

A COMPARISON OF SEASONAL ADJUSTMENT METHODS WHEN FORECASTING INTRADAY VOLATILITY*

Martin Martens**

University of New South Wales

Yuan-Chen Chang

National Chung-hsing University

Stephen J. Taylor

Lancaster University

JEL classification: C53; G15

* The authors would like to thank Torben Andersen, Tim Bollerslev, Eric Ghysels, Bruce Lehmann, editor William T. Moore, Diethelm Wurtz, an anonymous referee, and participants of the HFDF-II conference in Zurich, Switzerland, for useful comments. We are also grateful to Olsen & Associates for providing the exchange rate data. All remaining errors are our own responsibility.

** Corresponding author: tel. +61 2 9385 4291; fax +61 2 9385 6347; email: m.martens@unsw.edu.au

Abstract

In this study we compare volatility forecasts over a thirty-minute horizon for the spot exchange rates of the Deutsche Mark and the Japanese Yen against the US dollar. Explicitly modeling the intraday seasonal pattern improves the out-of-sample forecasting performance. We find that a seasonal estimated from the log of squared returns improves upon the use of simple squared returns, and that the flexible Fourier form (FFF) is an efficient way of determining the seasonal. The two-step approach that first estimates the seasonal using the FFF and then the parameters of the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model for the deseasonalized returns performs only marginally worse than the computationally expensive periodic GARCH model that includes the FFF.

I. Introduction

The intraday seasonal patterns in the volatility of foreign exchange and stock markets have important implications for modeling the volatility of high-frequency returns. The patterns are so distinctive there is a strong case for taking them into account before attempting to model the dynamics of volatility.

Standard time series models of volatility have proven inadequate when applied to high-frequency returns data. Andersen and Bollerslev (1997, 1998) argue that the reason for this is simply the systematic pattern in average volatility across the trading day. They also present a practical method for the estimation of the intraday seasonal pattern. Until now, however, there are no studies known to us that investigate the effect of the seasonal pattern on forecasting intraday volatility. Although Andersen and Bollerslev (1997) argue that persistence estimates for Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models estimated at different frequencies are in conflict with aggregation results for Autoregressive Conditional Heteroscedasticity (ARCH) effects (Drost and Nijman, 1993), these models could still be used to forecast the volatility in the next period.

Our objective is to test whether filtering out the seasonal pattern, that is deseasonalizing, improves the out-of-sample forecasting performance of volatility models. We also want to determine the best method for removing seasonality. We compare the forecasting performance of a GARCH model for the original intraday returns with the forecasting performance of a GARCH model for deseasonalized intraday returns. The seasonal is estimated by either simply averaging the volatility over a number of trading days for each intraday period, or using the flexible Fourier

form (FFF) advocated by Andersen and Bollerslev (1997, 1998). These approaches are two-step procedures. The seasonal is first estimated and extracted, and a GARCH model is then estimated. The natural question is then whether the one-step periodic GARCH (P-GARCH) model of Bollerslev and Ghysels (1996) can improve this.

We evaluate volatility forecasts over a thirty-minute horizon for the exchange rates of the Deutsche Mark against the US Dollar (DEM/USD) and Japanese Yen against the US Dollar (YEN/USD). Methods are developed from the returns for the period January to September 1996 and are assessed out-of-sample from October to December 1996. Our conclusions are the same for both currencies. The results show it is worthwhile to consider the seasonal explicitly when forecasting volatility. All three seasonal models outperform the GARCH model for the original returns. It is also better to estimate the FFF in contrast to the simple method of averaging the realized volatility over appropriate intervals. The smoothing effect of the FFF can explain this result. The two-step method using the FFF provides results that are only marginally worse than those of the P-GARCH model. As the two-step method using the FFF is computationally much more efficient than P-GARCH there is a strong case for giving preference to the FFF method.

There is a consensus in the literature that the total volatility is described by the product of two components, the seasonal and the remaining volatility dynamics. We find some support for the strategy of Andersen and Bollerslev (1997, 1998) that estimating the seasonal from the logarithms of squared returns improves over estimating the seasonal directly from squared returns.

II. Data

Olsen & Associates provided the database HFDF-II, which includes the most recent and subsequent Reuters FAFX quotes around half-hour marks for all of 1996. We study the two major exchange rates, the DEM/USD and the YEN/USD.

Since our primary objective is to forecast volatility, we use the last quote before the half-hour mark to reflect the exchange rate at the half-hour mark. The exchange rates on the half-hour marks are calculated as the midpoint of the bid- and ask-rates. From these half-hour prices we construct the half-hour return series:

$$r_t = \ln((p_t^{bid} + p_t^{ask}) / 2) - \ln((p_{t-1}^{bid} + p_{t-1}^{ask}) / 2). \quad (1)$$

Intervals that have no quote are assigned zero returns. There are only sixteen such zeros during the 260 days used to obtain empirical results.

Because of microstructure effects¹ the return series exhibit negative first-order autocorrelation. For this reason we first filter the returns before starting our analysis, estimating

$$r_t = \mathbf{r} r_{t-1} + e_t. \quad (2)$$

We find $\hat{\mathbf{r}}$ is -0.088 for the DEM/USD and -0.159 for the YEN/USD. The residual return, R_t , is then defined as

$$R_t = \hat{e}_t = r_t - \hat{\mathbf{r}} r_{t-1}. \quad (3)$$

¹ Because banks are often only willing to sell (buy) but not buy (sell), they will skew the bid-ask quotes to one side. This will result in negative autocorrelation when using midpoints of the bid- and ask-quotes (e.g., Goodhart and Figliuoli (1991), Bollerslev and Domowitz (1993)).

In the following text, residual returns are often simply called returns.

Next, we follow the adjustment process of Andersen and Bollerslev (1998) by excluding the returns from Friday 21:00 Greenwich Mean Time (GMT) through Sunday 21:00 GMT. A description of this weekend definition can be found in Bollerslev and Domowitz (1993). The last three hours of Sunday returns are regarded as Friday returns. As there is a discontinuity in the returns between consecutive weeks, we do not use forecasts for the last three hours on Friday, and we also remove the first two hours on Mondays to diminish the effect of this discontinuity. We also remove January 1 and December 25 from the sample as these holidays can have very low volatility (Taylor and Xu, 1997). As a result, for 1996 we have 260 days each with forty-eight (residual) half-hour returns.

Reuters' data are recorded in GMT, which does not reflect the daylight saving schemes followed by most of the major financial markets. Analysis of the volatility patterns without a suitable adjustment to GMT could be misleading during daylight saving periods. Consequently, we adjust GMT during the American daylight saving period, which is in effect from the first Sunday in April until the last Sunday in October. For 1996 this is from April 7 to October 27. The adjustment procedure is to add one hour to GMT during the American daylight saving period and to leave other times unchanged. We call this adjusted clock GMT_{adj} . The remainder of the analysis will be based on this adjusted time. US Eastern Time can be obtained by subtracting five hours from GMT_{adj} .

III. Alternative Models and Performance Measures

No Seasonal Adjustment

A natural question is whether out-of-sample forecasting of volatility is actually improved when using seasonal methods. To provide an answer, we first forecast volatility using the residual returns as defined in equation (3) using a GARCH(1,1) model for the thirty-minute returns. We use this model as our benchmark, as follows,

$$\begin{aligned} R_t &= \mathbf{m} + \mathbf{e}_t \\ \mathbf{e}_t | \Psi_{t-1} &\sim D(0, h_t) \\ h_t &= \mathbf{w} + \mathbf{a}\mathbf{e}_{t-1}^2 + \mathbf{b}h_{t-1} \end{aligned} \tag{4}$$

where Ψ_{t-1} is the information set at time $t-1$, and for D we use the scaled t-distribution.² The out-of-sample variance forecast, calculated using estimated parameters, is then simply

$$\hat{h}_{t+1} = \hat{\mathbf{w}} + \hat{\mathbf{a}}\hat{\mathbf{e}}_t^2 + \hat{\mathbf{b}}\hat{h}_t \tag{5}$$

Since we will be forecasting volatility related to realized volatility represented by $|R_t|$, it is necessary to adjust the level of the forecast in order to avoid bias. For GARCH models having conditional scaled t-distributions with v degrees of freedom, the density of $D(0,1)$ is

$$f(z) = \mathbf{p}^{-\frac{1}{2}} (v-2)^{-\frac{1}{2}} \frac{\Gamma[(v+1)/2]}{\Gamma[v/2]} \left(1 + \frac{z^2}{v-2} \right)^{-(v+1)/2} \quad (v > 2), \tag{6}$$

² We originally used the normal distribution. The t-distribution significantly improves the in-sample fit, the degrees of freedom being approximately four. However, out-of-sample variance forecasts hardly change, and neither does the forecasting performance. We therefore did not search for improvements upon the t-distribution.

where Γ is the gamma-function, and the expected absolute return is

$$E | R_{T+1} | = 2 \frac{\sqrt{v-2}}{\sqrt{p}} \frac{\Gamma[(v+1)/2]}{\Gamma[v/2](v-1)} \sqrt{\hat{h}_{T+1}} . \quad (7)$$

Thus the volatility forecast from the GARCH model is scaled to remove bias from the expected absolute return, which is our measure for realized volatility.

Seasonal Adjustment Using Averages of Squared Returns

Taylor and Xu (1997) use for each intraday period the appropriate average of the squared returns over all trading days. Let $R_{d,n}$ be the n th intraday return on day d , and suppose we have D days and N intraday periods. Then the seasonal variance estimate is given by

$$s_n^2 = \frac{1}{D} \sum_{d=1}^D R_{d,n}^2 \quad (n = 1..N) . \quad (8)$$

Note that, throughout the paper, returns are either denoted by R_t or $R_{d,n}$ with

$$t = 48(d-1) + n.$$

A second way to estimate the seasonal pattern takes account of the day of the week. Let S_d be the set of daily time indices that share the same day-of-the-week as time index d . Let N_d be the number of time indices to be found in S_d . Then the seasonal variance is given by

$$s_{d,n}^2(1) = \frac{1}{N_d} \sum_{s \in S_d} R_{s,n}^2 . \quad (9)$$

Andersen and Bollerslev (1997, 1998) use the logarithm of the squared returns to help estimate seasonal patterns, following the popular assumption that volatility is the product of the seasonal volatility and the time-varying non-seasonal component. Hence, alternatively,

$$s_{d,n}^2(2) = \exp \left[\frac{1}{N_d} \sum_{s \in S_d} \ln((R_{s,n} - \bar{R})^2) \right] \quad (10)$$

where \bar{R} is the overall mean taken over all returns.

Using the seasonal terms we filter the returns:

$$\tilde{R}_t \equiv \tilde{R}_{d,n} \equiv \frac{R_{d,n}}{s_{d,n}} \quad (11)$$

and then estimate a GARCH(1,1) model for the filtered returns:

$$\begin{aligned} \tilde{R}_t &= \mathbf{m} + \mathbf{e}_t \\ \mathbf{e}_t | \tilde{\Psi}_{t-1} &\sim D(0, \tilde{h}_t) \\ \tilde{h}_t &= \mathbf{w} + \mathbf{a}\mathbf{e}_{t-1}^2 + \mathbf{b}\tilde{h}_{t-1} \end{aligned} \quad (12)$$

with $\tilde{\Psi}_{t-1}$ the returns history Ψ_{t-1} augmented by the seasonal terms. To transform the volatility forecasts for the filtered returns into volatility forecasts for the original returns, we multiply the volatility forecasts by the appropriate seasonal term, $s_{d,n}$.

FFF

Andersen and Bollerslev (1997, 1998) advocate the FFF to model the intraday volatility pattern. They assume the following representation for the intraday returns:

$$R_{d,n} = E(R_{d,n}) + \frac{\mathbf{s}_d s_{d,n} Z_{d,n}}{N^{1/2}} \quad (13)$$

with the terms $Z_{d,n}$ being i.i.d. with zero mean and variance one. They then define

$$x_{d,n} \equiv 2\ln[|R_{d,n} - E(R_{d,n})|] - \ln \mathbf{s}_d^2 + \ln N = \ln s_{d,n}^2 + \ln Z_{d,n}^2. \quad (14)$$

Replacing $E(R_{d,n})$ by the average of all intraday returns, and \mathbf{s}_d by an estimate from a daily volatility model, $\hat{x}_{d,n}$ is obtained. The seasonal pattern is estimated by OLS using

$$\hat{x}_{d,n} = \sum_{j=0}^J \mathbf{s}_d^j \left[\mathbf{m}_{0j} + \mathbf{m}_{1j} \frac{n}{N_1} + \mathbf{m}_{2j} \frac{n^2}{N_2} + \sum_{i=1}^I \mathbf{l}_{ij} I_{t=d_i} + \sum_{i=1}^p \left(\mathbf{g}_{ij} \cos \frac{2pin}{N} + \mathbf{d}_{ij} \sin \frac{2pin}{N} \right) \right] \quad (15)$$

where $N_1 = (N+1)/2$ and $N_2 = (N+1)(N+2)/6$ are normalizing constants, and p is set equal to four.³ Each of the corresponding $J+1$ FFFs are parameterized by a quadratic component (terms with μ -coefficients) and a number of sinusoids. Moreover, it may be advantageous to also include time specific dummies for applications in which some intraday intervals do not fit well within the overall regular periodic pattern (the λ coefficients).

We follow a similar specification. Since we want to forecast volatility, we can not use σ_d based on future information. We therefore assume there is no interaction between the daily volatility level and the seasonal volatility pattern and hence $J=0$. We do, however, estimate a separate FFF for each day of the week. Ignoring σ_d also

³ Several different specifications are possible all based on the same idea; even Andersen and Bollerslev use different specifications in their 1997 and 1998 papers. All specifications lead to a similar smooth fit of the average of squared returns for each interval, see the results in section IV and Figures Ia to Id.

means that the term $\ln(\mathbf{s}_d^2)$ is removed from the definition of $\hat{x}_{d,n}$. Andersen and Bollerslev (1998) show that this has little impact on the seasonal.

For the dummy variables we use the three most important US macro-economic news announcements (Payne, 1996), i.e. unemployment, retail sales and international trade. These dummies reflect higher volatility for thirty minutes following news announcements. The news announcements are collected from Business Weekly. To provide a fair comparison, the averages in equations (8) and (9) are calculated separately for time intervals with and without news.

The intraday seasonal volatility pattern is determined using

$$s_{d,n}(3) = \exp(\hat{x}_{d,n} / 2) \quad (16)$$

Instead of using the logarithm of the squared returns, we also estimate the FFF directly from the squared returns. Either estimated seasonal pattern can then be used to filter the returns, and estimate the GARCH(1,1) model in equation (12).

P-GARCH

Bollerslev and Ghysels (1996) introduce the P-GARCH model explicitly designed to capture the repetitive seasonal time variation in the second-order moments. P-GARCH includes all GARCH models in which hourly dummies, for example, are used in the variance equation.

Extending the information set Ψ_{t-1} with a process defining the stage of the periodic cycle at each point in time, say to Ψ_{t-1}^s , the class of P-GARCH(1,1) models is defined as

$$\begin{aligned}
R_t &= \mathbf{m} + \mathbf{e}_t \\
\mathbf{e}_t | \Psi_{t-1}^s &\sim D(0, h_t) \\
h_t &= \mathbf{w}_{s(t)} + \mathbf{a}_{s(t)} \mathbf{e}_{t-1}^2 + \mathbf{b}_{s(t)} h_{t-1}
\end{aligned} \tag{17}$$

where $s(t)$ refers to the stage of the periodic cycle at time t . The periodic cycle of interest here is a repetitive cycle covering one week.

The P-GARCH model is potentially more efficient than the methods described above. These methods first estimate the seasonal and, after de-seasonalizing the returns, the volatility of these adjusted returns is estimated. Whereas this is a two-step method, the P-GARCH model allows a simultaneous estimation of the seasonal effects and the remaining time-varying volatility. On the other hand, however, we need a parsimonious model. Otherwise, estimating the P-GARCH model will become too difficult and time-consuming.

Two extreme cases are to use parameters $\omega_{s(t)}$ in (17) in such a way that they represent the role of (i) the average absolute/squared returns (e.g. 240 dummies); or (ii) the FFF. We estimate case (ii) where we only have one FFF for the entire week, instead of a separate FFF for each day of the week⁴:

$$s_n^2 = \mathbf{w}_0 + \mathbf{w}_1 \frac{n}{N_1} + \mathbf{w}_2 \frac{n^2}{N_2} + \sum_{i=1}^4 (\mathbf{g}_i \cos \frac{2\mathbf{p}in}{N} + \mathbf{d}_i \sin \frac{2\mathbf{p}in}{N}) \tag{18}$$

and

$$\begin{aligned}
R_t &= \mathbf{m} + \mathbf{e}_t \\
\mathbf{e}_t | \Psi_{t-1}^s &\sim D(0, h_t) \\
h_t &= s_n^2 + \mathbf{a} \mathbf{e}_{t-1}^2 + \mathbf{b} h_{t-1}.
\end{aligned} \tag{19}$$

An alternative specification for the variance equation is

$$h_t = \omega s_n^2 + \frac{s_n^2}{s_{n-1}^2} (ae_{t-1}^2 + bh_{t-1}) \quad (20)$$

with ω_0 set to one. The motivation for this specification is the GARCH(1,1) model for deseasonalized returns in equation (12), with $R_t = s_n \tilde{R}_t$ and $h_t = s_n^2 \tilde{h}_t$.

Forecasting Performance Measures

We compare the volatility forecasts from the four models described above with the realized volatility. Realized volatility is simply measured as the absolute realized return in the forecasted period.⁵

In previous literature various methods are provided to assess the performance of out-of-sample forecasts. See, for example, Day and Lewis (1992) and Canina and Figlewski (1993) who use the squared correlation between realized and forecasted volatility, as well as encompassing regressions, and Taylor and Xu (1997) who use measures of relative forecast errors. Since there is no consensus in the forecasting literature about the best method, we report several of the most popular performance measures.

The first method is to calculate the Root Mean Squared Error (RMSE). If \hat{s}_t is the forecasted volatility for period t , the RMSE for T forecasts is

⁴ As is shown in section IV, the FFF performs better than the simple R^2 -method. The latter would also require more parameters in the P-GARCH model as opposed to the FFF. We, therefore, focus on including the FFF in the P-GARCH model.

⁵ Alternatively the realized volatility can be measured by the squared return with forecasts given by conditional variances instead of by conditional standard deviations. The absolute return is preferred here because it reduces the impact of outliers. When all results are reproduced using squared returns and variances, the results and conclusions are very similar to those reported here.

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (|R_t| - \hat{\mathbf{s}}_t)^2}. \quad (21)$$

The second method is to calculate the logarithmic loss function employed in Pagan and Schwert (1990),

$$\text{LL} = \frac{1}{T} \sum_{t=1}^T (\ln |R_t - \bar{R}| - \ln(\hat{\mathbf{s}}_t))^2. \quad (22)$$

This alternative emphasizes the importance of accurate predictions when the realized volatility is close to zero, and it reduces the impact of outliers.

The third method we employ is to simply calculate the correlation between realized and forecasted volatility. The fourth method uses the regression

$$|R_t| = \mathbf{a} + \mathbf{b}\hat{\mathbf{s}}_t + \mathbf{e}_t \quad (23)$$

and reports the goodness of fit measured by the squared correlation, R^2 .

Finally, we also use the encompassing regression in which we include more than one forecast in the regression above and then check whether one forecast dominates the forecast from another method.

IV. Results

Intraday Seasonal Pattern

For two of the four methods we first filter (de-seasonalize) the returns. The seasonal pattern is estimated using the mean of the squared returns (equations (9), (10)) or the FFF (equations (14), (15)). The seasonal is estimated for both $\ln((R_t - \bar{R})^2)$ and simply R_t^2 . The seasonals are estimated for the first nine months of

1996. The averages taken over all thirty-nine weeks provide the resulting seasonals in Figures Ia and Ib for the DEM/USD and in Figures Ic and Id for the YEN/USD.

Insert Figure I about here

An obvious difference between using squared returns and their logarithms is that ‘outliers’ have less impact in the latter method. It is also clear that the FFF tries to find a smooth fit of the R^2 -seasonal. If the R^2 -seasonal is driven by outliers explaining the spikes on many occasions, we expect it to do well in-sample (since the effect of the outliers will be reduced), but worse out-of-sample (large returns are not repeated in the same thirty-minute intervals).

The seasonals for the YEN/USD seem to be less pronounced than the seasonals for the DEM/USD, partly reflected by many more spikes for the YEN/USD, and the average level of the seasonal is higher for the YEN/USD. This is probably caused by our sample period showing a relatively volatile year for the YEN/USD. In February 1996 the Bank of Japan intervened several times to support the dollar against the yen, and for much of the remainder of 1996 there was strong speculation of pending Japanese interest rate rises.

In-Sample Fit

Although we can not compare directly the in-sample fit for all the models, we can obtain some idea of their adequacy by estimating them for the first nine months. The seasonals presented in Figure I are employed for this purpose. The results are presented in Table 1.

Insert Table 1 about here

The parameter estimates for α and β are significantly different from zero at the 5 percent level for all six models. The inclusion of the FFF in the GARCH specification (the P-GARCH model) leads to a highly significant improvement over the use of a standard GARCH(1,1) model.

The R^2 - and FFF-methods estimate GARCH(1,1) parameters from the deseasonalized returns, $\tilde{R}_{d,n} = R_{d,n} / s_{d,n}$. This explains the substantial variation among the reported log-likelihoods. We adjust the log-likelihood by recalculating the log-likelihood after multiplying the mean-adjusted returns by the seasonal $s_{d,n}$ and the estimated conditional variances by the squared seasonal. This results in the R^2 -method (with seasonal based on the log of squared returns) having the highest log-likelihood. Of course we have to be cautious about drawing any conclusions from these likelihoods, since the models are generally non-nested. The R^2 -method is likely to be favored in this comparison, since the spikes in the seasonal reflect some large outliers which will now be fit well. This ‘advantage’ disappears when forecasting the volatility out-of-sample, as we will see below.

Interestingly, the use of the log of the squared returns (labeled ‘ $\ln R^2$ ’) instead of squared returns (labeled ‘ R^2 ’) to calculate the seasonals seems to improve the in-sample fit.

The P-GARCH model is obviously at a disadvantage compared to the FFF-method that makes use of a different FFF for each day-of-the-week, as well as news dummies. To check whether P-GARCH is more efficient than the two-step method based on the FFF, we extend equation (19) for the DEM/USD with a FFF for each day-of-the-week and the news dummies. The log-likelihood increases to 14,028, which is better than the two-step method regardless of whether the FFF is based on the

squared returns or the log of squared returns. Using an exponential version of the FFF in equation (19) with s_n^2 replaced by $\exp(s_n^2)$, gives a log-likelihood of 14,035. Repeating this exercise for equation (20) gives log-likelihoods of 14,039 and 14,045, respectively. Hence P-GARCH is more efficient in-sample, while here also the exponential form is an improvement although the effect is much smaller. Repeating the calculations for the YEN/USD give a highest log-likelihood of 11,455 for equation (12) and using $\exp(s_n^2)$, still short of the 11,462 for the two-step method based on the FFF for the log of squared returns. However, as we show below, out-of-sample it is better to use a more parsimonious model like the one provided in equation (19).

Out-of-Sample Forecasting Performance

Using the estimated seasonals in Figure I and assuming they can be used for the last three months of 1996, we now proceed with forecasting for each method. The GARCH(1,1) model is repeatedly estimated using data from the 60 days (2880 observations) preceding the forecasting period for the DEM/USD, and 195 days for the YEN/USD.⁶ The parameters are re-estimated every twenty-four hours. This results in 3120 forecasts. In the comparisons below we remove the last three hours on Fridays and the first two hours on Mondays, as well as intervals containing pre-scheduled news and the two subsequent intervals following the news announcements, leaving us with 2963 forecasts. The latter deletions are made to make the competition fair for the ‘raw’ and P-GARCH methods which do not use news dummies. It is also beyond the scope of this study to forecast volatility around news announcements.

⁶ A number of robustness checks revealed that for the DEM/USD the GARCH parameters changed more rapidly over the out-of-sample period than for the YEN/USD. For the YEN/USD all models performed better using 195 days preceding the forecasting period instead of sixty days.

The forecasts are compared using the measures discussed in section III. The results for the univariate measures are provided in Table 2.

Insert Table 2 about here

First of all it is important to note that modeling the seasonal pattern in volatility improves out-of-sample forecasting. It is obvious that for nearly all univariate measures the model using the (original) raw returns is the worst of the six models under consideration, the other five taking into account the intraday seasonal. The only exceptions are for the R^2 -method for the YEN/USD. Nevertheless, the forecasts using GARCH(1,1) for the raw returns still have a correlation of 0.245 with the realized volatility measure (the absolute returns) for the DEM/USD, and 0.237 percent for the YEN/USD.

Second, it is interesting to note that the seasonal based on $\ln((R_t - \bar{R})^2)$ improves out-of-sample forecasting compared with the seasonal based on R_t^2 . For both the R^2 - and the FFF-methods we find a higher correlation with realized volatility, as well as a lower Root Mean Squared Error and a higher coefficient of determination. For example, for the FFF the correlation is 0.294 versus 0.268 in the case of the DEM/USD, while for the YEN/USD it is 0.269 versus 0.264.

Third, it seems worthwhile to estimate the FFF rather than using the simpler R^2 -method. If we compare, for example, the results using the seasonal based on $\ln((R_t - \bar{R})^2)$, we find for the DEM/USD that the correlation for the FFF-method is 0.294, compared with 0.273 for the R^2 -method. For the YEN/USD the correlation for the FFF-method is 0.269 versus 0.251 for the R^2 -method. Also the RMSE is lowest for the FFF-method. A logical explanation is that the smooth fit of the FFF more

realistically describes the true underlying seasonal pattern. The R^2 -method will be more affected by outliers (resulting in the spiky seasonals in Figures Ia to Id), which are not likely to happen again in the same intraday intervals in other weeks.

Finally, it is worth noting the performance of the P-GARCH model using the specification in equation (19). P-GARCH ranks first for both the DEM/USD and the YEN/USD, using simply one FFF for the entire week (in contrast to different FFF's for every day of the week) and only an adjustment to the constant in the GARCH specification. P-GARCH outperforms the raw method thanks to including the seasonal. It, however, also outperforms the R^2 -method and the FFF-method. It is notable that P-GARCH ranks first and that the FFF method, with seasonal based on $\ln((R_t - \bar{R})^2)$, ranks second, for both currencies and for each of the performance measures given by the correlation, coefficient of determination, RMSE and LL.

Insert Table 3 about here

The results for the encompassing regressions are provided in Table 3. They support the conclusions based on correlation, coefficient of determination, RMSE and LL. First of all it is clear that the model ignoring the seasonal performs the worst. Including forecasts from P-GARCH makes the raw forecasts insignificant (see row 5 for both Panels A and B in Table 3), which is also the case using forecasts for the DEM/USD from the FFF-method based on the 'log'-seasonal (see row 4 in Panel A) and for the YEN/USD from the FFF-method based on the squared returns (see row 3 in Panel B). In the other cases, however, the raw forecasts do have some incremental information over the forecasts from the other methods. Second, the forecasts based on $\ln((R_t - \bar{R})^2)$ are better than the forecasts based on R_t^2 . For the FFF the latter are

even insignificant when both are included in one regression for the DEM/USD case (row 13 in Panel A of Table 3). Third, the FFF improves over the R^2 - method, although the latter does provide incremental information in the DEM/USD case (but not in the YEN/USD case).

V. Conclusions

In this study we investigate whether modeling the intraday seasonal volatility pattern in financial markets improves out-of-sample volatility forecasting. Using a GARCH(1,1) model for the original returns as the benchmark, it is shown that modeling the seasonal improves forecasting performance.

We also seek the best method of modeling the intraday seasonal volatility pattern for the DEM/USD and YEN/USD exchange rate returns, measured by the out-of-sample forecasting performance. First, we assess whether the seasonal pattern should be determined using squared returns or using the log of the squared returns. We recommend the latter as the best method. Second, we compare two methods that first determine the seasonal and then analyze the deseasonalized return series. The Fast Fourier Form improves over the use of the mean of the squared returns for each intraday interval. Third, these two-step methods are compared with the P-GARCH model of Bollerslev and Ghysels (1996). The latter could be seen as a simultaneous optimization of the seasonal pattern and the dynamic volatility pattern. We find that the P-GARCH model provides the best forecasts out-of-sample. All of these conclusions apply to both currencies.

There is a strong case for giving preference to the two-step method based on the FFF, rather than to the P-GARCH method. First, it is computationally more

efficient. Second, the forecasting performance of the P-GARCH model is only marginally better than the performance of the two-step FFF. Third, P-GARCH potentially suffers from the restriction that the amount of historical data used for the seasonal is the same as for the GARCH parameters.

The results provide important insights into the different methods of modeling the intraday volatility seasonal pattern. They show that it is worthwhile to do so when forecasting intraday volatility, and they also show the most promising way to do so. Our results help to solve the problem of selecting from the wide variety of approaches available for modeling seasonal volatility.

References

- Andersen, T. G. and T. Bollerslev, 1997, Intraday periodicity and volatility persistence in financial markets, *Journal of Empirical Finance* 4, 115-158.
- Andersen, T. G. and T. Bollerslev, 1998, DM-Dollar volatility: Intraday activity patterns, macroeconomic announcements and longer run dependencies, *Journal of Finance* 53, 219-265.
- Bollerslev, T. and I. Domowitz, 1993, Trading patterns and prices in the interbank foreign exchange market, *Journal of Finance* 48, 1421-1443.
- Bollerslev, T. and E. Ghysels, 1996, Periodic Autoregressive Conditional Heteroscedasticity, *Journal of Business & Economic Statistics* 14, 139-151.
- Canina, L. and S. Figlewski, 1993, The informational content of implied volatility, *Review of Financial Studies* 6, 659-681.
- Day, T. E. and C. M. Lewis, 1992, Stock market volatility and the information content of stock index options, *Journal of Econometrics* 52, 267-287.
- Drost, F. C. and T. E. Nijman, 1993, Temporal aggregation of GARCH processes, *Econometrica* 61, 909-927.
- Goodhart, C. A. E. and L. Figliuoli, 1991, Every minute counts in financial markets, *Journal of International Money and Finance* 10, 23-52.
- Pagan, A. R. and G. W. Schwert, 1990, Alternative models for conditional stock volatility, *Journal of Econometrics* 45, 267-290.
- Payne, R., 1996, Announcements effects and seasonality in the intraday foreign exchange market, London School of Economics Financial Markets Group discussion paper no. 238.
- Taylor, S. J. and X. Xu, 1997, The incremental volatility information in one million foreign exchange quotations, *Journal of Empirical Finance* 4, 317-340.

TABLE 1: In-Sample Fit, January - September 1996

Panel A: DEM/USD

	Raw	R ²		FFF		P-GARCH (eq. 19)
		R ²	lnR ²	R ²	lnR ²	
μ	0.001 (0.001)	0.020 (0.010)	0.047 (0.014)	0.019 (0.009)	0.043 (0.018)	0.010 (0.006)
ω	0.001 (0.000)	0.011 (0.024)	1.068 (0.337)	0.083 (0.124)	0.983 (0.405)	-
α	0.290 (0.028)	0.038 (0.037)	0.157 (0.031)	0.104 (0.088)	0.151 (0.038)	0.175 (0.044)
β	0.604 (0.032)	0.953 (0.058)	0.672 (0.082)	0.826 (0.196)	0.692 (0.101)	0.605 (0.130)
ν	3.716 (0.153)	4.813 (0.262)	4.117 (0.187)	4.232 (0.204)	4.193 (0.194)	4.226 (0.216)
Log-likelihood	13,745	-12,631	-20,476	-12,590	-20,531	-7,558
Adj. log-lik.	13,745	14,024	14,074	13,903	14,019	13,994

Panel B: YEN/USD

	Raw	R ²		FFF		P-GARCH (eq. 19)
		R ²	lnR ²	R ²	lnR ²	
μ	0.001 (0.001)	0.013 (0.006)	0.036 (0.029)	0.014 (0.131)	0.034 (0.018)	0.010 (0.005)
ω	0.001 (0.000)	0.053 (0.029)	0.586 (0.235)	0.082 (0.012)	0.578 (0.269)	-
α	0.229 (0.025)	0.104 (0.031)	0.143 (0.034)	0.131 (0.007)	0.140 (0.038)	0.157 (0.010)
β	0.654 (0.037)	0.852 (0.056)	0.771 (0.067)	0.803 (0.013)	0.776 (0.077)	0.715 (0.023)
ν	3.780 (0.155)	4.810 (0.235)	4.028 (0.170)	4.130 (0.002)	4.052 (0.176)	4.019 (0.001)
Log-likelihood	11,286	-12,537	-20,407	-12,300	-20,458	11,424
Adj. Log-lik.	11,286	11,472	11,513	11,388	11,462	11,424

Note: GARCH(1,1) models are estimated for the first nine months of 1996 (9,360 observations). The first column ("Raw") is for the original thirty-minute returns. The second column ("R²") is for returns divided by the seasonal based on averages of squared returns (R²) or averages of the log of squared returns (lnR²). The third column ("FFF") is for returns divided by the seasonal based on the FFF for squared returns (R²) or the log of squared returns (lnR²). The fourth column ("P-GARCH") is for the P-GARCH model defined in equations (18) and (19). The log-likelihood is based on the data used to estimate the GARCH model. Hence, for the R²- and FFF-methods, that use the deseasonalized returns, the adjusted log-likelihood is also reported, where the residuals are multiplied by the seasonal and the estimated conditional variances are multiplied by the square of the seasonal. μ is the conditional mean, ω , α and β are conditional variance parameters and ν is the degrees of freedom of the conditional (student-t) distribution. Heteroscedasticity consistent standard errors are shown inside parentheses.

TABLE 2: Out-Of-Sample Forecast Performance, October - December 1996**Panel A: DEM/USD**

	Raw	R ²		FFF		P-GARCH (eq. 19)
		R ²	lnR ²	R ²	lnR ²	
Correlation with realized ($ R_t $)	0.245	0.267	0.273	0.268	0.294	0.298
Mean forecast (realized: 0.048)	0.046	0.047	0.047	0.046	0.045	0.044
RMSE ($\times 10^{-2}$)	5.290	5.302	5.308	5.249	5.201	5.168
LL	1.741	1.696	1.687	1.691	1.642	1.636
Regression adjR ²	0.060	0.071	0.074	0.072	0.086	0.089
degrees of freedom						
mean	3.80	4.43	4.14	4.15	4.38	4.47
std	0.12	0.37	0.31	0.25	0.26	0.30
max	4.14	5.25	4.74	4.66	4.95	5.24
min	3.60	3.83	3.59	3.47	3.91	3.92

Panel B: YEN/USD

	Raw	R ²		FFF		P-GARCH (eq. 19)
		R ²	lnR ²	R ²	lnR ²	
Correlation with realized ($ R_t $)	0.237	0.227	0.251	0.264	0.269	0.270
Mean forecast (realized: 0.0575)	0.057	0.060	0.058	0.057	0.057	0.056
RMSE ($\times 10^{-2}$)	5.799	5.873	5.812	5.751	5.750	5.736
LL	1.736	1.760	1.715	1.712	1.697	1.686
Regression adjR ²	0.056	0.052	0.063	0.069	0.072	0.073
Degrees of freedom						
Mean	3.70	4.33	3.90	4.02	4.00	3.95
Std	0.07	0.27	0.11	0.07	0.07	0.08
Max	3.81	4.82	4.06	4.19	4.09	4.05
Min	3.56	3.91	3.67	3.81	3.83	3.77

Note: The 2963 forecasts for the last three months of 1996 are compared with the realized volatility measured by the absolute return. The models are repeatedly re-estimated every twenty-four hours using data from the 60 days (DEM/USD) or 195 days (YEN/USD) preceding the forecasting period. The first column ("Raw") is for the GARCH(1,1) model for the original thirty-minute returns. The second column ("R²") is for the GARCH(1,1) model for returns divided by the seasonal based on averages of squared returns (R²) or averages of the log of squared returns (lnR²). The third column ("FFF") is for returns divided by the seasonal based on the FFF for squared returns (R²) or the log of squared returns (lnR²). The fourth column ("P-GARCH") is for the P-GARCH model defined in equations (18) and (19). 'Regression adjR²' relates to the adjusted coefficient of determination in the regression $|R_t| = a + b\hat{s}_t + e_t$ where \hat{s}_t is the volatility forecast for one of the models. The last part of the table shows some statistics for the estimated degrees of freedom of the scaled t distribution.

TABLE 3: Encompassing Regressions, Coefficients for Explanatory Variables

Panel A: DEM/USD						Panel B: YEN/USD						
	Raw	R ²		FFF		P-GARCH	Raw	R ²		FFF		P-GARCH
row		R ²	lnR ²	R ²	lnR ²			R ²	lnR ²	R ²	lnR ²	
1	0.305 (4.8)	0.411 (7.7)					0.502 (6.5)	0.329 (5.3)				
2	0.327 (5.7)		0.414 (8.9)				0.379 (4.5)		0.441 (6.4)			
3	0.272 (4.0)			0.487 (7.4)			0.139 (1.2)			0.723 (6.6)		
4	0.124 (1.7)				0.640 (9.4)		0.211 (2.2)				0.640 (7.5)	
5	−0.027 (−0.32)					0.959 (9.7)	−0.090 (−0.69)					0.977 (7.4)
6		0.278 (3.8)	0.350 (5.1)					0.200 (2.6)	0.504 (6.5)			
7		0.302 (3.9)		0.380 (4.2)				0.118 (1.5)		0.724 (7.7)		
8		0.116 (1.5)			0.618 (7.1)			0.154 (2.2)			0.661 (8.4)	
9		0.209 (3.5)				0.707 (8.3)		0.181 (2.7)				0.732 (8.7)
10			0.340 (5.6)	0.353 (4.8)					0.291 (3.7)	0.559 (5.9)		
11			0.160 (2.2)		0.567 (6.5)				0.129 (1.2)		0.663 (5.6)	
12			0.232 (4.1)			0.670 (7.9)			0.233 (2.8)			0.657 (6.4)
13				0.085 (0.86)	0.657 (6.8)					0.380 (3.2)	0.482 (4.4)	
14				0.191 (2.5)		0.740 (7.8)				0.358 (3.0)		0.566 (4.5)
15					0.346 (3.7)	0.549 (4.7)					0.395 (3.3)	0.494 (3.6)
16	−0.031 (−0.37)		0.158 (2.1)	0.187 (1.5)	0.577 (4.2)		−0.053 (−0.41)		0.103 (0.97)	0.296 (1.9)	0.537 (2.9)	

Note: Coefficients and t-values (inside parentheses) for encompassing regressions. Using, for example, two models, we estimate the regression

$$|R_t| = \mathbf{a} + \mathbf{b}_1 \hat{\mathbf{S}}_t^{(1)} + \mathbf{b}_2 \hat{\mathbf{S}}_t^{(2)} + \mathbf{e}_t$$

using OLS where $|R_t|$ is the absolute realized 30-minute return, and $\hat{\mathbf{S}}_t^{(i)}$ is the volatility forecast for model i . The rows of the table then report the regression coefficients β_1 and β_2 and their t-values. In total there are 2963 forecasts for each model for the last three months of 1996. All models are estimated every twenty-four hours using data from the 60 days (DEM/USD) or 195 days (YEN/USD) preceding the forecasting period. “Raw” indicates the original thirty-minute returns are used in a GARCH(1,1) model. “R²” indicates the original returns are first divided by the seasonal based on averages of squared returns (R²) or the log of squared returns (lnR²) and only then the GARCH(1,1) model is estimated for the resulting deseasonalized returns. “FFF” indicates the original returns are divided by the seasonal based on the FFF for squared returns (R²) or the log of squared returns (lnR²). “P-GARCH” is the P-GARCH model defined in equations (18) and (19).

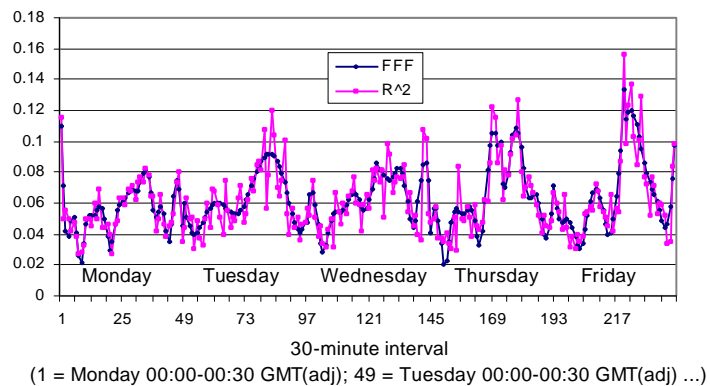


Figure 1a: DEM/USD volatility seasonals based on squared returns

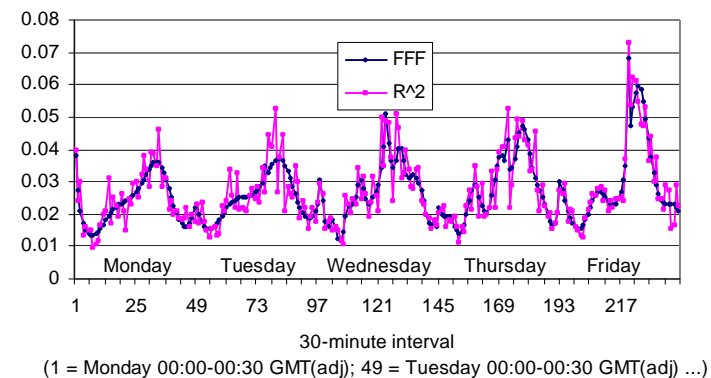


Figure 1b: DEM/USD volatility seasonals based on the log of squared returns

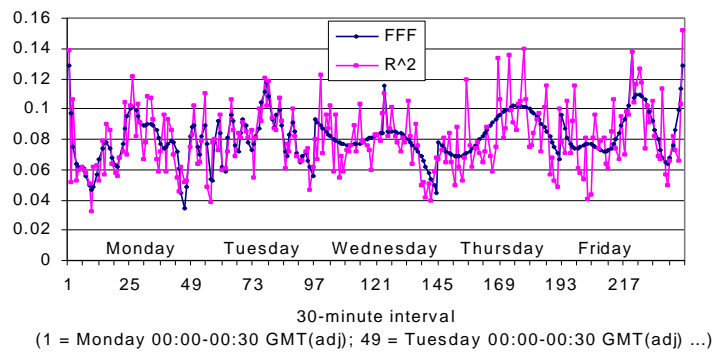


Figure 1c: YEN/USD volatility seasonals based on squared returns

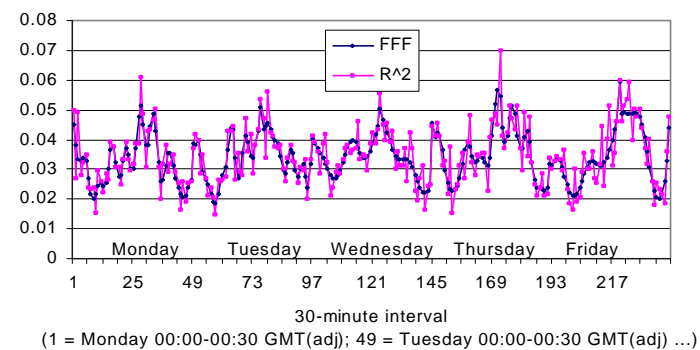


Figure 1d: YEN/USD volatility seasonals based on the log of squared returns