THE LEVERAGE EFFECT IN REACTIONS OF REALIZED VOLATILITY TO MACROECONOMIC NEWS SURPRISES

BACHELOR THESIS
DENNIS VAN DE WERKEN

Abstract

In this paper a HARST and AR(22) model are estimated for realized volatility for several indices and exchange rates. A comparison is made between models where only macroeconomic news announcement date dummies are included and models where actual news surprises are added as explanatory variables. The comparisons are made on both fit and forecast strength. It is shown that the newer HARST model outperforms the HAR and AR models in both forecasting and fitting strength. Further, it is shown that the used (leveraged) news surprises provide better fits but provide no significant increases in forecasting power.

1 Introduction

The realized volatility of an asset is important for several reasons. It is often used when estimating Value-at-Risk measures and can be of great importance to risk managers in general. Further, option traders can greatly profit from correctly estimating the realized volatility. Overall, it is a measure that can say a lot about a stock or market.

Several different models exist for fitting and forecasting realized volatility, of which some have been developed only recently. The realized volatilities of most assets share some interesting properties, among which relatively long memory, the leverage effect and seasonality effects. Most current research focuses on modelling these properties as accurately as possible. Another interesting and intuitively understandable feature of realized volatility is its (positive) response to macroeconomic announcements. This is usually accounted for by including dummies for news-announcements as explanatory variables.

The usual approach ignores the fact that news announcements might have different effects depending on the announced news. That is, it is assumed that positive macroeconomic news has the same effect as negative news. With this approach the magnitude of the announced variable is not taken into account either. To research this assumption the following hypothesis can be examined.

Hypothesis

The magnitude and sign of surprises on the announcement day of a macroeconomic variable have an influence on realized volatility beside that of the news occurring.

Macroeconomic news surprises might be of influence on realized volatility intuitively due to macroeconomic variables being taken into consideration by portfolio managers when estimating optimal portfolio weights. When a variable unexpectedly changes this must be of influence on the expected returns and risks for each of an investors underlying assets. The new estimation of their models might cause extra volatility in the stock prices for several days to come. Of course, other reasonings might also apply, as macroeconomic variables are likely to have an effect on several areas across the market.

A well-known fact about volatility is that it tends to rise more strongly on negative news than on positive news, the so-called leverage effect. In most current models news is taken as the lagged return, and variations in response to this lagged return are allowed to differ between positive sign and negative sign. It has been shown that incorporating this effect leads to significantly better models, see Martens, De Pooter & Van Dijk (2009), among others. Other models work with (smooth) regime switching or kernel estimation procedures to determine the effect that news has on realized volatility. Usually, the effect of macroeconomic news entering the market is taken to be constant. The idea that the reaction to macroeconomic news might incorporate a leverage effect as well is intuitively feasible.

1.1 Relevant Literature

Martens, De Pooter & Van Dijk (2009) compare several models in forecasting and fitting strength for S&P 500 realized volatility, among which the implied volatility in the Riskmetrics model, several Autoregressive Fractionally Integrated (ARFI) models, an Heterogeneous Autoregressive (HAR) model and higher order Autoregressive (AR) models. They also examine the effect of explicitly incorporating four stylized facts of realized volatility. Firstly, the leverage effect, which states that there is a negative relationship between news, as indicated by lagged returns, and volatility. Secondly, irregular level shifts, described in the same article as “occasional structural breaks”. Thirdly, seasonality effects, where different days of the week show different average volatility. Finally, the fact that volatility seems to be higher on days with macroeconomic announcements. They find that the higher order AR approximation with 22 lags gives the best fit and forecasts for several horizons.

McAleer & Medeiros (2008) propose an extension to the Heterogeneous Autoregressive model (HAR) by allowing a non-linear effect of past volatilities, where the effect is scaled by a smooth transition between regimes based on the return of the previous day. They name this model a Heterogeneous Autoregressive Smooth Transition (HARST) model. Due to it's multiple regimes based on past returns this model can account for the leverage effect as a response to lagged returns, which the original HAR model by Corsi (2004) can not.

Chuliá, Martens & Van Dijk (2010) study the effect of federal funds target rate decisions on individual S&P 100 stock returns, volatilities and correlations at the intra-day level and find that an asymmetry exists between reactions to positive and negative surprises. That is, a leverage effect in macroeconomic news. They do warn that these insights are the result of using intra-day data and that the effect is much less convincing on daily data.

1.2 Methodology

In this paper it is attempted to investigate the hypothesis by incorporating macroeconomic news surprises into recently developed models for realized volatility and comparing fit and forecast power. To do this, the HARST and AR(22) model as mentioned in Section 1.1 and detailed extensively in Section 3 are estimated for stock index realized volatility with several additions of macroeconomic news. Firstly, they are estimated for the case where only the announcement date dummies are included, as is usual in the current literature. Secondly, for the case where *absolute* macro-economic surprises are added. And finally, adding macro-economic surprises in a manner which allows a leverage effect to exist. The results are compared by F-tests for significant model improvement.

The surprises are added in a look-ahead manner, that is, the surprises at time *t* are added as explanatory variables on day *t*, even if they are not known in advance. This gives an examination of the real effect of surprises, rather than a realistically applicable method for prediction. If the surprises show to be of importance, obtaining real-time forecasts by proxying surprises by their expected values is considered.

To examine the effect of macroeconomic surprises on forecast strength the same process is repeated with a moving window estimation. The first 1200 observations are used for the initial parameter estimates, after which a month (22 days) of realized volatility is forecasted. Then, one-day-ahead forecasts are made, using the parameters estimated for that month. After this, the model parameters are re-estimated, moving the estimation window ahead by 22 days. This process is repeated for the rest of the sample period. Finally, the forecast errors of the forecasted series are compared by Diebold-Mariano tests.

For robustness and to compare differences between indices and exchange rates, the procedure described above is applied to the S&P500 index, the S&P400 index, the Dow Jones Index, the EUR/USD exchange rate and the JPY/USD exchange rate.

1.3 Main Results

When estimating the HAR, HARST and AR(22) models for the full-sample period for the S&P500 index, an increase in fit-strength is shown when comparing a model with absolute macroeconomic surprises to a model with announcement day dummies only. None of these increases are significant at the 5% level. When comparing a model with leveraged surprises to the model with announcement day dummies, significant increases in fit are shown at the 5% level.

Forecasting realized volatility gives multiple interesting results. The comparison between models shows that when predicting stock index realized volatility, the AR(22) model does not give significantly better or worse predictions than the HAR model at the 5% level. The HARST model with properly selected transition variable and number of limiting regimes outperforms the HAR and AR(22) models significantly in all cases.

Comparing the forecast strength for stock index realized volatility models with absolute and leveraged surprises to those with only announcement day dummies shows no significant differences at the 5% level. That is, adding preselected real-time surprises does not give rise to better predictions. In fact, even though none of the found p-values indicate significant differences, often, the effect of adding the macroeconomic surprises is negative rather than positive for forecast strength, as indicated by the sign of the Diebold-Mariano test statistic.

Estimating the models for exchange rates rather than stock indices shows different results. Here, adding macroeconomic surprises does not lead to a significant increase or decrease in forecast strength either. However, the comparison between models shows that HARST now does not outperform AR(22) or HAR significantly. This is intuitively feasible due to the fact that exchange rates contain a symmetry that stock indices do not: when, for instance, the EUR/USD exchange rate has a negative return, the USD/EUR exchange rate has a positive return, and vice versa.

The rest of this paper is organized as follows. Section 2 describes the used data. In Section 2.1, the existence of a leverage effect in response to returns is examined. Further, in Section 2.2, the available macro-economic variables are subjected to a first selection procedure for inclusion in the more extensive models. Section 3 gives details on the models into which the macroeconomic news is incorporated, specifically, the AR(22) model is described in Section 3.1 and the HARST model is given in Section 3.2. Section 4 shows the estimation process and results for both models, it examines the effect of incorporating macroeconomic news surprises for the fit strength of the full sample. Section 5 compares forecast strengths between several models by Diebold-Mariano tests. Section 6 gives the discussion of the results and methods used. Finally, Section 7 concludes.

2 Data

2.1 Realized Volatility and Returns

The dataset for realized volatility and returns is the Oxford-Man Realized Library[1] Version 0.2, of which the construction is described in Heber et al. (2009). It contains daily returns, realized volatility and realized kernel on numerous important stock indices and exchange rates. In this paper the main focus will be on the S&P500 index. The results are compared to the S&P400 index, the Dow Jones index and the EUR/USD and JPY/USD exchange rates, all of which have been obtained from this dataset. The data for each series are available for different time periods, but in this article the sample period January 1st, 2000 – May 23rd, 2008 is considered. Data is available for this sample period for all indices and exchange rates.

The two figures below give a histogram and descriptive statistics for the S&P500 returns and realized volatility series respectively. The descriptive statistics are typical for their types, with long tails for the stock returns and a decrease of occurrence of higher realized volatilities. The other indices and exchange rates show similar features.

[1] http://realized.oxford-man.ox.ac.uk/data

Figure 1. Histogram and Stats for S&P500 index returns

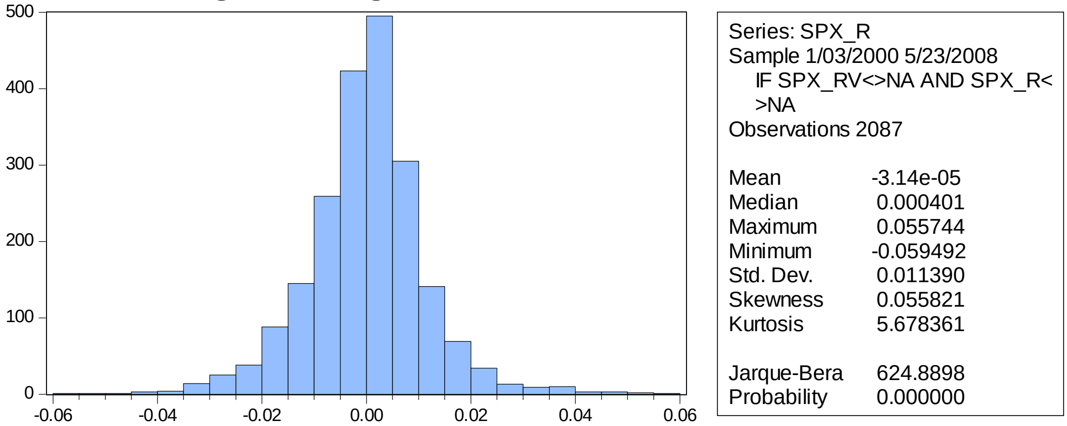
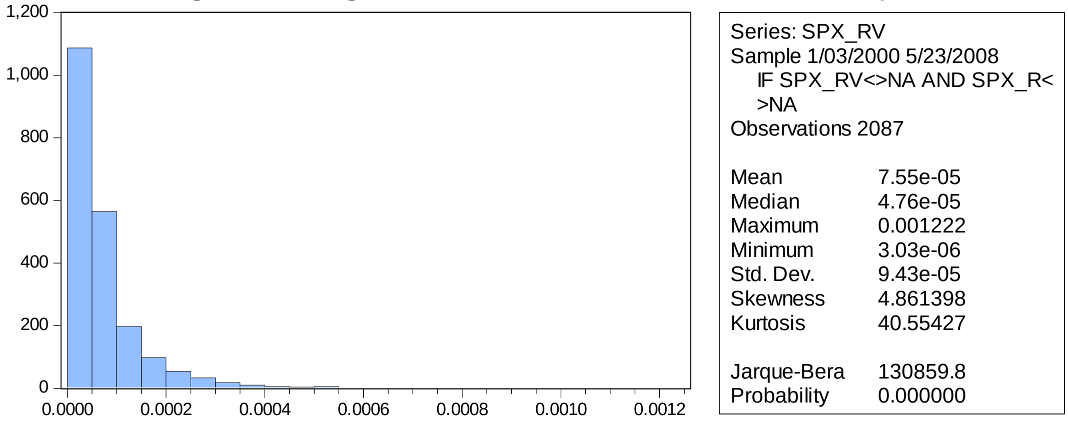


Figure 2. Histogram and Stats for S&P500 realized volatility



In the current literature it is common to model for the leverage effect in realized volatility. Realized volatility is usually higher in times of large absolute returns. The leverage effect says that this increase in volatility is more pronounced on negative returns than on positive returns. While modelling for the leverage-effect in reaction to macroeconomic news is the main focus of this article, the original leverage effect in response to lagged returns cannot be ignored. The two main models used in this paper both account for a leverage effect in response to lagged returns. To get a first indication of the existence of such an effect, a linear regression of the following model can be used. The so-called *news impact curve*, see Engle & Ng (1993). It is given by

$$y_t = \beta_0 + \beta_1 |z_t| + \beta_2 \, \mathrm{I}[z_t < 0] + \beta_3 |z_t| \, \mathrm{I}[z_t < 0] + \varepsilon_t, \qquad (1)$$

where y_t is log realized volatility, $z_t = (r_{t-1})/\hat{\sigma}_r$ is the lagged return divided by return sample standard deviation, I[A] is an indicator function equal to 1 if event A is true and 0 otherwise, and ε_t is an error term. In this model, β_1 indicates the effect of absolute return and the combination of β_2 and β_3 gives the strength of the levarge effect. That is, the realized volatility on days with negative (lagged) returns can have a different mean and a different slope in the response to absolute returns. The estimated news impact curve is given in Table 1 below for all examined indices and exchange rates. A Wald-test is applied for the restriction $\beta_2 = \beta_3 = 0$, giving an indication of whether the leverage effect exists for that series.

Table 1. News impact curve for several indices and exchange rates

	S&P500	S&P400	DJI	USD/EUR	USD/JPY
Lev. Eff. Wald p-value	0.0000	0.0000	0.0000	0.1108	0.0049
β_0	**-10.4288** **(0.0349)**	**-10.4730** **(0.0295)**	**-10.4577** **(0.0332)**	**-10.4081** **(0.0327)**	**-10.4057** **(0.0313)**
β_1	**0.5029** **(0.0309)**	**0.3350** **(0.0282)**	**0.4769** **(0.0306)**	**0.2811** **(0.0328)**	**0.2422** **(0.0327)**
β_2	**0.1539** **(0.0521)**	**0.1288** **(0.0469)**	**0.1698** **(0.0555)**	0.0197 (0.0476)	0.0441 (0.0431)
β_3	**0.1872** **(0.0488)**	**0.1946** **(0.0477)**	**0.1262** **(0.0591)**	0.0495 (0.0443)	0.0537 (0.0443)

Note: White estimation errors are given between parentheses. Estimates significantly differing from zero at the 5% significance level are highlighted in bold.

The leverage effect is very pronounced in all three stock indices. The Dow Jones Index seems to react more to the occurrence of a negative return than to the magnitude of the return, when compared to the S&P indices. For the exchange rates, the leverage effect is less pronounced. Separately, the coefficients for terms including negative returns cannot be considered to significantly differ from zero. This is expected, as a leverage effect in, for instance, the USD/EUR exchange rate would signify a reversed leverage effect, where volatility increases more after high returns, in the EUR/USD exchange rate. This symmetry in the exchange rate would cause the leverage effect to be non-existent. When applying the Wald test to investigate the significance of the leverage effect as a whole, the existence of a leverage effect in the USD/JPY exchange rate can however not be ruled out.

2.2 Macroeconomic surprises

As in Balduzzi, Elton & Green (2001), macroeconomic surprises are generated from a series of announced macroeconomic variables and investor expectations of the announcements as

$$S_{i,t} = \frac{A_{i,t} - E_{i,t}}{\sigma_i} \tag{2}$$

where $A_{i,t}$ is the actual announcement and $E_{i,t}$ is the investor expectation. The surprises are standardized by dividing by σ_i, the sample standard deviation of $A_{i,t} - E_{i,t}$, to facilitate comparison.

On days without news it is assumed that $S_{i,t} = 0$. Note that when adding this series as an explanatory variable it is often important to also include dummies for the event that news actually occurred. A zero-surprise news announcement might yield a different average volatility than no announcement occurring at all.

The used data contains announcement dates and surprises for 24 macroeconomic variables, the exact types of which can be found in the Appendix. The surprises are available for the entire sample period of January 1st, 2000 – May 23rd, 2008.

The federal funds target rate decision has been shown to be a variable of great influence on realized volatility. However, no surprise data on these announcements is readily available for the selected time period. Because the announcements are discrete rather than continuous, in that they are normally a rise or fall of 0.25%, and sometimes happen on unexpected dates, for instance after big news events, proxying them by a simple AR model as in Gutker et al.

(2011) is not an option. Further, the surprises in changes in the federal funds target rate have generally been small in recent years, so adding these is not expected to give greatly significant results. For this reason, but to not ignore this important factor, the federal funds target rate announcements are not added as surprises but simply as announcement-day dummies, which are available for the entire period.

2.2.1 Variable selection

Because the models for realized volatility in this paper are non-linear by nature it is important to limit the amount of variables to estimate from the start. Further, for prediction accuracy, the amount of non-influential variables should be limited. To gain insight in which macroeconomic announcements to include in the broader models, the following model is estimated separately for all news-types. Note that it resembles the news impact curve in (1). The main differences are the inclusion of a news dummy and 5 autoregressive terms. The model is given by

$$\begin{aligned} y_t = \beta_0 + \beta_1 D_{i,t} + \beta_2 \left|S_{i,t}\right| + \beta_3\, \mathrm{I}\left[S_{i,t} < 0\right] + \beta_4 \left|S_{i,t}\right| \mathrm{I}\left[S_{i,t} < 0\right] \\ + \sum_{k=1}^{5} \varphi_k y_{t-k} + \varepsilon_t, \end{aligned} \tag{3}$$

where y_t is log realized volatility, $\left|S_{i,t}\right|$ is absolute news surprise of type i on day t and $\varepsilon_{i,t}$ is an error term. The estimations are compared to the restricted model where $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ and if these four variables are jointly significant, a Wald test is applied with the restriction $\beta_3 = \beta_4 = 0$, to see if a leverage effect is present. The limit significance level is taken as 10% to not discard valuable information.

Results of the selection process are given in Table 2 for the S&P 500 index, results for other indices and exchange rates and the full names of the variables corresponding to each news type are given in the Appendix.

Table 2. Test for macroeconomic announcement inclusion – S&P500 - All/Leverage Terms

News type	*p*-value	News type	*p*-value
GDP_A	**0.0054** (**0.0570**)	CS	0.6394
GDP_P	0.8468	FO	**0.0524** (0.4975)
GDP_F	0.6732	BI	0.7363
NFPE	**0.0000** (0.4500)	GBD	0.5275
RS	0.9225	TB	**0.0407** (**0.0971**)
IP	0.3916	PPI	**0.0674** (0.3669)
CU	0.7589	CPI	0.3524
PI	0.5529	CCI	**0.0186** (0.3604)
CC	**0.0424** (0.7668)	NAPM	**0.0453** (**0.0788**)
PCE	0.6404	HS	**0.0519** (**0.0847**)
NHSI	0.1708	ILI	**0.0007** (0.1327)
DGO	0.1330	IUC	**0.0236** (0.7129)

Note: The first value indicates an effect of the combined terms. If significant at the 10% level, the second value inidicates significance of the two leverage terms.

A total of 11 out of 24 variables are significantly of influence at the 10% level. For 4 out of the 11 variables, a leverage effect cannot be rejected at the 10% level.

A similar test has been applied for the FOMC target rate announcement dummies. Here, no surprises are included and an autoregressive model with and without the dummies are compared by a Wald test. The Wald test rejects the null-hypothesis of no influence of the target rate announcement dummies at a significance level of 0.0000.

The results of this selection process are used throughout the rest of the article. When macroeconomic announcement date dummies are added, only those that show significance of the first test are included. When adding absolute or leveraged surprises, only those with a significant second test result are included.

3 Models

This section contains an overview of the models used for fitting and predicting the realized volatility series. Note that in some cases different symbols are used than in the original paper/proposition, here they have been chosen such that they are consistent throughout the paper.

3.1 Higher order Autoregressive (AR) model

Martens, Van Dijk & De Pooter (2009) find that an AR(22)-DAXRL model when compared to ARFI-DAXRL and HAR-DAXRL models give slightly better results in both fit and forecast. The name DAXRL signifies the inclusion of day-of-the-week dummies (D), announcement-day dummies (A) and lagged returns (R). It further implies that the lagged returns are added as exogenous regressors (X), rather than in the time-varying μ_t and that the model can account for leverage effects (L).

The article gives the model in the form

$$\varphi(L)(y_t - \mu_t) = \beta_1 |r_{t-1}| + \beta_2 \, \mathrm{I}[r_{t-1} < 0] + \beta_3 |r_{t-1}| \, \mathrm{I}[r_{t-1} < 0] + \varepsilon_t \qquad (4)$$

where $\varphi(L)$ is a lag-polynomial, y_t is log realized volatility, μ_t is a time-varying mean, $\mathrm{I}[A]$ is an indicator function which equals 1 if event A is true and 0 otherwise, r_t is the asset return at time t and ε_t is an error-term. In the case of the AR(22)-DAXRL model, the time-varying mean is given by

$$\mu_t = c(t) + \boldsymbol{\alpha}' \boldsymbol{w}_t, \qquad (5)$$

with $\boldsymbol{\alpha}$ a parameter vector and w_t a vector containing the day-of-the-week dummies, announcement day dummies and holiday dummies. They find that the time-varying $c(t)$ does not add significant improvement to the model, so it is replaced in this paper by adding a constant term 1 to $\mathbf{w}_t$ and removing $c(t)$ from (2). Note that the AR(22) model nests a HAR model.

The model can be extended to include (leveraged) macroeconomic news surprises. For each news type i the following three terms are added to $\mathbf{w}_t$: $|S_{i,t}|$, $\mathrm{I}[S_{i,t} < 0]$ and $|S_{i,t}| \, \mathrm{I}[S_{i,t} < 0]$. In the case that no leverage effect is allowed for, only the first term is added. The parameter vector $\boldsymbol{\alpha}$ is extended to include extra parameters for the added terms. The dummies for the actual news announcements in $\mathbf{w}_t$ should not be removed as days without news might still have a different average volatility than days with a zero news surprise.

3.2 Heterogeneous Autoregressive Smooth Transition (HARST) model

McAleer & Medeiros (2008) propose an extension of the original HAR model by Corsi (2004) with non-linear effects of past returns modelled by smooth transitions between regimes.

They define this HARST model as

$$y_t = \boldsymbol{\alpha}' \boldsymbol{w}_t + \boldsymbol{\varphi}_0' \boldsymbol{x}_t + \sum_{m=1}^{M} \boldsymbol{\varphi}_m{}' \boldsymbol{x}_t f(z_t; \gamma_m, c_m) + \varepsilon_t, \tag{6}$$

where the variables in w_t are exogenous regressors, the variables in $\mathbf{x}_t$ are average realized volatility over previous periods, and ε_t is an error term. The periods for $\mathbf{x}_t$ are often chosen as 1 day, 5 days and 22 days, as they will be in this paper. Note that $\boldsymbol{\varphi}_0$ and $\boldsymbol{\varphi}_m$ are vectors containing three parameters, one each for the 1 day, 5 day and 22 day effects. The variable z_t determines the current regime, it can potentially be any variable, but in the original article the lagged return is proposed to capture the leverage effect. A possible function $f(z_t; \gamma_m, c_m)$ is the logistic transition function given by

$$f(z_t; \gamma_m, c_m) = \frac{1}{1 + \exp\left(-\gamma_m (z_t - c_m)\right)}, \tag{7}$$

which defines the smooth regime transitions. The parameter estimates for the γ and c parameters can be estimated by non-linear least squares. The amount of limiting regimes $M + 1$ remains to be chosen, for this McAleer & Medeiros (2008) provide a detailed specific-to-general procedure which is followed directly.

The motivation for using regime switches is that it allows the model to capture a leverage effect in response to lagged returns. The effect of the HAR-terms in $\boldsymbol{x}_t$ can be different after (very) positive lagged returns versus (very) negative lagged returns. It allows for a leverage effect without being too strict in its definition; it can occur about any value, and with any smoothness. The possible existence of more than two regimes is also a positive feature.

To include macroeconomic surprises, a similar extension can be made as in 3.1, including them exogenously in $\boldsymbol{w}_t$.

4 Estimation

4.1 HARST Model

4.1.1 Transition Variable

The transition variable to be used for the rest of the comparisons in the HARST-model can be chosen by approximating the logistic transition function by a third-order Taylor Expansion around zero. That is, the realized volatility series is regressed on $\boldsymbol{x}_t$, $\boldsymbol{x}_t z_t$, $\boldsymbol{x}_t z_t^2$ and $\boldsymbol{x}_t z_t^3$. The non-linear effect of the transition variable can then be tested by comparing the full model to a model where the coefficients for $\boldsymbol{x}_t z_t^2$ and $\boldsymbol{x}_t z_t^3$ are restricted to zero. The best choice of transition variable is given by the transition variable for which an LM-test rejects the null-hypothesis of non-linearity the strongest.

Table 3 shows the results for all five series and 6 possible transition variables. The Dow Jones and S&P500 index both give a selection of r_{t-1}, the lagged return, which gives a model that models the most often examined leverage effect. The S&P400 index is better modelled by allowing regime switches based on lagged realized volatility. For the exchange-rates longer lags are selected for the moving average. With the USD/EUR exchange rate being modelled best by different regimes when the past 22 day moving average shows differences, and the USD/JPY rate having a preferred transition variable of the past 5 day average realized volatility.

Table 3. Transition Variable non-linearity test

	r_{t-1}	$r_{t-1}^{(5)}$	$r_{t-1}^{(22)}$	y_{t-1}	$y_{t-1}^{(5)}$	$y_{t-1}^{(22)}$
S&P 500	**5.4201**	1.8169	1.4676	3.1718	2.7355	1.1505
S&P 400	4.9045	2.6996	2.7966	**7.7992**	3.2770	1.6848
DJI	**5.5808**	2.6516	2.7595	4.1101	3.0439	1.6044
USD/EUR	2.6985	3.9011	**4.9169**	1.5856	1.6351	1.3652
USD/JPY	2.4934	3.6129	5.2101	1.8957	**5.6011**	0.7712

4.1.2 Number of limiting regimes

The number of regimes in the HARST model is chosen by the procedure outlined in McAleer & Medeiros (2008). It develops a test based on the null-hypothesis that no non-linear effect remains in the residuals, after using M - 1 regimes to estimate a HARST-model, or $\gamma_M = 0$. The test is applied by regressing the residuals of the estimate with M - 1 regimes on a third order Taylor-expansion of the logistic function around $\gamma_M = 0$ and comparing the sum of the resulting residuals to the sum of the original residual series.

The resulting LM test-statistic is asymptotically distributed as χ^2 with $3(p + 1)$ degrees of freedom, where p is the amount of the HAR-terms used, in this case, p = 3. Of importance when applying this test when using one of the HAR-terms as the transition variable is to not include a constant in $\boldsymbol{x}_t$ as this leads to a singular matrix **v**, using the notation of McAleer & Medeiros (2008).

The test can be applied consecutively for increasing values of M. It is repeated until the first rejection outcome, which then gives the optimal value for M as the last value for which it was not rejected. McAleer & Medeiros (2008) suggest using a decreasing significance level when applying the test to limit the overall significance level. When the rejection significance level is defined as $\lambda_J = \lambda_1 C^J$, where J is the currently tested regime and C is some arbitrary constant, the overall significance level has an upper bound of $\sum_{j=1}^{J} \lambda_j$. Here λ_1 is chosen as 0.025 and C as 0.5, such that the overall significance level has an upper bound of 0.5 as $J \rightarrow \infty$.

For simplicity, it is assumed that the addition of exogenous variables to the model will not have an effect on the optimal amount of regimes. Under this assumption, it is possible to estimate the optimal amount of regimes once, and apply the comparisons in the rest of the paper to the optimal amount of regimes only.

The following table details the results of the LM-test. The values for non-selected transition variables are added for comparison purposes.

Table 4. LM-tests for selection of Limiting Regimes

	r_{t-1}		$r_{t-1}^{(5)}$		$r_{t-1}^{(22)}$		y_{t-1}		$y_{t-1}^{(5)}$		$y_{t-1}^{(22)}$	
S&P 500	**1**	**0.0000**	1	0.6018	1	0.6049	1	0.1977	1	0.0686	1	0.7456
	2	**0.0264**										
S&P 400	1	0.0000	1	0.0161	1	0.2014	**1**	**0.0458**	1	0.0019	1	0.7630
	2	0.1482	2	0.0171					2	0.0284		
DJI	**1**	**0.0000**	1	0.4275	1	0.5537	1	0.3439	1	0.1391	1	0.9055
	2	**0.0167**										
USD/EUR	1	0.0000	1	0.0113	**1**	**0.0106**	1	0.0038	1	0.0000	1	0.0020
	2	0.0015	2	0.1294	**2**	**0.4306**	2	1.0000	2	0.2884	2	0.4292
	3	0.1609										
USD/JPY	1	0.0000	1	0.0337	1	0.0005	1	0.0144	**1**	**0.0000**	1	0.6462
	2	0.0506			2	0.0557	2	0.9423	**2**	**0.0822**		

Note: When a p-value is high, this implies that using (M+1) regimes is not expected to give a significantly better fit than (M+1)-1 regimes. The parameter M signifies the number of transition functions, thus the total number of regimes used is (M+1)-1=M for the first rejection value of M.

When the result for $M = 1$ is not significant, as in most cases, this indicates that a HARST-model is not able to perform significantly better then an ordinary HAR-model. In other words, no regime switches based on that transition variable seem to take place. It is visible that two regimes are preferred for all indices and exchange rates with the selected transition variable, except for the S&P400 index, for which one regime, i.e., a HAR model, suffices.

4.1.3 Coefficient Interpretations

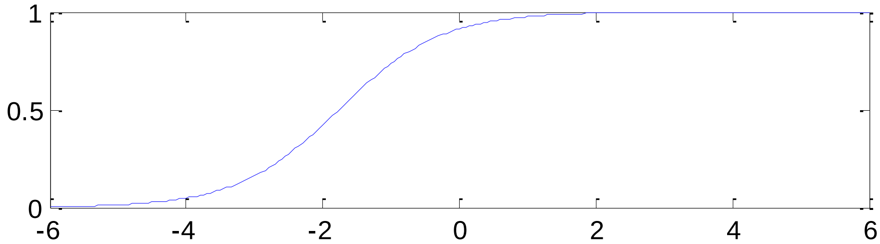
When interpreting coefficient estimates for the HARST-model one should be careful not to see the constants as a direct indication of average log volatility in a limiting regime. A typical estimation of a HARST(2) model without exogenous variables is given below.

Table 5. Typical Phi estimates in HARST model
(S&P 500 full sample, no exogenous regressors)

	Regime 1	Regime 2
γ_m		1.3372
c_m		-1.7670
$\varphi_{0,m}$	-2.7793	2.0608
$\varphi_{1,m}$	0.3803	-0.1102
$\varphi_{5,m}$	0.0663	0.4270
$\varphi_{22,m}$	0.1738	0.0053

In this estimate the lagged returns have been divided by their standard deviation for easy interpretation. The found transition function parameters give a transition function as in the following figure.

Figure 3. Logistic transition function for $\gamma = 1.3372$, $c = -1.7671$



When naively interpreting the sum of constants $\varphi_{0,m}$ as an indication of average realized volatility in regime m+1, this would seem to indicate that regime 2 shows higher volatility. That is, the volatility after high/positive returns is higher than that after low/negative returns, in contrast with the leverage effect in realized volatility. However, one should take note that Regime 2 shows a higher effect of the 5-day HAR-term than Regime 1. This HAR-term is generally negative due to the fact that it is an average of past log realized volatility, which has an average of approximately -9. When taking this into account, the average fitted value for realized volatility in Regime 2 is actually lower than in Regime 1.

To show this, one can estimate the same HARST model as before, but now using as HAR-terms the deviation from the mean log realized volatility over the past h days, rather than the average log realized volatility over the past h days. This modification has no influence on the actual model estimates, but gives a different interpretation to the sum of $\varphi_{0,m}$ coefficients. They now do give an indication of average log realized volatility level in regime m+1. The following table shows an estimate of the same model as before, but using the alternative definition of the HAR-terms.

Table 6. Typical Phi estimates in HARST model – alternative HAR-terms (S&P 500 full sample, no exogenous regressors)

	Regime 1 (m = 0)	Regime 2 (m = 1)
γ_m		1.3372
c_m		-1.7671
$\varphi_{0,m}$	-8.9406	-1.1382
$\varphi_{1,m}$	0.3803	-0.1102
$\varphi_{5,m}$	0.0663	0.4270
$\varphi_{22,m}$	0.1738	0.0053

Now, it is visible that the average realized volatility in Regime 2 is indeed quite remarkably lower than in Regime 1. All other estimates have remained the same.

Table 5 shows interesting results for the HARST(2)-model besides accounting for a leverage effect in response to lagged returns. It also shows that after positive returns, weekly average realized volatility is more indicative for future volatility, while after negative returns, daily volatility plays a more important role.

The monthly volatility keeps an influence that does not differ much between the two regimes. This is an intuitive result. The HAR-model is based on the economic intuition that different kinds of traders have an influence on realized volatility, see Corsi (2004). Long-term traders, which are indicated here as having a monthly influence on realized volatility, are not likely to respond to the present-day lagged return. Rather, they examine monthly patterns in returns. Estimating the model where the average return over the past 22 days is used for z_t confirms this. See the table below.

In this table, a comparable transition function is found, however, in Regime 2 a definite increase in influence of the 22-day and 5-day average realized volatility is seen after high returns. That is, if there were higher than average returns in the past 22 days, there will be more dependence of current volatility on the past 22 day average volatility.

Table 7. Coefficient estimates for HARST model – alternative HAR-terms (S&P 500 full sample, no exogenous regressors, tr. var.: 22-day MA return)

	Regime 1	Regime 2
γ_m		1.5898
c_m		-1.5074
$\varphi_{0,m}$	-9.1423	-0.8958
$\varphi_{1,m}$	0.4000	-0.0533
$\varphi_{5,m}$	0.0071	0.2992
$\varphi_{22,m}$	0.1442	0.1214

4.2 AR(22)

Estimating the AR(22) model as given in (2) and (3) for the S&P500 log realized volatility, using endogenous weekday dummies and adding the news impact curve exogenously, yields the following results.

Table 8. Parameter estimation AR(22) model, full sample S&P 500

Parameter	Estimation	Std. Error	Parameter	Estimation	Std. Error
$\alpha_{CONSTANT}$	-10.655	0.0118	φ_8	0.0315	0.0232
α_{MONDAY}	-0.2059	0.0491	φ_9	-0.0001	0.0232
$\alpha_{TUESDAY}$	-0.0997	0.0464	φ_{10}	0.0146	0.0232
$\alpha_{THURSDAY}$	-0.0371	0.0467	φ_{11}	0.0209	0.0232
α_{FRIDAY}	-0.1537	0.0467	φ_{12}	0.0460	0.0232
β_1	-0.0091	0.0227	φ_{13}	0.0235	0.0232
β_2	0.0345	0.0311	φ_{14}	-0.0452	0.0232
β_3	0.2279	0.0310	φ_{15}	0.0099	0.0232
φ_1	0.2780	0.0225	φ_{16}	0.0130	0.0232
φ_2	0.2556	0.0229	φ_{17}	0.0073	0.0232
φ_3	0.0438	0.0233	φ_{18}	-0.0207	0.0232
φ_4	0.0738	0.0232	φ_{19}	0.0297	0.0232
φ_5	0.0344	0.0233	φ_{20}	-0.0346	0.0232
φ_6	0.0322	0.0233	φ_{21}	0.0187	0.0232
φ_7	0.0170	0.0225	φ_{22}	0.0099	0.0227

The φ estimates are according to expectations; they are positive where significant and decrease gradually in size as the lag length becomes higher. The day-of-the-week effects are quite pronounced. Surprisingly they show the highest volatility on Wednesday and Thursday, and lower volatilities on Monday, Thursday and Friday.

4.3 Incorporating Macroeconomic News

To examine the effect of including macroeconomic announcement dates and surprises, the models are estimated for several combinations of announcements and surprises. If surprises have an influence on realized volatility, a lower SSR value is expected when these are included in the model. The difference between SSR when allowing for a leverage effect in surprises versus just including absolute surprises unleveraged is also examined.

To test the main hypothesis of leveraged surprises providing additional information beside macroeconomic announcement dates, the third, fourth and fifth model in Table 9 below need to be compared. This can be done by an ordinary F-test for joint significance of the added terms.

It is of note that adding absolute surprises beside announcement date dummies does not seem to result in a significantly better fit at the 5% significance level. However, when also adding the selected leveraged surprises, which in the case of the S&P 500 are GDP_A, TB, NAPM and HS, a significant improvement in fit is obtained. If a pronounced leverage effect in response to these macroeconomic variables exists, this is an expected result.

Table 9. Effects on SSR of adding marcoeconomic surprises – S&P500

Macro-economic news	HAR	HARST(2) $z_t = r_{t-1}$	AR(22)
None	477.12	439.36	462.10
+ FOMC	462.43	425.90	447.64
+ dummies	451.86	417.18	436.22
+ absolute surprises	448.05 (0.0949)	413.49 (0.0744)	432.77 (0.1277)
+ leveraged surprises	445.12 (0.0410)	410.94 (0.0401)	429.86 (0.0489)

Note: The p-values between parentheses given in the last 2 rows indicate the result of an F-test comparing the model with dummies to the extended model as given in the first column. For example, the p-value given in the last row is the result of a comparison of the model with an FOMC dummy and the selected dummies in 2.2.1 against the model with an FOMC dummy, selected dummies, selected absolute surprises and selected leveraged surprises.

5 Forecasting

The models are first estimated for the first 1200 observations, corresponding to the time period January 3rd, 2001 – November 15th, 2005. The one-day-ahead forecasts $\hat{y}_{t+1|t}$ are then made consecutively for the remaining 887 observations, corresponding to the period November 16th 2005 – May 23rd, 2008. Here $\hat{y}_{t+1|t}$ is the estimate for y_{t+1} given all information up to time t and where the model parameters are not re-estimated after each forecasted day but rather after each forecasted month (22 days) for calculation convenience.

5.1 Surprise Information: real-time and look-ahead

If there is a macroeconomic announcement on day t + 1, the magnitude of surprise for this announcement is obviously not known on day t. One option is to proxy $|S_{i,t+1}|$ by its expected value, which, if the surprises are assumed to be normally distributed, equals $\sqrt{2/\pi}$ because of the normalization of the generated surprises. The sign indicator functions should have expected value 0.5 as $E[S_{i,t}] = 0$.

This method would likely not provide results that vary much from the original models, as this constant prediction effect of news can be captured in the news-dummies. However, the estimation of the model using surprise magnitudes can give a more precise estimate for all variables. Specifically, it can remove a bias that occurs when the estimation period contains mostly large news surprises, while the forecast period contains mostly small surprises. Ordinarily the effect of a news announcement of that type would be overestimated whereas in this model it is not, which could lead to better forecast results.

Here, we first examine look-ahead surprises. If these add significant improvement, this indicates that surprises are of importance to model-estimation, and thus the proxied surprises can be tested for a more realistic forecasting procedure. However, as shown in the next Section, the look-ahead surprises turn out to not give significant improvements, indicating that real-time surprises are even less likely to provide an improvement. They are not examined further.

5.2 Results

5.2.1 S&P500

Table 10 and Table 11 on page 15 show the results of estimating the moving window forecast for all the discussed models for the S&P 500 index. The results are given for the case where look-ahead data is used for the surprises.

Table 10 gives several measures of forecasting performance. Firstly, in terms of MSPE, all models beat the random walk model. Surprisingly, adding surprises to the models does not give rise to big de- or increases in MSPE.

Secondly, none of the models seem heavily biased. The largest average prediction error found is 0.0106 for the HAR model with leveraged surprises. This is a small value when compared with the variance of the prediction errors and the variance in the log realized volatility series, which is 0.8295. Interesting is that all average prediction errors are positive, indicating a slight positive bias for all models.

Table 11 compares the forecasts by means of Diebold-Mariano tests. The Diebold-Mariano test is robust to covariance in the residuals and shows whether one model significantly outperforms another in terms of prediction power.

The earlier observation that all models perform better than the Random Walk model is confirmed, as seen in the first row. The HARST(2) model significantly outperforms the HAR and AR(22) models in all cases. The HAR and AR(22) model predictions show no significant differences.

Adding macroeconomic surprises to the models gives no significant improvements. In fact, all signs of the Diebold-Mariano statistics are negative, except in the case of the AR(22) model, indicating that they may actually have a negative effect on prediction power in these models.

5.2.2 Other indices and Exchange rates

The results for the other two stock indices and two exchange rates are given in the Appendix. The S&P400 and DJI forecasts confirm the results above. Here, the same conclusions hold about both the differences between models and the differences between news surprise inclusions. Note that this makes the results about macroeconomic news surprises quite robust, as other (preselected) sets of macroeconomic variables as well as different transition variables have been used for these forecasts.

The forecasts of the exchange rate realized volatilies show almost no significant differences in forecast strength between the different models. The HARST model seems less proper for forecasting in this case than in the case of stock indices. Adding surprises does not give an increase in forecast strength.

Table 10. Forecast measures for several models – S&P500 – Look-ahead surprises

Model	RW	HAR			HARST(2)			AR(22)		
News	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
$\overline{e^2}$	0.3487	0.2571	0.2574	0.2591	0.2434	0.2431	0.2448	0.2584	0.2582	0.2581
$\bar{e}$	0.0011	0.0062	0.0090	0.0106	0.0024	0.0041	0.0058	0.0073	0.0081	0.0069
$\mathrm{var}[e]$	0.3491	0.2573	0.2576	0.2593	0.2437	0.2434	0.2450	0.2586	0.2584	0.2584

Table 11. Diebold-Mariano test outcomes – S&P500 – Look-ahead surprises

	Model	RW	HAR			HARST(2)			AR(22)		
Model	*News*	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
RW	-		**7.9547** **(0.0000)**	**7.9426** **(0.0000)**	**8.1080** **(0.0000)**	**8.4554** **(0.0000)**	**8.5929** **(0.0000)**	**8.7417** **(0.0000)**	**7.0618** **(0.0000)**	**7.2106** **(0.0000)**	**7.3965** **(0.0000)**
HAR	Ann. Day Dummies			-0.3391 (0.7346)	-1.8512 (0.0641)	**4.2763** **(0.0000)**	**4.0733** **(0.0000)**	**4.2430** **(0.0000)**	-0.6625 (0.5077)	-0.4629 (0.6435)	-0.4889 (0.6250)
	Absolute Surprises				-1.7543 (0.0794)	**4.5954** **(0.0000)**	**4.7320** **(0.0000)**	**4.9720** **(0.0000)**	-0.4481 (0.6541)	-0.3758 (0.7071)	-0.3243 (0.7457)
	Leveraged Surprises					**4.6839** **(0.0000)**	**4.7409** **(0.0000)**	**5.4465** **(0.0000)**	0.2544 (0.7992)	0.3609 (0.7182)	0.4248 (0.6710)
HARST	Ann. Day Dummies						0.2921 (0.7702)	-0.9560 (0.3391)	**-5.1814** **(0.0000)**	**-6.3383** **(0.0000)**	**-4.0290** **(0.0000)**
	Absolute Surprises							-1.4602 (0.1442)	**-4.5857** **(0.0000)**	**-6.1634** **(0.0000)**	**-3.9204** **(0.0000)**
	Leveraged Surprises								**-5.1814** **(0.0000)**	**-6.3383** **(0.0000)**	**-4.0290** **(0.0000)**
AR(22)	Ann. Day Dummies									0.1551 (0.8767)	0.1787 (0.8582)
	Absolute Surprises										0.0226 (0.9820)
	Leveraged Surprises										

Note: P-values are in parentheses. A low p-value indicates a significant difference in prediction power between the two models. If the coefficient of the DM-statistic is positive, this implies that the model indicated in that column performs better than the model indicated in that row. If the statistic is negative, the opposite holds. The implied values for the empty side of the table are equal to the values given, except with opposite sign.

6 Discussion

The results in this paper show an increase in fit strength when adding leveraged macroeconomic surprises to recent models for predicting realized volatility. When used for forecasting the models do not show any improvements for three different stock indices and two exchange rates.

This discrepancy could be caused by a non-linear or time-varying effect of macroeconomic surprises. The prediction period includes the beginning of 2008, the year of the global financial crisis, which often shows different behaviour than calmer periods. It is possible that realized volatility responds differently to negative and positive news in bull or bear markets. This is also shown in Andersen et al (2007), who state that "we see that positive real economic shocks are met with a negative response in expansions and a positive response in contractions". Another possibility is that the importance to a stock index of one variable decreases over time, while that of another increases. Other variables might influence the importance of the macroeconomic surprises, leading to a surprise being very important to realized volatility one time, while it is mostly ignored another time. If for instance, a positive surprise comes after several negative surprises, this might lead to a bigger response.

The non-positive effect of adding news surprises on prediction strength is concluded only for stock indices and exchange rates. It is important not to extrapolate this result to individual stocks which might react differently. Intuitively this is feasible. The method of measurement of realized volatility is such that a positive intra-day return on one stock in the index, in conjunction with a negative intra-day return on another stock in the index, gives rise to zero increase in realized volatility if the returns happen in the same measure period. If then, the surprises have negative effects on some stocks, but positive effects on others, the total realized volatility does not increase, while the underlying volatility has increased. For exchange rates, a similar reasoning applies, as some parties benefit from the surprise, and increase the exchange rate, while others detriment from it, and decrease it.

All in all, the used surprise series play a very important role in the conclusions. If another method of surprise or expectation generation is used, this could likely lead to more significant effects. A larger sample period might lead to similar increases.

7 Conclusions

In Section 4.3 it was shown that adding present day absolute surprises to the models gave an overall better fit than leaving them out. However, the increase in fit was not significant for any of the compared models. Proceeding by accounting for a possible leverage effect in the realized volatility's reaction to macroeconomic surprises showed a significant increase in fit.

In Section 5.2 the forecast power of the different models has been compared, this lead to several conclusions. Firstly, the HARST model provided a significant improvement in forecasting power to the HAR and AR(22) model when applied to the Dow Jones Index, the S&P500 Index and the S&P400 Index. The applied Diebold-Mariano tests were inconclusive on any differences in forecasting between the AR(22) model and the HAR model. Secondly, incorporating look-ahead absolute and leveraged surprises did not show any significant improvement in forecasting power. That is, even if the surprise on day t was known beforehand, this did not lead to a significantly better prediction for day t on day t - 1.

The estimations and comparisons have been repeated for three indices and two exchange rates. This showed the results for inclusion of macroeconomic surprises to be robust to all series. The exchange rates showed different behaviour for which model gave the best forecasts, in this case the differences between the models seemed to disappear.

The hypothesis was given by "*The magnitude and sign of surprises on the announcement day of a macroeconomic variable have an influence on realized volatility beside that of the news occurring".* The hypothesis must be rejected when considering the look-ahead forecasts. In none of the five cases, for the HAR, HARST and AR(22) model, accounting for effects in the reaction to macroeconomic surprises can improve forecast quality. That is, in a total of 15 forecasts no significant improvements could be made by incorporating macroeconomic surprises.

In terms of modelling, adding macro-economic surprises, absolute and/or leveraged, seems to give a better fit. Due to this, the hypothesis might hold when considering it for a model, rather than a forecast-method. However, due to the nature of the predictions, it is advised to proceed with care when adding them, and to take the found effects lightly. It is likely that if the effects do exist, that their strength is not linear, and might depend on other variables not included in this paper.

8 References

Andersen, T. G., Bollerslev, T., Diebold, F.X. & Vega, C. (2007). Real-time price discovery in global stock, bond and foreign exchange markets. Journal of International Economics, 73, 251-277.

Balduzzi, P., Elton, E. J., & Green, T. C. (2001). Economic news and the yield curve: Evidence from the U.S. treasury market. Journal of Financial and Quantitative Analysis, 36, 523–543.

Chuliá, H., Martens, M. & Van Dijk, D. (2010). Asymmetric effects of federal funds target rate changes on S&P100 stock returns, volatilities and correlations. Journal of Banking & Finance, 34, 834–839.

Corsi, F. (2004). A simple long memory model of realized volatility. Manuscript, University of Southern Switzerland.

Engle, R.F. & Ng, V.K. (1993). Measuring and testing the impact of news on volatility. Journal of finance, 48, 1161-1191.

Gutker, S., Li, Y., Van de Werken, D. & Wissink, P.L.J. (2011). Lead-lag relation between small and large stocks and the influence of macroeconomic announcements. Paper, Erasmus University Rotterdam.

Heber, G., Lunde, A., Shephard, N. & Sheppard, K.K. (2009). Oxford-Man Institute's realized library. Version 0.2. Oxford-Man Institute. University of Oxford.

Martens, M., Van Dijk, D. & De Pooter, M. (2009). Forecasting S&P 500 volatility: Long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements. International Journal of Forecasting, 25, 282–303.

McAleer, M. & Medeiros, M.C. (2008). A multiple regime smooth transition Heterogeneous Autoregressive model for long memory and asymmetries. Journal of Econometrics, 147, 104–119.

Van Dijk, D., Teräsvirta, T. & Franses, P.H. (2002). Smooth transition autoregressive models – a survey of recent developments. Econometric Reviews, 21(1), 1–47.

9 Appendix

Table 12. Macroeconomic variables

Index	Type	Index	Type
GDP_A	GDP Advance	CS	Construction spending
GDP_P	GDP Preliminary	FO	Factory orders
GDP_F	GDP final	BI	Business inventories
NFPE	Nonfarm payroll employment	GBD	Government budget deficit
RS	Retail sales	TB	Trade balance
IP	Industrial production	PPI	Producer price index
CU	Capacity utilization	CPI	Consumer price index
PI	Personal income	CCI	Consumer confidence index
CC	Consumer credit	NAPM	NAPM index
PCE	Personal consumption expenditures	HS	Housing starts
NHSI	New home sales Investment	ILI	Index of leading indicators
DGO	Durable goods orders	IUC	Initial unemployment claims

Table 13. Test for macroeconomic announcement inclusion – S&P400 - All/Leverage Terms

News type	*p*-value	News type	*p*-value
GDP_A	0.1391	CS	0.5487
GDP_P	0.5648	FO	0.1056
GDP_F	0.9478	BI	0.9616
NFPE	0.5027	GBD	0.2787
RS	0.9317	TB	0.3848
IP	0.2757	PPI	0.3072
CU	0.5166	CPI	0.3440
PI	0.5805	CCI	**0.0935** (0.2708)
CC	0.6382	NAPM	0.4494
PCE	0.6153	HS	0.2523
NHSI	0.1995	ILI	**0.0345** (**0.0491**)
DGO	0.1674	IUC	0.1701

Table 14. Test for macroeconomic announcement inclusion – DJI - All/Leverage Terms

News type	*p*-value	News type	*p*-value
GDP_A	**0.0494** (0.1176)	CS	0.7908
GDP_P	0.7797	FO	**0.0954** (0.2185)
GDP_F	0.4720	BI	0.6291
NFPE	**0.0001** (0.5203)	GBD	0.3557
RS	0.9350	TB	0.3004
IP	0.6295	PPI	**0.0674** (0.1985)
CU	0.9120	CPI	0.2071
PI	0.6694	CCI	**0.0840** (0.4322)
CC	0.1539	NAPM	**0.0945** (0.1144)
PCE	0.6781	HS	**0.0199** (**0.0417**)
NHSI	0.2111	ILI	0.1397
DGO	**0.0043** (**0.0057**)	IUC	0.1276

Table 15. Test for macroeconomic announcement inclusion – EUR/USD - All/Leverage Terms

News type	*p*-value	News type	*p*-value
GDP_A	0.9157	CS	0.2168
GDP_P	0.9468	FO	**0.0173** (**0.0291**)
GDP_F	0.6504	BI	0.1795
NFPE	-*	GBD	0.2064
RS	0.2953	TB	0.1668
IP	0.5734	PPI	0.4183
CU	0.5242	CPI	0.6881
PI	0.1105	CCI	-*
CC	0.1587	NAPM	0.9167
PCE	0.5423	HS	0.8625
NHSI	0.1748	ILI	**0.0002** (0.7471)
DGO	0.1179	IUC	0.3015

*: All announcements for NFPE and CCI were made on days that no data is available for the EUR/USD realized volatility. Thus, they are not examined here.

Table 16. Test for macroeconomic announcement inclusion – JPY/USD - All/Leverage Terms

News type	*p*-value	News type	*p*-value
GDP_A	0.9822	CS	0.4730
GDP_P	0.9599	FO	0.3125
GDP_F	0.3198	BI	0.7843
NFPE	-*	GBD	0.5079
RS	0.8947	TB	0.2062
IP	0.1884	PPI	0.2353
CU	0.2146	CPI	0.8421
PI	**0.0217** (**0.0049**)	CCI	-*
CC	0.4929	NAPM	**0.0454** (**0.0227**)
PCE	0.9915	HS	0.9710
NHSI	0.6310	ILI	0.1810
DGO	**0.0355** (**0.0746**)	IUC	0.8123

*: All announcements for NFPE and CCI were made on days that no data is available for the JPY/USD realized volatility. Thus, they are not examined here.

Table 17. Forecast measures for several models – S&P400 – Look-ahead surprises

Model	RW	HAR			HARST(2)			AR(22)		
News	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
$\overline{e^2}$	0.3462	0.2614	0.2615	0.2627	0.2463	0.2461	0.2478	0.2624	0.2624	0.2619
$\bar{e}$	0.0020	0.0273	0.0270	0.0274	0.0353	0.0353	0.0358	0.0280	0.0280	0.0281
$\mathrm{var}[e]$	0.3466	0.2610	0.2610	0.2623	0.2454	0.2452	0.2468	0.2619	0.2619	0.2614

Table 18. Diebold-Mariano test outcomes – S&P400 – Look-ahead surprises

	Model	RW	HAR			HARST(2)			AR(22)		
Model	*News*	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
RW	-		**7.3699** **(0.0000)**	**7.5665** **(0.0000)**	**7.5227** **(0.0000)**	**8.3679** **(0.0000)**	**8.5493** **(0.0000)**	**8.4456** **(0.0000)**	**6.5357** **(0.0000)**	**6.9177** **(0.0000)**	**6.8730** **(0.0000)**
HAR	Ann. Day Dummies			-0.1421 (0.8870)	-0.8249 (0.4094)	**5.1476** **(0.0000)**	**5.2222** **(0.0000)**	**3.9930** **(0.0000)**	-0.4463 (0.6554)	-0.5529 (0.5803)	-0.2804 (0.7791)
	Absolute Surprises				-0.852 (0.3942)	**5.0353** **(0.0000)**	**5.1705** **(0.0000)**	**4.0054** **(0.0000)**	-0.3825 (0.7021)	-0.4936 (0.6216)	-0.2370 (0.8127)
	Leveraged Surprises					**4.9959** **(0.0000)**	**5.0734** **(0.0000)**	**4.9304** **(0.0000)**	0.1121 (0.9115)	0.1244 (0.9010)	0.3606 (0.7184)
HARST	Ann. Day Dummies						0.5926 (0.5535)	-0.9788 (0.3277)	**-4.8613** **(0.0000)**	**-4.8584** **(0.0000)**	**-4.5945** **(0.0000)**
	Absolute Surprises							-1.0946 (0.2737)	**-4.7727** **(0.0000)**	**-4.8706** **(0.0000)**	**-4.5837** **(0.0000)**
	Leveraged Surprises								**-3.8662** **(0.0001)**	**-3.9681** **(0.0000)**	**-3.9260** **(0.0000)**
AR(22)	Ann. Day Dummies									-0.0245 (0.9805)	0.4478 (0.6543)
	Absolute Surprises										0.7703 (0.4411)
	Leveraged Surprises										

Note: P-values are in parentheses. A low p-value indicates a significant difference in prediction power between the two models. If the coefficient of the DM-statistic is positive, this implies that the model indicated in that column performs better than the model indicated in that row. If the statistic is negative, the opposite holds. The implied values for the empty side of the table are equal to the values given, except with opposite sign. To compare the HAR and HARST model, here a HARST(2) model is estimated for S&P400 volatility; the LM-test in table 4 showed a HAR model would be sufficient, but the HARST model with 2 limiting regimes proves to be quite succesfully able to better predict realized volatility.

Table 19. Forecast measures for several models – DJI – Look-ahead surprises

Model	RW	HAR			HARST(2)			AR(22)		
News	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
$\overline{e^2}$	0.3389	0.2519	0.2525	0.2521	0.2403	0.2403	0.2403	0.2495	0.2494	0.2479
$\bar{e}$	0.0009	0.0031	0.0059	0.0060	-0.0008	0.0011	0.0013	0.0043	0.0044	0.0040
$\mathrm{var}[e]$	0.3393	0.2522	0.2527	0.2524	0.2406	0.2406	0.2405	0.2498	0.2497	0.2482

Table 20. Diebold-Mariano test outcomes – DJI – Look-ahead surprises

	Model	RW	HAR			HARST(2)			AR(22)		
Model	*News*	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
RW	-		**6.1087** **(0.0000)**	**6.0800** **(0.0000)**	**5.9224** **(0.0000)**	**6.0876** **(0.0000)**	**6.0698** **(0.0000)**	**5.9177** **(0.0000)**	**5.4528** **(0.0000)**	**5.5951** **(0.0000)**	**5.5874** **(0.0000)**
HAR	Ann. Day Dummies			-1.2448 (0.2132)	-0.1101 (0.9123)	**3.8531** **(0.0001)**	**3.6134** **(0.0003)**	**2.6802** **(0.0074)**	0.7975 (0.4252)	0.9129 (0.3613)	1.2651 (0.2058)
	Absolute Surprises				0.1256 (0.9001)	**4.0988** **(0.0000)**	**3.1987** **(0.0000)**	**2.8943** **(0.0038)**	0.9455 (0.3444)	1.0852 (0.2779)	1.3826 (0.1668)
	Leveraged Surprises					**3.3677** **(0.0008)**	**3.3386** **(0.0008)**	**3.8365** **(0.0001)**	0.5869 (0.5573)	0.6792 (0.4970)	1.3291 (0.1838)
HARST	Ann. Day Dummies						-0.0011 (0.9991)	0.0174 (0.9861)	**-3.4034** **(0.0007)**	**-4.0013** **(0.0000)**	**-3.2058** **(0.0013)**
	Absolute Surprises							0.0192 (0.9847)	**-3.1306** **(0.0017)**	**-3.6942** **(0.0002)**	**-2.9508** **(0.0032)**
	Leveraged Surprises								**-1.9949** **(0.0461)**	**-2.2342** **(0.0255)**	**-2.5234** **(0.0116)**
AR(22)	Ann. Day Dummies									0.1183 (0.9058)	0.6445 (0.5193)
	Absolute Surprises										0.7285 (0.4663)
	Leveraged Surprises										

Note: P-values are in parentheses. A low p-value indicates a significant difference in prediction power between the two models. If the coefficient of the DM-statistic is positive, this implies that the model indicated in that column performs better than the model indicated in that row. If the statistic is negative, the opposite holds. The implied values for the empty side of the table are equal to the values given, except with opposite sign.

Table 21. Forecast measures for several models – USD/EUR – Look-ahead surprises

Model	RW	HAR			HARST(2)			AR(22)		
News	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
$\overline{e^2}$	0.2011	0.1103	0.1130	0.1116	0.1085	0.1113	0.1100	0.1100	0.1123	0.1118
$\bar{e}$	-0.0004	-0.0303	-0.0316	-0.0287	-0.0338	-0.0351	-0.0336	-0.0300	-0.0305	-0.0294
$\mathrm{var}[e]$	0.2013	0.1095	0.1121	0.1109	0.1075	0.1102	0.1090	0.1092	0.1115	0.1111

Table 22. Diebold-Mariano test outcomes – USD/EUR – Look-ahead surprises

	Model	RW	HAR			HARST(2)			AR(22)		
Model	*News*	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
RW	-		**11.1890** **(0.0000)**	**9.5741** **(0.0000)**	**11.2800** **(0.0000)**	**11.8640** **(0.0000)**	**10.0720** **(0.0000)**	**11.7370** **(0.0000)**	**10.6090** **(0.0000)**	**9.1647** **(0.0000)**	**10.8720** **(0.0000)**
HAR	Ann. Day Dummies			-1.6095 (0.1075)	-1.8419 (0.0655)	1.9097 (0.0562)	-0.4830 (0.6291)	0.4395 (0.6603)	0.2591 (0.7955)	-0.8012 (0.4230)	-1.8627 (0.0625)
	Absolute Surprises				0.6350 (0.5254)	**2.5653** **(0.0103)**	1.9485 (0.0514)	1.4950 (0.1349)	2.9598 (0.0031)	0.7100 (0.4777)	0.9054 (0.3653)
	Leveraged Surprises					**2.2952** **(0.0217)**	0.1319 (0.8951)	2.3583 (0.0184)	1.0229 (0.3064)	-0.2342 (0.8148)	-0.1828 (0.8550)
HARST	Ann. Day Dummies						-1.6713 (0.0947)	-1.9082 (0.0564)	-1.3451 (0.1786)	-1.5538 (0.1202)	-3.5253 (0.0004)
	Absolute Surprises							0.5961 (0.5511)	0.9926 (0.2309)	-0.9062 (0.3648)	-0.3461 (0.7292)
	Leveraged Surprises								-0.0164 (0.9869)	-0.8423 (0.3996)	-1.9833 (0.04733)
AR(22)	Ann. Day Dummies									-1.4894 (0.1364)	**-3.021** **(0.0025)**
	Absolute Surprises										0.2493 (0.8032)
	Leveraged Surprises										

Note: P-values are in parentheses. A low p-value indicates a significant difference in prediction power between the two models. If the coefficient of the DM-statistic is positive, this implies that the model indicated in that column performs better than the model indicated in that row. If the statistic is negative, the opposite holds. The implied values for the empty side of the table are equal to the values given, except with opposite sign.

Table 23. Forecast measures for several models – USD/JPY – Look-ahead surprises

Model	RW	HAR			HARST(2)			AR(22)		
News	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
$\overline{e^2}$	0.2051	0.1450	0.1462	0.1449	0.1703	0.1500	0.1481	0.1468	0.1492	0.1470
$\bar{e}$	0.0001	-0.0040	-0.0052	-0.0043	-0.0195	-0.0052	-0.0055	-0.0033	-0.0041	-0.0034
$\mathrm{var}[e]$	0.2054	0.1451	0.1463	0.1450	0.1702	0.1502	0.1483	0.1470	0.1493	0.1471

Table 24. Diebold-Mariano test outcomes – USD/JPY – Look-ahead surprises

	Model	RW	HAR			HARST(2)			AR(22)		
Model	*News*	-	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises	Ann. Day Dummies	Absolute Surprises	Leveraged Surprises
RW	-		**5.1172** **(0.0000)**	**5.0830** **(0.0000)**	**5.4867** **(0.0000)**	1.2657 (0.2056)	**4.9716** **(0.0000)**	**5.4581** **(0.0000)**	**5.1072** **(0.0000)**	**4.9606** **(0.0000)**	**5.3777** **(0.0000)**
HAR	Ann. Day Dummies			-1.4215 (0.1552)	0.1124 (0.9105)	-1.3249 (0.1852)	-1.9635 (0.0496)	-1.0306 (0.3027)	-1.0934 (0.2742)	-1.4278 (0.1533)	-0.8460 (0.3976)
	Absolute Surprises				1.1807 (0.2377)	-1.2341 (0.2172)	-1.2481 (0.2120)	-0.5529 (0.5804)	-0.6545 (0.5128)	-1.4027 (0.1608)	-0.4792 (0.6318)
	Leveraged Surprises					-1.2651 (0.2059)	-1.6702 (0.0949)	-1.0273 (0.3043)	-1.2589 (0.2081)	-1.6623 (0.0964)	-1.3306 (0.1833)
HARST	Ann. Day Dummies						1.1080 (0.2679)	1.1666 (0.2434)	1.1815 (0.2374)	1.0226 (0.3065)	1.1288 (0.2590)
	Absolute Surprises							1.7240 (0.0847)	0.8583 (0.3907)	0.1764 (0.8600)	0.7282 (0.4665)
	Leveraged Surprises								0.3123 (0.7548)	-0.2071 (0.8359)	0.2623 (0.7931)
AR(22)	Ann. Day Dummies									-1.7222 (0.0850)	-0.1392 (0.8893)
	Absolute Surprises										1.6816 (0.0927)
	Leveraged Surprises										

Note: P-values are in parentheses. A low p-value indicates a significant difference in prediction power between the two models. If the coefficient of the DM-statistic is positive, this implies that the model indicated in that column performs better than the model indicated in that row. If the statistic is negative, the opposite holds. The implied values for the empty side of the table are equal to the values given, except with opposite sign.