1.1183	1.1123

Notes. The bandwidth is 2 in each of the 63 tests.

of long run dependence was detected as the statistics never proved significant. Their average values over the 63 sets of non overlapping observations that correspond to each volatility are set out at the bottom of table 7. Estimation in terms of volatility levels thus seems to be justified.

Panel estimates of model (13) are set out in table 8. The random effect GLS estimator was selected, assuming that the individual tenor specific effects be random variables. The consistency pre-conditions are

Table 8

$$\widetilde{\sigma}_{it,T} = \mu + b\widetilde{\sigma}^{F}_{it,T} + \varepsilon_{it,T} \qquad \textbf{\textit{(13)}}$$

$$\widetilde{\sigma}_{it,T} = \mu + b\widetilde{ISD}_{it} + c\widetilde{\sigma}_{P,it} + \varepsilon_{it,T} \qquad \textbf{\textit{(15)}}$$

$$(\sigma_{it,T} - \overline{\sigma}_{i.,T}) = b(\sigma^{F}_{it,T} - \overline{\sigma}^{F}_{i.,T}) + (u_{it,T} - \overline{u}_{i.,T}) \qquad \textbf{\textit{(10)}}$$

Model (13)	$\widetilde{\sigma}^{F}_{it,T}$	μ	b	$\overline{R}^2$	*S.D.*	*D.W.*	λ	*W.H.*	*W.C.H.*
S.Sterling									
[1]	$\widetilde{ISD}_{it}$	5.579 (0.283)	0.431 (0.021)	0.243	3.192	2.131	0.658	0.075 [0.78]	3.143 [0.06]
[2]	$\widetilde{\sigma}_{H,it}$	6.868 (0.360)	0.324 (0.027)	0.094	3.493	1.854	0.544	0.861 [0.35]	0.050 [0.82]
[3]	$\widetilde{\sigma}_{G,it}$	6.945 (0.351)	0.320 (0.027)	0.098	3.487	1.790	0.578	0.055 [0.81]	2.748 [0.09]
Euromark									
[4]	$\widetilde{ISD}_{it}$	5.156 (0.424)	0.383 (0.029)	0.117	3.013	1.612	0.830	0.044 [0.99]	28.434 [0.00]
[5]	$\widetilde{\sigma}_{H,it}$	9.139 (0.226)	0.090 (0.013)	0.035	3.151	1.600	0.692	0.704 [0.40]	19.239 [0.00]
[6]	$\widetilde{\sigma}_{G,it}$	9.707 (0.241)	0.052 (0.014)	0.010	3.191	1.550	0.727	0.012 [0.91]	23.808 [0.00]

Model (15)	$\widetilde{\sigma}_{P,it}$	μ	b	c	$\overline{R}^2$	*S.D.*	*D.W.*	λ	*W.H.*
S.Sterling									
[7]	$\widetilde{\sigma}_{H,it}$	5.726 (0.336)	0.446 (0.028)	-0.027 (0.034)	0.244	3.192	2.128	0.660	0.142 [0.93]
[8]	$\widetilde{\sigma}_{G,it}$	5.254 (0.338)	0.405 (0.026)	0.052 (0.030)	0.245	3.190	2.126	0.651	0.525 [0.77]
Euromark									
[9]	$\widetilde{\sigma}_{H,it}$	4.469 (0.439)	0.359 (0.029)	0.066 (0.012)	0.136	2.981	1.663	0.818	0.265 [0.87]
[10]	$\widetilde{\sigma}_{G,it}$	4.564 (0.460)	0.377 (0.029)	0.043 (0.013)	0.124	3.001	1.625	0.835	0.075 [0.96]

Model (10)	$\sigma^{F}_{it,T}$		b		$\overline{R}^2$	*S.D.*	*D.W.*
[11]	ISD_{it}		0.421 (0.030)		0.126	3.076	1.718
[12]	$\sigma_{H,it}$		0.111 (0.009)		0.005	3.199	1.725
[13]	$\sigma_{G,it}$		0.074 (0.010)		0.027	3.245	1.657

Notes. *W.H.*: Wu-Hausman Wald test for the orthogonality of the random effects and the regressors; $\lambda = \dfrac{\sigma_u^2}{\sigma_u^2 + z\sigma_\alpha^2}$, z being the number of observations of the ith unit; *W.C.H.:* White's cross sectional heteroskedasticity test. Probability values in square brackets.

satisfied since the Wu-Hausman tests for orthogonality are not significant and, at the same time, the λ ratio estimates lie in the 0-1 interval, indicating that the GLS estimates do differ from the less efficient fixed effect within units or pooled data regression counterparts.

The findings are more accurate (the coefficient standard errors being much smaller) than those set forth in table 6, computed with quarterly non overlapping data. They are broadly similar to the estimates of equation (7) of table 4, obtained with daily overlapping data in terms of levels and using the GMM procedure in order to adjust the standard errors for the serial correlation of the residuals. We can thus conclude that spurious regression effects do not seem to affect daily estimates, in spite of the persistence detected in the time series. The quality of the Short Sterling estimates (rows 1 to 3) is more satisfactory, the residuals being homoskedastic and serially uncorrelated throughout. Here too implied volatilities are more efficient, as realised volatility predictors, than the historical and GARCH out of sample proxies. Encompassing regressions corroborate the finding of a differing information content of the volatilities across contracts; the coefficients of the historical and GARCH out of sample forecasts are not significantly different from zero in the case of the Short Sterling only (rows 7 and 8). Euromark implied volatility forecast errors (rows 9 and 10) do not seem to be orthogonal to the market information set. These results, too, are reasonably similar to the daily overlapping data level estimates of table 5.

The implementation of the random effect GLS estimation procedure might be inappropriate in the case of the Euromark contract since

White's cross sectional heteroskedasticity test statistics are significant.[13] In rows 11, 12, and 13 are set out GLS within units fixed effect estimates of model (10) adjusted for cross section heteroskedasticity. (Each unit equation is weighted by an estimate of the cross section standard deviation obtained from a first step pooled data OLS regression.) White's estimator is also used in order to obtain coefficient standard errors that are robust to heteroskedasticity within each cross section unit. The results do not differ in a significant way from those of rows 4, 5, and 6 and corroborate the previous conclusions on the reliability of overlapping data levels estimates.[14]

The weights that enter the fixed effect pooled data and random effect slope estimators b^t and b^G in equations (11) and (14) are finally computed for both implied volatilities in order to analyse the homogeneity of the panel, they are set out in table 9.

Table 9

***ISD$_{it}$* Slope Coefficient Decomposition**

$$b^t = \frac{S^w}{S^w + S^b} b^w + \frac{S^b}{S^w + S^b} b^b \qquad (11) \qquad\qquad b^G = \frac{S^w}{S^w + \lambda S^b} b^w + \frac{\lambda S^b}{S^w + \lambda S^b} b^b \qquad (14)$$

	$\frac{S^w}{S^w + S^b}$	$\frac{S^b}{S^w + S^b}$	$\frac{S^w}{S^w + \lambda S^b}$	$\frac{\lambda S^b}{S^w + \lambda S^b}$
S.Sterling	0.9630	0.0370	0.9753	0.0246
Euromark	0.9261	0.0739	0.9475	0.0524

It turns out that most of the variation is within units. Panel data analysis suggests that the relevance of shifts in volatility over the estimation interval (which bring about the within units variation) is very large compared to the impact of changes in contract tenor. These findings hold for both the fixed and random effects estimation procedures even if, as expected, the weight attributed to the between units slope estimator is lower in the case of the GLS random effect estimation approach.

4. Conclusion

The accuracy of volatility forecast estimators was assessed using daily overlapping and non overlapping observations on two major short-term interest rate futures contracts traded in London. The use of a panelized data set has eliminated some of the drawbacks usually associated with non overlapping data estimation, such as the lack of accuracy due to an insufficient number of observations or the arbitrariness of the choice of tenor. Non stationarity and long memory characteristics of daily overlapping time series are disposed of in the same way, along with their potential distortive effects.

The empirical findings suggest that information content estimation in levels associated with the Hansen (1982) variance covariance matrix estimator provides reasonably accurate results, broadly similar to the corresponding benchmark panel data estimates. The criticisms of estimation in levels with overlapping data advanced by Christensen and Prabhala (1998) and ap Gwilym and Buckle (1999) among others do not seem here to be justified.

Implied volatility is not an unbiased and efficient predictor of future realised volatility. (Unbiasedness and efficiency are obtained mostly in non overlapping data studies and are possibly due to an inappropriate selection of contract tenor.) It has consistently more explanatory power than both the historical and GARCH out of sample forecasts - a ranking that reproduces the results of Jorion (1995), Fleming et al. (1995) and Guo (1996). Figlewski (1997) suggests that the inefficiency of implied volatility is not due to a lack of information. Rather, it is due to the difficulty in using it, i.e. in implementing the arbitrage trading that would bring option prices back to equilibrium after a disturbance. Expensive and risky with stock or stock index option contracts, this arbitrage is much easier to perform with currency or with futures option contracts. Indeed, most findings presented in the literature suggest that implied volatilities extracted from foreign exchange or from futures options have a greater information content.

The results of this paper can be interpreted according to this paradigm. The implied volatility from the Short Sterling contracts is more accurate as a future volatility predictor than implied volatility from the 3 Month Euromark contracts. Implied volatilities have thus a country-specific pattern as LIFFE traders, reacting to the same inflow of information, seem to be more proficient in predicting domestic than foreign interest rate volatility. This finding can be attributed to the greater cost (and associated risk) of options arbitrage trading across currencies.

Bibliography

Alexander C. O. (2001), “Market Models, A Guide to Financial Data Analysis”, John Wiley & Sons, Sussex.

Alexander C. O. and C. T. Leigh (1997), “On the Covariance Matrices used in Value-at-Risk Models”, Journal of Derivatives, 4, 50-62.

Amin K. and V. K. Ng (1997), “Inferring Future Volatility from the Information in Implied Volatility in Eurodollar Options: A New Approach”, Review of Financial Studies, 10, 333-367.

Andrews D. W. K. (1991), “Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation”, Econometrica, 59, 817-858.

Ané T. and H. Geman (1998), “Stochastic Volatility and Transaction Time: An Activity - Based Volatility Estimator”, mimeo, “Forecasting Financial Markets” 1998 Conference, London.

Ap Gwilym O. and M. Buckle (1999), “Volatility Forecasting in the Framework of the Option Expiry Cycle”, European Journal of Finance, 5, 73-94.

Bahra B. (1998), “Implied Risk–Neutral Probability Density Functions from Option Prices: A Central Bank Perspective”, in “Forecasting Volatility in the Financial Markets”, J. Knight and S. Satchell (eds.), Butterworth – Heinemann, Oxford.

Beran J. (1994), Statistics for Long Memory Processes. Chapman and Hall, New York.

Beran J. (1995), “Maximum Likelihood Estimation of the Differencing Parameter for Invertible Short and Long Memory ARIMA Models”, Journal of the Royal Statistical Society Series B, 57, 659-672.

Black F. (1976), “The Pricing of Commodity Contracts”, Journal of Financial Economics, 3, 167-179.

Black F. and M. Scholes (1973), “The Pricing of Options and Corporate Liabilities”, Journal of Political Economy, 81, 637-659.

Canina L. and S. Figlewski (1993), “The Informational Content of Implied Volatility”, Review of Financial Studies, 6, 659-681.

Christensen B. J. and N. R. Prabhala (1998), “The Relation Between Implied and Realized Volatility”, Journal of Financial Economics, 50, 125-150.

Day T. and C. Lewis (1993), “Forecasting Futures Markets Volatility”, Journal of Derivatives, 1, 51-63.

Fair R. C. and R. J. Shiller (1990), “Comparing Information in Forecasts from Econometric Models”, American Economic Review, 80, 375-389.

Feinstein S. (1989), “The Black-Scholes Formula is Nearly Linear in σ for At-The-Money Options”, Presentation at the AFA Conference, Washington, D.C..

Figlewski S. (1997), “Forecasting Volatility”, Financial Markets Institutions & Instruments, 6, 1-88.

Fleming J. (1993), “The Quality of Market Volatility Forecasts Implied by S&P 100 Index Option Prices”, Working Paper, Duke University, N.C..

Fleming J, B. Ostdiek and R. E. Whaley (1995), “Predicting Stock Market Volatility: a New Measure”, Journal of Futures Markets, 15, 265-302.

Gemmill G. (1986), “The Forecasting Performance of Stock Options on the London Traded Option Market”, Journal of Business Finance and Accounting, 13, 535-546.

Granger C. W. J. (1980), “Long Memory Relationships and the Aggregation of Dynamic Models”, Journal of Econometrics, 14, 227-238.

Granger C. W. J. and P. Newbold (1974), “Spurious Regressions in Econometrics”, Journal of Econometrics, 2, 111-120.

Granger C. W. J. and R. Joyeux (1980), “An Introduction to Long Memory Time Series Models and Fractional Differencing”, Journal of Time Series Analysis, 1, 15-39.

Greene W. H. (1993), Econometric Analysis, Prentice-Hall, New Jersey.

Guo D. (1996), “The Predictive Power of Implied Stochastic Variance from Currency Options”, Journal of Futures Markets, 8, 915-942.

Hamilton J. D. (1994), “Time Series Analysis”, Princeton University Press, New Jersey.

Hansen L. P. (1982), “Large Sample Properties of Generalized Methods of Moments Estimators”, Econometrica, 50, 1029-1054.

Hansen L. P. and R. J. Hodrick (1980), “Forward Exchange Rates as Optimal Predictors of Future Spot Rates: an Econometric Analysis”, Journal of Political Economy, 88, 829-853.

Hansen L.P. and R. J. Hodrick (1983), “Risk Averse Speculation in Forward Exchange Markets: an Econometric Analysis”, in “Exchange Rates and International Macroeconomics”, J. A. Frankel (ed.), University of Chicago Press, Chicago.

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

Hwang S. and S. E. Satchell (1998), “Implied Volatility Forecasting: A Comparison of Different Procedures Including Fractionally Integrated Models with Applications to UK Equity Options”, in “Forecasting Volatility in the Financial Markets”, J. Knight and S. Satchell (eds.), Butterworth - Heinemann, Oxford.

Im K. S, M. H. Pesaran and Y. Shin (2003), “Testing for Unit Roots in Heterogeneous Panels”, Journal of Econometrics, 115, 53-74.

Jorion P. (1995), “Predicting Volatility in the Foreign Exchange Market”, Journal of Finance, 50, 507-528.

Lamoureux C. G. and W. D. Lastrapes (1993), “Forecasting Stock - Return Variance: Toward an Understanding of Stochastic Implied Volatilities”, Review of Financial Studies 6, 293-326.

Li K. (2002), “Long-Memory versus Option-Implied Volatility Predictions”, Journal of Derivatives, 9, 9-25.

Lieu D. (1990), “Option Pricing with Futures-Style Margining”, Journal of Futures Markets, 10, 327-338.

Lo A. W. (1991), “Long Term Memory in Stock Market Prices”, Econometrica, 59, 1279-1313.

Neuhaus H. (1995), “The Information Content of Derivatives for Monetary Policy”, Discussion Paper 3/95, Economic Research Group of the Deutsche Bundesbank.

Pesaran M. H. (2003), “A Simple Panel Unit Root Test in the Presence of Cross Section Dependence”, mimeo, University of Southern

California & Cambridge University. Available at http://www.econ.cam.ac.uk./faculty/pesaran/panelcadf.pdf.

Phillips P. C. B. and B. Hansen (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes", Review of Economic Studies, 57, 473-495.

Richardson M. and T. Smith (1991), "Tests of Financial Models in the Presence of Overlapping Observations", Review of Financial Studies, 4, 227-254.

Scott L. O. (1992), "The Information Content of Prices in the Derivative Security Markets", International Monetary Fund Staff Papers, 39, 596-625.

Stein J. C. (1989), "Overreactions in the Option Market", Journal of Finance, 44, 1011-1023.

Acknowledgements

The author would like to thank an anonymous referee for useful suggestions.

Notes

[1] LIFFE options are of the American style but, because of low transaction costs, margining eliminates significant differences between European and American style options profit opportunities and Black's formula above can be applied. For more details, see Lieu (1990).

[2] It should be noticed that, owing to the linear relationship between the two variables, it is irrelevant whether F_t or i_t is used as underlying asset; a call (put) option on an interest futures price is equivalent to a put (call) option on the implied interest rate.

[3] Black and Scholes (1973) type models should be inconsistent with stochastic volatilities. However, as shown by Hull and White (1987), if volatility is uncorrelated with aggregate consumption, an option (call) price is equal to the expected Black-Scholes price integrated over the average variance during the life of the option. More generally, implied volatility obtained inverting a Black-Scholes type formula will be, in a stochastic context, an unbiased estimator of average expected volatility of the underlying over the remaining life of the option if there is no risk premium and if the price of the option is linear in volatility. For more details see Stein (1989) and Feinstein (1989), among others.

[4] Alexander and Leigh (1997) point out that extreme market movements tend to distort volatility measures based on equally weighted averages. A N-day past squared returns moving average will be affected in the same way,

irrespectively of whether the abnormal return shift occurred 1 or *N-1* periods ago. Volatility estimates may thus be kept artificially high even if the true volatility has dropped to a lower level.

[5] In the presence of serial correlation (short memory), but not of long memory, of the x_t time series, Q_T converges weakly to the range of a Brownian bridge on the unit interval. The corresponding quantiles can be found in Lo (1991).

[6] A fractionally integrated ARIMA(*p,d,q*) or ARFIMA(*p,d,q*) process reads as $\Phi(L)(1-L)^d x_t = \Psi(L)\varepsilon_t$. All roots of $\Phi(L)$ and $\Psi(L)$ lie outside the unit circle and ε_t is iid (0, σ^2). The fractional difference operator is defined as $(1-L)^d = \sum_{k=0,\infty}\{\Gamma(k-d)L^k/[\Gamma(k+1)\ \Gamma(-d)]\}$where $\Gamma(.)$ is the gamma function. For $-1/2<d<1/2$ the process is covariance stationary, while $d<1$ implies mean reversion. This is in contrast to a unit root process which is both covariance non-stationary and not mean-reverting. Interesting ARFIMA analyses of implied volatility performance can be found in Hwang and Satchell (1998) and in Li (2002).

[7] A word of caution is called for here. Some of the time series do lie on the borderline separating stationarity from non stationarity, the BIC statistics of their non stationary parameterisations being only marginally larger than those of the stationary short memory ones.

[8] Multicollinearity between ISD_t and σ_{Pt} might affect coefficient estimation in encompassing tests of this kind.

[9] The regressor $\sigma^F_{t,T}$ in equation (7) is predetermined. It is not necessarily exogenous, which may lead to a finite sample bias in the coefficient estimates. Indeed $\sigma^F_{t,T}$ is correlated with past error terms even if it is not correlated with contemporaneous or future error terms. These correlations exist since shocks to the regressor are correlated with shocks to the regressand and the regressor is highly persistent.

[10] The GLS estimator is consistent if $\mathrm{cov}(\alpha_i, \sigma^F_{it,T}) = 0$. Under the null of orthogonality the within units OLS and GLS estimators are both consistent (but the former is inefficient) whereas under the alternative the within units OLS estimator is consistent but the GLS is not. Under the null the two estimates should not differ too much and a Wald test is derived in order to assess the statistical relevance of their difference (and thus of the null). Let $\Omega = \mathrm{var}(b^w) - \mathrm{var}(b^G)$, the Wu-Hausman Wald test statistic is defined as $(b^w - b^G)\Omega^{-1}(b^w - b^G)$. It has a chi squared distribution with as degrees of freedom the number of regressors.

[11] Relative out of sample forecasting accuracy has been assessed also using the Root Mean Square Error minimisation criterion. It has to be interpreted with caution; Alexander (2001, pages 122-123) points out that RMSE statistics will give poor results when applied to second moments because of excessive noise. The Mean Absolute Error is also set forth since, as suggested by Gemmill (1986), it implies a linear loss function and should provide a better measure of

forecasting performance. We obtain the following statistics over the 1993-1997 time period

	S. Sterling			Euromark		
	$\sigma_{H,t}$	$\sigma_{G,t}$	ISD_t	$\sigma_{H,t}$	$\sigma_{G,t}$	ISD_t
MAE	3.3093	3.4800	3.2785	5.8995	6.0159	4.1230
RMSE	4.3260	4.3468	4.1784	8.4246	8.4401	4.9456

They confirm the ranking obtained with the regression analysis. Forecast errors are lower in the case of the Short Sterling contract. Implied volatilities are more accurate than historical forecasts and, here too, GARCH out of sample forecasts provide the worst results.

[12] $t_i(N,T)$ being the t-ratio of b_i in the CADF regressions, $t_i^*(N,T)$ is selected as follows $\begin{cases} t_i^*(N,T)=t_i(N,T) & if \quad -k_1 < t_i(N,T) < k_2 \\ t_i^*(N,T)=-k_1 & if \quad t_i(N,T) \le -k_1 \\ t_i^*(N,T)=k_2 & if \quad t_i(N,t) \ge k_2 \end{cases}$ The boundaries k_2 and k_1 are determined in such a way that $\Pr[-k_1 < t_i(N,T) < k_2] > 0.999$, using the normal approximation of $t_i(N,T)$ as a benchmark. Their use is justified in Pesaran (2003, theorem 2). In the tests above, $k_1 = 6.7195$ and $k_2 = 3.2595$.

[13] White's cross sectional heteroskedasticity test reads as MR^2 where M is the total number of observations in the panel and R^2 is provided by the regression of the squared residuals of an OLS first stage pooled data estimation of the model on a constant and all unique variables in $x \otimes x$ where x are the model regressors. It is asymptotically distributed as a chi-square with *g-1* degrees of freedom, where *g* is the number of regressors (not including the constant). Significance implies rejection of the null of cross sectional homogeneity.

[14] Panel data estimates may, however, be affected by some cross sectional correlation of the residuals. Residuals are cross sectionally uncorrelated here if $\text{cov}(u_{it,T}, u_{js,T}) = 0$ (i) $\forall t \neq s$ and (ii) $\forall i \neq j$. Condition (i) ensures that the *(20x20)* residual submatrices that constitute the *(1260x1260)* covariance matrix of the residuals from the pooled data regression are diagonal. Condition (ii) transforms the unrestricted pooled data regression matrix in a diagonal matrix of 63 *(20x20)* unit regression residual blocks. The combination of both conditions ensures that also the latter are diagonal. Condition (ii) may be violated, at least in the case of adjoining trading days, because of the overlapping nature of the data whereas condition (i) is likely to be satisfied because of the quarterly sampling procedure. The covariance matrix of the residuals from the pooled data regression will not be wholly diagonal as some of the *(20x20)* off diagonal submatrices that constitute it will have entries that are different from zero. It should be noticed, however, that since these submatrices are diagonal, because of condition (i), most of the off-diagonal entries of the pooled regression residual matrix are likely to be nil.