Scheduled Macroeconomic News Announcements and Forex Volatility Forecasting

Tomáš Plíhal[a,*]

[a]*Institute of Financial Complex Systems, Masaryk University, Brno, Czech Republic*

Abstract

In the world of finance, the volatility of asset prices plays a crucial role, e.g., in portfolio optimization or the valuation of derivatives. Macroeconomic news announcements are among the most important factors that influence volatility in financial markets. This paper focuses on the effect of scheduled macroeconomic news announcements on the realized volatility of the most traded currency pairs, EUR/USD, GBP/USD, and USD/JPY, from 2009 to 2017. Realized volatility is analysed on a daily basis, and it is also decomposed into continuous and jump components that are analysed separately. We focus on out-of-sample forecasting and provide strong evidence that scheduled macroeconomic news announcements play a statistically significant role in volatility models. Forecasting accuracy is improved by up to 12.4%. These results are important for future practical applications in various areas of finance.

Keywords: high-frequency data, realized volatility, scheduled announcements, forecasting, Forex

[⋆]The author appreciates the support provided by the Czech Grant Agency under Grant No. 18-05829S.

[*]Corresponding author; email: t.plihal@mail.muni.cz

JEL classification: C53, G15, G17

---

1. Introduction

The volatility of asset prices plays a crucial role in finance. It is inevitable in, for example, portfolio optimization problems or the valuation of derivatives. On the other hand, volatility presents challenges for researchers because it is an elementary input variable in many models, but we are not able to measure it precisely. We usually work only with estimates of volatility, and its real value is unknown.

Volatility, in general, represents the variation in the prices of any given asset. In other words, it measures how much and how quickly prices change over time. Volatility is also often seen as uncertainty (Schwert, 1989) or a measure of risk because rapid changes in the value of assets make them less predictable. As a result, future potential returns are less certain, which makes asset and risk management more costly.

Macroeconomic news announcements are among the most important factors that influence volatility in financial markets. Market participants react to the new information and adjust their expectations and hence positions accordingly. Therefore, when some new information appears, it usually increases the number of trades, prices, and volatility.

The fact that new information affects asset prices is an essential component of modern financial theories. It is closely related to the efficient market hypothesis and the theory of rational expectations. Data containing some form of news announcements are often used to test the validity of these hypotheses in the academic literature. The development of macroeconomic news influences almost all financial assets. However, some of them are more sensitive to economic fundamentals than others. Currencies traded on foreign exchange markets are likely to be more sensitive to macroeconomic news announcements, as they are directly related to economic fundamentals and policy decisions.

In this paper, we focus on the three most traded currency pairs in the world, EUR/USD, GBP/USD, and USD/JPY. We cover a broad time frame of the nine years from 2009 to 2017. We use the specification of heterogeneous autoregressive models developed by Corsi (2009) that have increased in popularity over the last decade due to its superior performance, simple estimation process, and good interpretability of the results. There already exist various specifications of this model that address different statistical properties of realized volatility and attempt to further improve the model's accuracy. We have enriched these models by introducing a wide range of scheduled macroeconomic news announcements and used the least absolute squares shrinkage operator (LASSO) to choose the optimal model specification for out-of-sample forecasting.

We contribute to the current literature by providing strong evidence that macroeconomic news announcements are statistically significant determinants of volatility even on a daily level. Most important, these effects are useful for out-of-sample forecasting, and their inclusion improves forecasts by up to 12.4%.

The remainder of the paper is structured as follows. Section 2 contains a literature review and presents the contributions of our paper. The description of the data and methodology is provided in Sections 3 and 4, respectively.

Section 6 contains the preliminary data analysis and results from the out-of-sample forecasting procedure. Finally, Section 7 concludes and presents the possible implications of our findings.

2. Literature Review

The effect of macroeconomic reporting on financial markets has been analysed extensively from the perspective of the first or second moments. In other words, we can measure the reactions of the market in terms of the development of prices or its volatility.

Studies that focus on macroeconomic news announcements and price developments in the Forex market [1] usually focus on scheduled news. These studies incorporate news announcements into their models in the form of standardized surprises that extract unexpected information. It should be the most crucial part of the news that influences prices. A common practice is to perform an event study that considers a short time window around the news announcement. It is usually approximately one hour before and after the announcement. In other words, the studies are predominantly intraday and focus on in-sample analysis. The results suggest that macroeconomic news surprises have a statistically significant impact on foreign exchange rates. The value of the standardized surprises is unknown in advance, which decreases its usefulness for out-of-sample forecasting.

The reaction of Forex market volatility to macroeconomic news announce-

[1] e.g., Andersen et al. (2003b); Ehrmann and Fratzscher (2005); Evans and Lyons (2005); Dominguez and Panthaki (2006); Andersen et al. (2007b); Faust et al. (2007); Cai et al. (2009); Rosa (2011); Fatum et al. (2012); Kočenda and Moravcova (2016); Gau and Wu (2017)

ments has been analysed, for example, by DeGennaro and Shrieves (1997); Andersen and Bollerslev (1998b); Melvin and Yin (2000); Cai et al. (2001); Bauwens et al. (2005); Evans and Lyons (2008); Chen and Gau (2010); Lahaye et al. (2011); Hutchison and Sushko (2013); Ben Omrane and Hafner (2015).

All of the mentioned studies, except Hutchison and Sushko (2013), used high-frequency data. High-frequency data improve the precision of daily volatility estimation and the analysis of the development of volatility during the day. Most studies have focused on only one or two years of data, while Lahaye et al. (2011) examined up to 18 years of foreign exchange data along with 6.5-7 years of bond and equity futures.

When we look at the methodology applied in the papers, we notice that all of them focus on intraday and in-sample analysis. Intraday analysis attempts to answer the question of what happens with volatility during the day. Such analysis typically considers a time window around the news announcements and examines the development of volatility before and after the release of the news. This helps to separate the effect of private and public information and the speed and magnitude of the market reactions before the new information is incorporated into prices. This means that these studies do not attempt to predict future volatility with news announcements but specialize in describing the inference.

The most popular models for this type of analysis come from the family of GARCH models developed by Engle (1982). GARCH is historically probably the most famous model of volatility. However, we can find other methods, such as flexible Fourier form regression (Cai et al., 2001), generalized method of moments (GMM, Evans and Lyons, 2008), impulse response decomposition (Ben Omrane and Hafner, 2015) or simple OLS (Chen and Gau, 2010; Hutchison and Sushko, 2013). Each model is selected according to the specific needs of each research question because the authors often explain slightly different hypotheses.

All of the studies mentioned above found at least some contribution of news announcements to foreign exchange rate volatility. DeGennaro and Shrieves (1997) concluded that scheduled news announcements are more important than unscheduled news and that it causes a dramatic increase in volatility. The volatility continues to rise for the next ten minutes. Andersen and Bollerslev (1998b) also found that the effect of increased volatility lasts for only a very short time period. On the other hand, Bauwens et al. (2005) claim that volatility increases just before scheduled news announcements, while after news is released, there are no significant changes in Forex volatility.

Another important factor closely related to macroeconomic news announcements is market activity, often measured via order flows as in Cai et al. (2001), Bauwens et al. (2005), and Evans and Lyons (2008). According to Cai et al. (2001), news announcements and order flows are both important determinants of Forex volatility. However, order flows seem to be more important than the news. This statement is supported by Evans and Lyons (2008), who claimed that almost two-thirds of the effect of macroeconomic news announcements on the volatility of returns in the foreign exchange market is transmitted via order flow. Only the remaining one-third of the effect is caused directly by the news.

Our paper provides several contributions to the current academic literature. The most important contribution is that we investigate the effect of macroeconomic news announcements on a daily basis and measure performance in an out-of-sample forecasting procedure. The majority of the existing literature focuses only on inference and applies event study analysis that examines a short time window around the news announcement during the day. These studies offer valuable insight into market behaviour in relation to news announcements. However, for the further practical application of the findings, we need to measure whether the news is helpful in determining the future. This paper fills this gap and incorporates the out-of-sample forecasting approach.

Moreover, as a second contribution, we decompose realized volatility into its continuous component and jumps and analyse each component separately (similarly to Busch et al., 2011). Due to the different statistical properties of these two components (as Andersen et al., 2007a, demonstrated), we could obtain various results. It helps us to understand what part of realized volatility is the most affected by external information in the form of scheduled announcements.

Our third contribution is that we use a large scale of macroeconomic news and focus on the volatility of the Forex market, while most studies focus on the price discovery process or the stock market.

3. Data

In this analysis, we focus on the foreign exchange market. It is the market with the largest turnover in the world. We choose the top three most traded currency pairs according to the Bank of International Settlements. The latest available statistics are from April 2016 [2] when turnover on the foreign exchange market averaged approximately 5.1 trillion per day.

The most dominant currency is the US dollar, which has been on one side of almost 88% of all trades. The three most actively traded currency pairs are, as expected, USD/EUR, USD/JPY, and USD/GBP. The most liquid currency pair is USD/EUR, which accounts for approximately 23% of the total volume of transactions in the foreign exchange market. Second place belongs to USD/JPY with a share of 17.7% followed by USD/GBP with a share of 9.2% of all transactions.

We analysed nine years of data. To be more specific, our dataset ranges from 1 January 2009 to 31 December 2017. For our analysis, we require two types of data. The first part consists of high-frequency data from the foreign exchange market for selected currency pairs. We need these data for to calculate the volatility measures. The second part of the dataset consists of macroeconomic news announcements.

3.1. Realized Volatility Measures

We work with realized variance (RV), which is often used in the literature even when talking about volatility. Realized volatility is, in fact, the square root of realized variance. To follow common practice in the literature, we use high-frequency data. Due to the high persistence of volatility, high-frequency data provide a more precise measure of actual volatility, which is beneficial

[2] See https://www.bis.org/publ/rpfx16fx.pdf

for forecasting future levels of volatility Hansen and Lunde (2012).[3]

Raw high-frequency data for all currency pairs are available from DukasCopy [4] in tick-by-tick format. This is the highest frequency possible. Because such data are prone to errors caused by humans or computers, we need to apply an appropriate cleaning procedure to eliminate these inaccuracies. Our cleaning procedure is based on the work of Barndorff-Nielsen et al. (2009), who studied the effect of cleaning on realized kernels.

The foreign exchange market works continuously without breaks. However, we can observe regular changes in trading volume known as stylized facts. Dacorogna et al. (2001) and Aloud et al. (2013) provided clear evidence of intraday and intraweek seasonality. They also found that the trade volume is closely related to stock market opening hours in a given country. However, because we focus on the three currency pairs from around the world, we use the whole day from 00:00 to 23:59 GMT and only remove weekends. All release dates and times of macroeconomic news announcements are converted to the GMT timezone before we create a daily dataset.

Another question is how to calculate volatility, because there exist a large number of volatility estimators that use high-frequency data. [5] The most common practice is to measure volatility using realized variance (RV), which is defined as the sum of squared intraday returns. In the rest of the text, we will refer to this estimator simply as realized volatility or volatility. The

[3]Further statistical theory is presented, for example, in the papers by Andersen and Bollerslev (1998a); Barndorff-Nielsen and Shephard (2002); Meddahi (2002); Andersen et al. (2003a); Mykland and Zhang (2009)

[4]See https://www.dukascopy.com/swiss/english/marketwatch/historical/

[5]For a comprehensive list of estimators, see Liu et al. (2015)

exact formula is described in Equation 1

$$RV_t^N = \sum_{i=1}^{n} r_{t,i}^2 \ , \tag{1}$$

where $r_{t,i}$ represents intraday exchange rate returns and N is the number of intraday returns. The crucial part of this process is to select the appropriate N, the sampling frequency. This issue was deeply analysed by Liu et al. (2015). After considering more than 400 different estimators for 11 years of data in 31 different asset classes, they concluded that it is difficult to beat the 5-minute RV estimator in terms of estimation accuracy. Therefore, we choose a 5-minute sampling frequency and this RV estimator for our analysis because it is widely used and provides good performance in volatility estimation.

3.2. Macroeconomic News Announcements

Our data about macroeconomic news announcements come from Bloomberg. It offers thousands of various news items for each country. However, not all of these news items are suitable for our analysis. Because we want to show the effect of news announcements in an out-of-sample analysis, we choose only scheduled news. Then, we select the announcements that were released for the majority of our observed period (news for which the first observation was at least in the year 2010 and the last no sooner than in 2016). Another requirement was the presence of market expectations in the form of a Bloomberg survey. This condition selects only the most important news that grabs the most attention from market participants. Finally, all news items were manually checked to remove duplicated variables that represent the same event only in different units.

For our three currency pairs (EUR/USD, GBP/USD, JPY/USD), we need data for four countries, namely, the United States (US), euro area (EA), United Kingdom (UK), and Japan (JP). The euro area represents all countries that use the euro as their official currency. The full list of all news items along with Bloomberg tickers and the number of observations are available in Appendix A.

In Table 1, there are a number of different types of news announcements for each country sorted by sector. As Table 1 shows, we ultimately have a total of 187 different types of news announcements. The most news items, 72, are related to the United States, followed by the euro area with 57 news items. For the United Kingdom and Japan, we have only 32 and 26 observations, respectively. The exact number of observations for each news type is presented in Appendix A, along with more detailed information including the name of each news article and Bloomberg ticker.

Table 1: The Number of Different News Announcements by Sector

#	Sector	US	EA	UK	JP
1	Agriculture	10	0	0	0
2	Energy	5	0	0	0
3	Housing and Real Estate	9	0	5	2
4	Industrial Sector	6	8	2	5
5	Intl Trade & BoP	3	6	2	4
6	Labor Market	12	6	4	2
7	Monetary Sector	1	2	2	2
8	National Accounts (GDP)	3	10	7	1
9	Personal/Household Sector	3	0	1	1
10	Prices	5	7	5	4
11	Retail & Wholesale Sector	5	2	3	3
12	Surveys/Cyclical Indicators	10	16	1	2
	Total	**72**	**57**	**32**	**26**

Because we focus on out-of-sample analysis, we cannot use standardized surprises, which is common in the research literature. Instead, we represent news with simple dummy variables. The dummy variable for individual news items takes value one when there is an announcement on a given day and zero otherwise. Because we use only scheduled news, we know the value of the dummy variables for each day before it occurs.

Because some news announcements are released on the same days, the dummy variables for these news items are identical. Therefore, we merge the news with a correlation higher than 90% and retain only one dummy variable for the whole group. As a result, one dummy variable might represent a few different announcements that are released at the same time. We can see this in Appendix A, where the index in the first column sometimes stands for more than one news item.

The dummy variables contribute to the models with external information. Therefore, it could provide increased explanatory power over a model specification based only on data on historical prices. Before every news announcement, uncertainty exists, and it generally increases turmoil and volatility in markets. The dummy variables help us to detect the days when there are scheduled announcements and the danger of market surprise increases.

4. Methodology

4.1. Benchmark Volatility Models

We use two benchmark volatility models that we attempt to outperform in out-of-sample forecasting. The first model that serves as our benchmark comes from the original version of the heterogeneous autoregressive model of realized volatility (HAR-RV) developed by Corsi (2009). The model consists of three historical volatility components that should represent the short-and long-term volatility behaviour or the behaviour of individual market participants that have different investment horizons (Muller et al., 1997).

The HAR model is easy to modify. Therefore, we can add other variables that we want to analyse or control for in the model. Due to the presence of intraweek seasonality in the Forex market (e.g., see Andersen and Bollerslev (1998b); Aloud et al. (2013)), we exclude weekends from our analysis and control for day-of-the-week dependencies. To avoid the multicollinearity trap, we include dummy variables for Monday, Tuesday, Thursday, and Friday. The choice of the working day to exclude does not influence our results, so we choose Wednesday. The resulting benchmark model is specified in the following Equation 2.

$$\begin{aligned} RV_{t+1} = & \beta_1 + \beta_2 RV_t^D + \beta_3 RV_t^W + \beta_4 RV_t^M + \\ & \beta_5 (MON_{t+1} \times RV_t^D) + \beta_6 (TUE_{t+1} \times RV_t^D) + \\ & \beta_7 (THU_{t+1} \times RV_t^D) + \beta_8 (FRI_{t+1} \times RV_t^D) + \epsilon_t \end{aligned} \tag{2}$$

We label this model simply as HAR. All dummy variables are included in the model specification as interactive regressors (we multiply them by lagged daily volatility) because we want to control for the actual volatility level. This type of model specification allows us to observe whether macroeconomic news announcements contain additional information for realized volatility forecasting or if we are able to capture the most crucial part of volatility in simple day-of-the-week dummies.

The second type of HAR model decomposes realized volatility into continuous (CC) and jump (JC) components. According to Andersen et al.

(2007a), continuous and jump components have different statistical properties, and as a result, the effect on the future levels of volatility might vary. Another reason for this decomposition is the fact that realized volatility is a non-consistent estimator of price variation when jumps (discontinuous price changes) are present.

Several approaches exist to detect volatility due to jumps. We follow the procedure suggested by Andersen et al. (2012) that uses median realized volatility (MRV). Median realized volatility is a jump-robust estimator of integrated volatility and is defined as

$$MRV_{t,f} = \frac{\pi}{6 - 4\sqrt{3} + \pi} \left(\frac{f}{f-2}\right) \sum_{j=2}^{f-1} \left[med|r_{t,j-1}|, |r_{t,j}|, |r_{t,j+1}|\right]^2 \ . \tag{3}$$

For the estimation of continuous and jump components, we also need to define the median realized quarticity MRQ (Equation 4). Subsequently, we are able to calculate the test statistic JT (Equation 5) related to the jump component.

$$MRQ_{t,f} = \frac{3\pi f}{9\pi + 72 - 52\sqrt{3}} \left(\frac{f}{f-2}\right) \sum_{j=2}^{f-1} \left[med|r_{t,j-1}|, |r_{t,j}|, |r_{t,j+1}|\right]^4 \tag{4}$$

$$JT_t = \sqrt{f} \frac{(RV_t - MRV_r)RV_t^{-1}}{(0.96 max\{1, MRQ_t/RV_t^2\})^{1/2}} \tag{5}$$

The last step is to calculate continuous and jump components defined as

$$JC_t = max\left\{0, (RV_t - MRV_t)I\left[JT_t > \phi_{1-\alpha}\right]\right\} \tag{6}$$

$$CC_t = MRV_t I\left[JT_t > \phi_{1-\alpha}\right] + RVI\left[JT_t \leq \phi_{1-\alpha}\right] \ . \tag{7}$$

$I[.]$ denotes an indicator function that takes value 1 if the test statistic JT_t exceeds the 1 - α quantile of the standard normal distribution ($\phi_{1-\alpha}$), where α is set to 0.05 (see Eq. 6 in Andersen et al., 2012, for further details). Our HAR-CJ model also contains day-of-the-week variables as in the previous HAR model and is specified as follows:

$$\begin{aligned} RV_{t+1} =& \beta_1 + \beta_2 CC_t^D + \beta_3 JC_t^D + \beta_4 CC_t^W + \beta_5 JC_t^W + \beta_6 CC_t^M + \beta_7 JC_t^M + \\ & \beta_8 (MON_{t+1} \times RV_t^D) + \beta_9 (TUE_{t+1} \times RV_t^D) + \\ & \beta_{10} (THU_{t+1} \times RV_t^D) + \beta_{11} (FRI_{t+1} \times RV_t^D) + \epsilon_t \end{aligned} \tag{8}$$

Andersen et al. (2007a) proved that CC and JC have different statistical properties. Therefore, it could be valuable to forecast these parts of realized volatility separately. Our HAR models could be easily modified to achieve this goal. A similar approach was used, for example, by Busch et al. (2011).

We use the model specifications as in the case of HAR and HAR-CJ (Equations 2 and 8) and only replace the dependent variable RV_{t+1} with CJ_{t+1} for forecasting continuous components or JC_{t+1} for jump component forecasts. All other regressors remain the same. These model specifications allow us to measure the incremental information in macroeconomic news announcements separately for continuous and jump components. This approach could provide important insights into the information content of news announcements for volatility forecasting.

4.2. Least Absolute Shrinkage and Selection Operator (LASSO)

We enrich our benchmark models by including macroeconomic news announcement dummy variables. By adding these regressors, our models contain up to 96 variables. We do not face a dimensionality issue because the number of observations (1000-day window) is still larger than the number of regressors. However, with a large number of variables in a model, it becomes difficult to interpret the results, and out-of-sample performance is usually negatively affected in models that are too complex. Moreover, including too many variables in a regression model could lead to overspecification bias.

We address this issue by the least absolute shrinkage and selection operator (LASSO) introduced by Tibshirani (1996) that helps us to select only the most useful variables for explaining volatility. This machine learning algorithm is well supported by theory (e.g., see Hastie et al., 2015) and provides variable selection and regularization to improve the predictive ability and interpretability of the model. Another advantage of LASSO is that the algorithm is computationally cheap even with a large number of variables involved. Moreover, LASSO simultaneously provides model estimates and model selection, which saves resources required for computation (Friedman et al., 2010).

LASSO uses the L1 regularization technique, which adds the absolute value of a magnitude of coefficients as a penalty term to the loss function, as we can see in Equation 9. It helps to avoid overfitting by shrinking the less important features' coefficients to zero. When a coefficient is zero, it basically removes the regressor from the model. λ determines the size of the penalty. If λ is equal to zero, the resulting model is a standard OLS regression. On the other hand, a very large value makes most coefficients zero, and the model will underfit.

$$\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda\sum_{j=1}^{p}|\beta_j| \tag{9}$$

However, LASSO suffers from some drawbacks. For example, Zou and Hastie (2005) demonstrates that in the case of highly correlated predictors, LASSO selects one at random. This behaviour could alter the predictive accuracy of the model and its interpretation. Another criticism comes from Castle et al. (2011). They claim that LASSO coefficients are biased and that LASSO could exclude some important variables with negative dependencies because its choice depends on the order of inclusion in a model.[6]

The result of LASSO depends on the selected λ. The optimal λ is usually selected via cross-validation. However, the standard cross-validation method selects validation data randomly from the whole dataset. Due to the statistical nature of the time series and the frequent presence of autocorrelation, nearby observations are dependent. As a result, there is a danger of overfitting when the estimated errors from random cross-validation will be excessively optimistic.[7]

The models that we use in this study have an autoregressive nature. All of them contain a lagged dependent variable of some form. To eliminate

[6]Naturally, there exist some other statistical techniques that help to handle overfitting and provide variable selection such as cross-validation and stepwise regression. However, these methods perform well only with a small number of variables. We can also mention the Elastic Net (Zou and Hastie, 2005), Group LASSO (Yuan and Lin, 2006), Fused LASSO (Tibshirani et al., 2005), Adaptive LASSO (Zou, 2006), or general-to-specific model selection method (Castle et al., 2011).

[7]For more details see, e.g., Burman et al. (1994); Bergmeir and Benítez (2012); Bergmeir et al. (2018)

some further possible issues with time series, we adopt suggestions from Bergmeir and Benítez (2012) and apply 10-fold cross-validation on the blocks of data. In other words, instead of selecting the data for the validation part at random, we create sequential blocks. In this way, the issue with dependent observations is minimized for the majority of our data.

In all our models, we are working with logarithms of RV, CJ, and JC. Because the jump component is sometimes zero, which prevents the use of simple logarithmic transformation, JC is calculated as $log(1 + JC)$. This is a standard procedure in the academic literature (e.g., see Corsi and Renò, 2012). We use a rolling window of 1000 observations. We estimate the model on the selected 1000 observations, create a one-day-ahead forecast and move the window one observation ahead. Then, the model is re-estimated again. In this way, we end up with 1319 forecasts.

5. Forecasting Performance Comparison

The most common approach for forecast comparison is to use statistical loss functions. The most widely used loss functions in the academic literature related to volatility forecasting are the mean squared error (MSE) and the QLIKE. Patton (2011) evaluated different statistical loss functions and examine whether they are robust to the presence of noise in the volatility proxies. He confirmed that the MSE and the QLIKE give a consistent ranking of forecasts even if the proxy of the underlying latent volatility is measured with noise.

The exact specifications of both loss functions are in the following Equations 10 and 11.

$$MSE_i = \frac{1}{n}\sum_{i=1}^{n}(\widehat{RV}_{i,t} - RV_t)^2 \tag{10}$$

$$QLIKE_i = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{RV_t}{\widehat{RV}_{i,t}} - \ln\frac{RV_t}{\widehat{RV}_{i,t}} - 1\right), \tag{11}$$

where $\widehat{RV}_{i,t}$ is a vector of predictions and RV_t represents a vector of observed values. The mean squared error (MSE) penalized the forecast error in both directions with the same weight. In other words, forecasting higher volatility has the same negative effect on the results as forecasting lower volatility. Extreme forecasts are also increasingly penalized due to the construction of MSE. On the other hand, the QLIKE loss function is less sensitive to extremes and penalizes the underestimation of volatility, which is usually worse in practical applications than overestimation.

For formal comparison of individual models, we use the model confidence set (MCS) suggested by Hansen et al. (2011). The input values for the MCS are daily loss functions (in our case, MSE and QLIKE). The MCS test is formally defined as follows. Equation 12 shows the loss function L calculated for the i^{th} model on day t. $\widehat{RV}_{i,t}$ represents our forecast, and RV_t is the observed value.

$$L(i,t) = L(\widehat{RV}_{i,t}, RV_t) \tag{12}$$

To compare two competing models i and j, we have to calculate a loss differential $D(i,j,t)$ (Equation 13 and the average loss of model i at time t with respect to the remaining models represented by $D(i,.,t)$ (Equation 14 as follows:

$$D(i,j,t) = L(i,t) - L(j,t) \tag{13}$$

$$D(i,.,t) = \frac{1}{m-1} \sum_{j,j \neq i} D(i,j,t) \tag{14}$$

Then, we test the hypothesis defined in Equations 15 and 16. The null hypothesis stands for the equal predictive ability of the models.

$$H_0 : E[D(i,.)] = 0, \text{ for all}, \quad i = 1, ..., m \tag{15}$$

$$H_1 : E[D(i,.)] \neq 0, \text{ for some}, \quad i = 1, ..., m \tag{16}$$

The test statistic T is defined as follows:

$$T_{max} = max_{(i)} t_{i,.} \tag{17}$$

$$t_{i,.} = \frac{\bar{d}_{i,.}}{\sqrt{v\bar{a}r\left(\bar{d}_{i,.}\right)}} \tag{18}$$

where $\bar{d}_{i,.}$ is an average of the average loss differential of the i^{th} model with respect to other competing models and the denominator is a block bootstrap estimator of the variance of $\bar{d}_{i,.}$. The distribution under the null hypothesis of the test statistics T_{max} is bootstrapped as the variance before. If the null hypothesis is rejected, the model with the highest $t_{i,.}$ is eliminated, and the testing procedure is repeated until the null is not rejected (we set $\alpha = 0.1$).

6. Results

6.1. Preliminary Data Analysis

First, we examine the elementary statistical properties of our dependent variables for individual countries. The results are presented in Table 2. All variables are shown after the logarithmic transformation in a form that enters our models. We can spot similar trends in all analysed currency pairs.

Realized volatility and its long-memory properties are clearly noticeable. Even at lag 100, the values of the autocorrelation coefficients are 0.33 for EUR/USD, 0.4 for GBP/USD, and 0.15 in the case of USD/JPY. These values indicate the high persistence of the data generating process. Even the lowest value for USD/JPY is considerable. The mean values for volatility are almost the same for all currencies and centre on a value of 4.25. The probability distributions are leptokurtic in all cases, which indicates the presence of 'fat tails.' All of these findings are in line with the stylized facts about realized volatility.

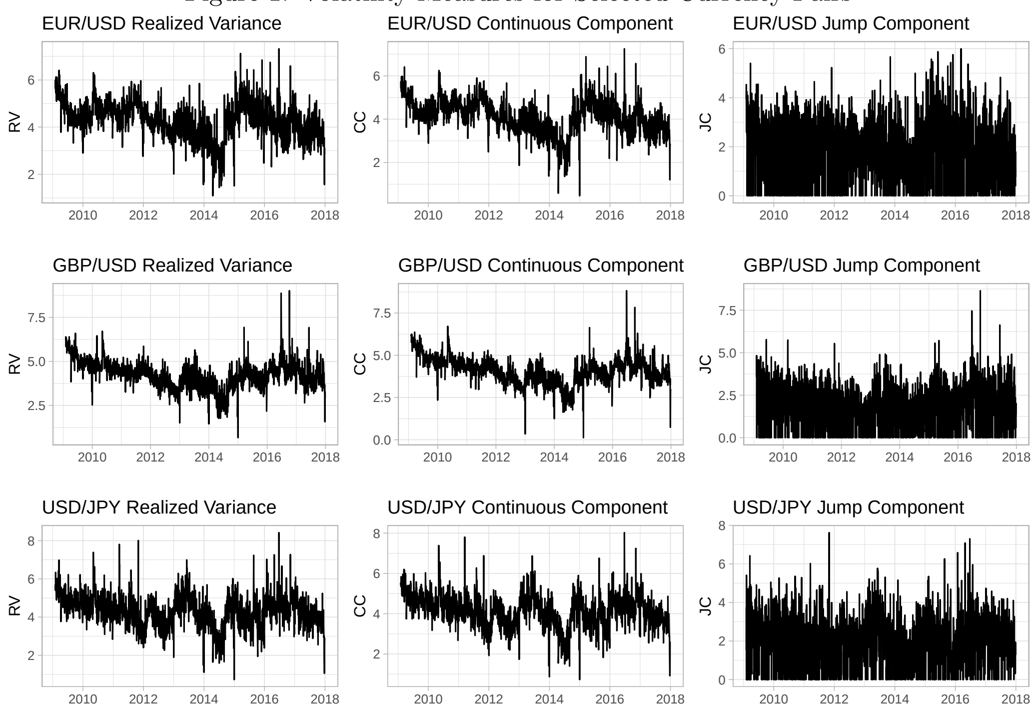
Table 2: Descriptive Statistics of Volatility Measures

		Mean	SD	Skew.	Kurt.	Max.	Med.	Min.	AR1	AR22	AR100	AR250
	RV	4.26	0.76	-0.24	0.63	7.32	4.31	1.09	0.72	0.48	0.33	0.08
EUR/USD	CC	4.11	0.77	-0.30	0.75	7.25	4.15	0.47	0.77	0.53	0.36	0.10
	JC	1.92	1.21	0.01	-0.56	5.99	2.01	0.00	0.05	0.07	0.08	0.00
	RV	4.23	0.77	0.15	1.85	9.01	4.24	0.67	0.76	0.53	0.37	0.20
GBP/USD	CC	4.08	0.77	0.02	1.73	8.82	4.08	0.12	0.80	0.58	0.40	0.21
	JC	1.90	1.23	0.15	-0.05	8.65	1.98	0.00	0.13	0.06	0.05	0.03
	RV	4.27	0.83	0.05	1.14	8.42	4.29	0.73	0.66	0.36	0.15	-0.03
USD/JPY	CC	4.11	0.81	-0.04	1.01	8.03	4.13	0.73	0.72	0.39	0.15	-0.02
	JC	2.04	1.28	0.17	-0.10	7.62	2.11	0.00	0.14	0.05	0.02	-0.01

The continuous and jump components for all countries confirm the findings of Andersen et al. (2007a), who claim that these two components have different statistical properties. This distinction could cause different degrees of predictive power for realized volatility forecasting.

The continuous components exhibit similar characteristics as realized volatility, but they are more persistent when we check the autocorrelation coefficients, which is intuitively clear because one removes the jump component that, on the other hand, does not exhibit a long-memory or strong short-memory property. The mean jump component value is approximately half of the CC with a higher standard deviation (SD) in all cases. Additionally, note that jumps are sometimes zero.

Figure 1: Volatility Measures for Selected Currency Pairs



In Figure 6.1, we can see how each of these components (RV, CC, JC) developed during the observed period from 2009 to 2017 for each currency pair. Individual currency pairs evolved in the same manner. We can also see the long memory of realized variance and continuous components, while jump components seem to be very noisy with low autocorrelation.

6.2. Out-of-Sample Analysis

We forecast one day ahead. The forecasts are based on the rolling window of one thousand observations, and the model is re-estimated after every observation. We use LASSO to calibrate the model with 10-block cross-validation. We measure model performance according to the loss functions MSE and QLIKE. Moreover, we perform a model confidence test to statistically prove which model is best at a 90% confidence level. The best models are denoted by the ‡ sign.

The results are shown in Table 3. Our benchmark is the HAR model with the day-of-the-week regressors. The HAR-CJ model decomposes RV regressors into CC and JC components. The other models, N-HAR and N-HAR-CJ, incorporate macroeconomic news variables and allow LASSO to choose the best model via 10-block cross-validation. For forecasting the continuous component (CC) and jump component (JC), we use only the HAR-CJ model and N-HAR-CJ model. The specifications are the same as in the case of realized volatility, except that the dependent variable is replaced by CC or JC. The ‡ sign in a table indicates the model that was selected for inclusion in the superior set of models according to the model confidence set (MCS).

When we look at the overall results, realized volatility forecasting is always selected as the best model N-HAR-CJ. This model provides improvements of up to 12.398% in comparison to the benchmark HAR for EUR/USD. The other two currency pairs confirm these findings, but the improvements are lower and centre on approximately 9% for the GBP/USD and approximately 8% for the USD/JPY currency pair.

For the continuous component, the improvements are not as substantial as in the case of realized volatility. The best results are again provided by the N-HAR-CJ model for all currency pairs. The greatest improvement occurs in the case of EUR/USD and is up to 8.107%. The accuracy of the GBP/USD and USD/JPY forecasts improved by approximately 4% and 3%, respectively.

The analysis of the jump component provides the least contribution to forecasting performance. The N-HAR-CJ model is always included in the superior set of models. However, it is often not better than our benchmark model that is also chosen as the best at the 10% significance level. Forecasting accuracy improvements are only approximately 2%.

In all cases, for all currency pairs, our benchmark models that contain the day-of-the-week variables (HAR and HAR-CJ) are always outperformed by the models that incorporate news variables and LASSO for estimation. Therefore, the day-of-the-week variables are not sufficient, and news announcements provide additional information for forecasting realized volatility and continuous components. However, the results are not statistically significant for the forecasting of the jump component.

Table 3: Out-of-sample Loss Functions for Individual News Dummy Variables

		EUR/USD						**GBP/USD**						**USD/JPY**					
		MSE			QLIKE			MSE			QLIKE			MSE			QLIKE		
RV	HAR	27.777			0.979			25.119			0.854			37.227			1.146		
	N-HAR LASSO	24.879	-10.432	‡	0.883	-9.779		23.785	-5.310	‡	0.803	-5.944		32.250	-13.370	‡	1.078	-5.907	
	N-HAR BMA	24.287	-10.762	‡	0.966	-1.314		23.926	-4.749	‡	0.829	-2.967		32.658	-12.273		1.123	-1.981	
	MCS p-value	0.547			0.02			0.360			0.02			0.103			0.01		
CC	HAR-CJ	23.257			0.995			20.416			0.995			30.448		‡	1.090		
	N-HAR-CJ LASSO	21.372	-8.107	‡	0.925	-7.049	‡	19.576	-4.112	‡	0.957	-3.762	‡	30.368	-0.263	‡	1.053	-3.412	‡
	N-HAR-CJ BMA	23.848	2.540		0.966	-1.314		33.547	64.320		0.829	-2.967		92.578	204.051		1.123	-1.981	
	MCS p-value	0			0			0.005			0			0.665			0		
JC	HAR-CJ	140.106			10.467		‡	138.984			10.327			535.801		‡	9.339		
	N-HAR-CJ LASSO	136.322	-2.701	‡	10.277	-1.819	‡	136.885	-1.511	‡	10.167	-1.552	‡	532.012	-0.707	‡	9.181	-2.327	‡
	N-HAR-CJ BMA	169.475	20.962		0.966	-1.314		164.681	18.489		0.829	-2.967		449.364	-16.132	‡	1.123	-1.981	
	MCS p-value	0.016			0.31			0.102			0			0.352			0.01		

‡ denotes models selected to the superior set of models at the level of $\alpha = 0.1$
All values of loss functions are multiplied by 100.

6.3. The Variables that Make the Greatest Contribution

We also want to distinguish the macroeconomic variables that are the most beneficial for the model. Because we re-estimate the model every day using LASSO, the regressors that enter the best models could vary greatly, as we have a new model every day. To find the best variables, we note the selected model for each day and calculate the ratio of how often each regressor is included in the model. We present only simplified results and show the top-five variables in terms of contribution. The results are shown in Appendix B in Tables 8, 9, and 10. The first column (Sel.%) shows how often a given variable occurs in a model in percentage terms. The second and third columns present the full names of the regressors and related sectors, respectively.

When we look at the results, it is obvious that news from the United States is selected more often, which supports the hypothesis of Ehrmann and Fratzscher (2005); Ben Omrane and Hafner (2015) that U.S. news is more important than announcements from other countries. In the case of the EUR/USD in the first few places, there is only one announcement from the euro area and is related to the interest rate policy of the European Central Bank. For the GBP/USD, the situation is more balanced at first glance. However, the USD/JPY simplified table does not contain any news announcements from Japan, which contradicts Fatum et al. (2012), who claim that US and Japanese news have the same importance. To generalize the results, we can conclude that the United States announcements play a more important role in realized volatility and its continuous and jump components.

Among the most important macroeconomic news announcements according to the academic literature belong to the nonfarm payrolls, federal funds rate, GDP, and trade balance (e.g., Lahaye et al., 2011; Caporin and Poli, 2017). Information about the change in nonfarm payrolls is released on the same days as other news related to the labor market, such as change in Private Payrolls and Change in Manufact. Payrolls, Average Hourly Earnings MoM, Average Weekly Hours All Employees, and Unemployment Rate. In our daily analysis, we are not able to distinguish which of these news items is the most important. However, we can confirm that news related to the labor market, including nonfarm payrolls, is a significant determinant of exchange rate volatility and its components. The strongest results occur for USD/JPY, where these news items are always included in all models. The high importance of this information is also present in the EUR/USD currency pair. For GBP/USD, the results are not as strong, and these announcements from the U.S. labor market are strongest only for the HAR-CJ model of realized volatility. However, we observe a substantial contribution of labor market news from the United Kingdom, specifically Weekly Earnings ex Bonus 3M/YoY, Average Weekly Earnings 3M/YoY, ILO Unemployment Rate 3Mths, and Jobless Claims Change. These UK news items are also released during the same days and are nearly always present in all models. Therefore, it is obvious that the labor market, in general, is an important determinant of exchange rate volatility for the examined currency pairs. We can also confirm the importance of the federal funds rate, which often belongs among top news announcements. However, the information about GDP or trade balance does not seem to be the most influential news releases.

Regarding the other news often present in our simplified results, we highlight, for example, the energy sector from the United States that is often present among the best news and is represented by announcements about DOE U.S. Refinery Utilization, DOE U.S. Crude Oil Inventories, DOE U.S. Distillate Inventory, and DOE U.S. Gasoline Inventories.

We can also observe the short memory of the jump component. The values in the first column are lower than in the case of realized volatility or the continuous component, which indicates that the selected models in each step are not very stable and the regressors often change.

7. Conclusion

This paper focuses on measuring the effect of scheduled macroeconomic news announcements on realized volatility in the foreign exchange market. We choose three of the most traded currency pairs in the world, EUR/USD, GBP/USD, and USD/JPY. Daily realized volatility for these currency pairs is calculated from the high-frequency data from Ducascopy. The scheduled news comes from Bloomberg, and we obtain 72, 57, 32, and 26 different types of announcements from the United States, the euro area, the United Kingdom, and Japan, respectively. The analysed time period ranges from 2009 to 2017.

The primary goal was to analyse the effect of scheduled macroeconomic news announcements on foreign exchange market volatility in an out-of-sample setting. We wanted to measure the contribution of news announcements, identify the most important news, and separately examine the continuous and jump components of realized volatility to observe the possible differences.

As a benchmark, we use the heterogeneous autoregressive model developed by Corsi (2009) and its modification that decomposes realized volatility into continuous and jump components from Andersen et al. (2007a). We control for the day-of-the-week seasonality and the level of volatility. Because of the large number of macroeconomic news announcements in our analysis, we use the power of the least squares shrinkage operator that selects the best model and provides a variable selection. It shows us the most frequently selected variables and the best model according to 10-block cross-validation that minimizes the MSE loss function.

With the use of the model confidence set test, we found significant improvements in forecast accuracy. These findings are stable across all examined currency pairs. The lowest increases were observed when forecasting the jump components. On the other hand, realized volatility improves by up to 12.4% in the case of EUR/USD.

All our models beat the benchmark HAR model. The macroeconomic news announcement variables contain additional valuable information for out-of-sample forecasting because the models that include news announcement regressors provided the most accurate results and were often preferred by the MCS test as the best model.

We also show the macroeconomic news announcements that were selected for the forecasting model most of the time. The results confirm the dominance of the news from the United States, especially for the USD/JPY currency pair. The news from the U.S. labor market (Change in Private Payrolls, Change in Manufact. Payrolls, Change in Non-farm Payrolls, Average Hourly Earnings MoM, Average Weekly Hours All Employees, Unemployment Rate) and energy sector (DOE U.S. Refinery Utilization, DOE U.S. Crude Oil Inventories, DOE U.S. Distillate Inventory, DOE U.S. Gasoline Inventories) has demonstrable importance for realized volatility as well as continuous and jump component forecasting. Additionally, news related to monetary policy in all countries is often present among the best news variables.

In this paper, we provide strong evidence that scheduled macroeconomic news announcements play an important role in volatility forecasting on a daily basis. Announcement dummy variables make statistically significant contributions to the models that are not captured by simple day-of-the-week dummies. These results contribute to the research literature by focusing on the out-of-sample performance of news announcements on a daily level, which is useful for practical applications in finance.

Data Availability Statement

Raw high-frequency data for analyzed currency pairs are openly available from DukasCopy (See https://www.dukascopy.com/swiss/english/marketwatch/historical/) in tick-by-tick format. Data about macroeconomic news announcements come from Bloomberg (tickers are provided in Appendix A in Tables 4, 5, 6, 7).

References

References

Aloud, M., Fasli, M., Tsang, E., Dupuis, R., Olsen, R., 2013. Stylized facts of trading activity in the high frequency fx market: An empirical study. Journal of Finance and Investment Analysis 2, 145–183.

Andersen, T.G., Bollerslev, T., 1998a. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review , 885.

Andersen, T.G., Bollerslev, T., 1998b. Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. The Journal of Finance , 219.

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007a. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. The Review of Economics and Statistics , 701.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003a. Modeling and forecasting realized volatility. Econometrica , 579.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Vega, C., 2003b. Micro effects of macro announcements: Real-time price discovery in foreign exchange. The American Economic Review , 38.

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

Andersen, T.G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. Journal of Econometrics 169, 75 – 93.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2009. Realized kernels in practice: trades and quotes. Econometrics Journal 12.

Barndorff-Nielsen, O.E., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society. Series B (Statistical Methodology) , 253.

Bauwens, L., Ben Omrane, W., Giot, P., 2005. News announcements, market activity and volatility in the euro/dollar foreign exchange market. Journal of International Money and Finance 24, 1108 – 1125.

Ben Omrane, W., Hafner, C., 2015. Macroeconomic news surprises and volatility spillover in foreign exchange markets. Empirical Economics 48, 577.

Bergmeir, C., Benítez, J.M., 2012. On the use of cross-validation for time series predictor evaluation. Information Sciences 191, 192 – 213.

Bergmeir, C., Hyndman, R.J., Koo, B., 2018. A note on the validity of cross-validation for evaluating autoregressive time series prediction. Computational Statistics and Data Analysis 120, 70–83.

Burman, P., Chow, E., Nolan, D., 1994. A cross-validatory method for dependent data. Biometrika 81, 351.

Busch, T., Christensen, B.J., Nielsen, M.Ø., 2011. The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. Journal of Econometrics 160, 48 – 57. URL: http://www.sciencedirect.com/science/article/pii/S0304407610000564.

Cai, F., Joo, H., Zhang, Z., 2009. The impact of macroeconomic announcements on real time foreign exchange rates in emerging markets. International Finance Discussion Papers 973. Board of Governors of the Federal Reserve System (U.S.).

Cai, J., Cheung, Y.L., Lee, R.S., Melvin, M., 2001. ‘once-in-a-generation’ yen volatility in 1998: fundamentals, intervention, and order flow. Journal of International Money and Finance 20, 327 – 347. doi:https://doi.org/10.1016/S0261-5606(00)00056-5.

Caporin, M., Poli, F., 2017. Building news measures from textual data and an application to volatility forecasting. Econometrics 5, 1.

Castle, J.L., Doornik, J.A., Hendry, D.F., 2011. Evaluating automatic model selection. Journal of time series econometrics 3, 1 – 31.

Chen, Y.L., Gau, Y.F., 2010. News announcements and price discovery in foreign exchange spot and futures markets. Journal of Banking & Finance 34, 1628 – 1636.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7, 174 – 196.

Corsi, F., Renò, R., 2012. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. Journal of Business & Economic Statistics , 368.

Dacorogna, M., Müller, U., Pictet, O., de Vries, C., 2001. Extremal forex returns in extremely large data sets. Extremes 4, 105.

DeGennaro, R.P., Shrieves, R.E., 1997. Public information releases, private information arrival and volatility in the foreign exchange market. Journal of Empirical Finance 4, 295 – 315. doi:https://doi.org/10.1016/S0927-5398(97)00012-1.

Dominguez, K.M., Panthaki, F., 2006. What defines 'news' in foreign exchange markets? Journal of International Money and Finance 25, 168 – 198.

Ehrmann, M., Fratzscher, M., 2005. Exchange rates and fundamentals: New evidence from real-time data. Journal of International Money and Finance 24, 317–341.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica , 987.

Evans, M.D., Lyons, R.K., 2005. Do currency markets absorb news quickly? Journal of International Money and Finance 24, 197 – 217.

Evans, M.D., Lyons, R.K., 2008. How is macro news transmitted to exchange rates? Journal of Financial Economics 88, 26 – 50.

Fatum, R., Hutchison, M., Wu, T., 2012. Asymmetries and state dependence: The impact of macro surprises on intraday exchange rates. Journal of The Japanese and International Economies 26, 542 – 560.

Faust, J., Rogers, J.H., Wang, S.Y.B., Wright, J.H., 2007. The high-frequency response of exchange rates and interest rates to macroeconomic announcements. Journal of Monetary Economics 54, 1051 – 1068.

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software 33.

Gau, Y.F., Wu, Z.X., 2017. Macroeconomic announcements and price discovery in the foreign exchange market. Journal of International Money and Finance 79, 232 – 254.

Hansen, P.R., Lunde, A., 2012. Forecasting Volatility Using High-Frequency Data. Oxford University Press, Stanford University.

Hansen, P.R., Lunde, A., Nason, J.M., 2011. The model confidence set. Econometrica 79, 453–497.

Hastie, T., Tibshirani, R., Wainwright, M., 2015. Statistical learning with sparsity. Monographs on Statistics and Applied Probability, 143., CRC Press, Boca Raton, FL.

Hutchison, M., Sushko, V., 2013. Impact of macro-economic surprises on carry trade activity. Journal of Banking & Finance 37, 1133 – 1147.

Kočenda, E., Moravcova, M., 2016. Intraday Effect of News on Emerging European Forex Markets: An Event Study Analysis. Working Papers IES 2016/20. Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies.

Lahaye, J., Laurent, S., Neely, C.J., 2011. Jumps, cojumps and macro announcements. Journal of Applied Econometrics 26, 893 – 921.

Liu, L.Y., Patton, A.J., Sheppard, K., 2015. Does anything beat 5-minute rv?

a comparison of realized measures across multiple asset classes. Journal of Econometrics 187, 293 – 311.

Meddahi, N., 2002. A theoretical comparison between integrated and realized volatility. Journal of Applied Econometrics 17, 479–508. doi:10.1002/jae.689.

Melvin, M., Yin, X...., 2000. Public information arrival, exchange rate volatility, and quote frequency. Economic Journal 110, 644–661.

Muller, U.A., Dacorogna, M.M., Dave, R.D., Olsen, R.B., Pictet, O.V., von Weizsacker, J.E., 1997. Volatilities of different time resolutions — analyzing the dynamics of market components. Journal of Empirical Finance 4, 213 – 239. High Frequency Data in Finance, Part 1.

Mykland, P.A., Zhang, L., 2009. Inference for continuous semimartingales observed at high frequency. Econometrica , 1403.

Patton, A.J., 2011. Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics 160, 246 – 256.

Rosa, C., 2011. The high-frequency response of exchange rates to monetary policy actions and statements. Journal of Banking and Finance 35, 478 – 489.

Schwert, G.W., 1989. Why does stock market volatility change over time?. The Journal of Finance , 1115.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) , 267.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., Knight, K., 2005. Sparsity and smoothness via the fused lasso. Journal of the Royal Statistical Society. Series B (Statistical Methodology) , 91.

Yuan, M., Lin, Y., 2006. Model selection and estimation in regression with grouped variables. Journal of the Royal Statistical Society. Series B (Statistical Methodology) , 49.

Zou, H., 2006. The adaptive lasso and its oracle properties. Journal of the American Statistical Association , 1418.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society. Series B (Statistical Methodology) , 301.

Appendix A

Table 4: List of US News Announcements

	Event	Ticker	Sector	No. Obs.
1	ADP Employment Change	ADP CHNG Index	Labor Market	108
2	Acreage Corn Planted	ACRECRNP Index	Agriculture	9
	Acreage Soybean Planted	ACRESOYP Index	Agriculture	9
3	Average Hourly Earnings MoM	AHE MOM% Index	Labor Market	95
	Average Weekly Hours All Employees	AWH TOTL Index	Labor Market	95
	Change in Manufact. Payrolls	USMMMNCH Index	Labor Market	108
	Change in Nonfarm Payrolls	NFP TCH Index	Labor Market	108
	Change in Private Payrolls	NFP PCH Index	Labor Market	93
	Unemployment Rate	USURTOT Index	Labor Market	108
4	Building Permits	NHSPATOT Index	Housing and Real Estate	107
	Housing Starts	NHSPSTOT Index	Housing and Real Estate	106
5	Business Inventories	MTIBCHNG Index	Retail & Wholesale Sector	108
	Retail Sales Advance MoM	RSTAMOM Index	Retail & Wholesale Sector	108
	Retail Sales Ex Auto MoM	RSTAXMOM Index	Retail & Wholesale Sector	108
6	CPI Ex Food and Energy MoM	CPUPXCHG Index	Prices	108
	CPI MoM	CPI CHNG Index	Prices	108
7	Cap Goods Orders Nondef Ex Air	CGNOXAI% Index	Industrial Sector	87
8	Capacity Utilization	CPTICHNG Index	Industrial Sector	108
	Industrial Production MoM	IP CHNG Index	Industrial Sector	108
9	Chicago Purchasing Manager	CHPMINDX Index	Surveys/Cyclical Indicators	108
10	Conf. Board Consumer Confidence	CONCCONF Index	Surveys/Cyclical Indicators	108
	S&P CoreLogic CS 20-City YoY NSA	SPCS20Y% Index	Housing and Real Estate	108
11	Construction Spending MoM	CNSTTMOM Index	Housing and Real Estate	107
	ISM Manufacturing	NAPMPMI Index	Surveys/Cyclical Indicators	108
	ISM Prices Paid	NAPMPRIC Index	Prices	108
12	Consumer Credit	CICRTOT Index	Personal/Household Sector	108
13	Continuing Claims	INJCSP Index	Labor Market	469
	EIA Natural Gas Storage Change	DOENUSCH Index	Energy	469

	Initial Jobless Claims	INJCJC Index	Labor Market	469
14	Current Account Balance	USCABAL Index	Intl Trade & BoP	36
15	DOE U.S. Crude Oil Inventories	DOEASCRD Index	Energy	469
	DOE U.S. Distillate Inventory	DOEASDIS Index	Energy	469
	DOE U.S. Gasoline Inventories	DOEASMGS Index	Energy	469
	DOE U.S. Refinery Utilization	DOEAUTIL Index	Energy	469
16	Dallas Fed Manf. Activity	DFEDGBA Index	Surveys/Cyclical Indicators	108
17	Durable Goods Orders	DGNOCHNG Index	Industrial Sector	108
	Durables Ex Transportation	DGNOXTCH Index	Industrial Sector	108
18	Empire Manufacturing	EMPRGBCI Index	Surveys/Cyclical Indicators	108
19	Employment Cost Index	ECI SA% Index	Labor Market	36
20	Existing Home Sales	ETSLTOTL Index	Housing and Real Estate	108
21	FHFA House Price Index MoM	HPIMMOM% Index	Housing and Real Estate	108
22	FOMC Rate Decision (Upper Bound)	FDTR Index	Monetary Sector	72
23	Factory Orders	TMNOCHNG Index	Industrial Sector	107
24	GDP Annualized QoQ	GDP CQOQ Index	National Accounts (GDP)	108
	Personal Consumption	GDPCTOT% Index	National Accounts (GDP)	108
25	ISM Non-Manf. Composite	NAPMNMI Index	Surveys/Cyclical Indicators	108
26	Import Price Index MoM	IMP1CHNG Index	Intl Trade & BoP	108
27	Industrial Production MoM	IP CHNG Index	Industrial Sector	108
28	Leading Index	LEI CHNG Index	Surveys/Cyclical Indicators	108
29	Monthly Budget Statement	FDDSSD Index	National Accounts (GDP)	108
30	NAHB Housing Market Index	USHBMIDX Index	Housing and Real Estate	108
31	New Home Sales	NHSLTOT Index	Housing and Real Estate	107
32	Nonfarm Productivity	PRODNFR% Index	Labor Market	72
	Unit Labor Costs	COSTNFR% Index	Labor Market	72
33	PPI Ex Food and Energy MoM	FDIDSGMO Index	Prices	108
34	Pending Home Sales MoM	USPHTMOM Index	Housing and Real Estate	109

35	Personal Consumption	GDPCTOT% Index	National Accounts (GDP)	108
36	PCE Core MoM	PCE CMOM Index	Prices	108
	Personal Income	PITLCHNG Index	Personal/Household Sector	108
	Personal Spending	PCE CRCH Index	Personal/Household Sector	108
37	Philadelphia Fed Business Outlook	OUTFGAF Index	Surveys/Cyclical Indicators	108
38	Richmond Fed Manufact. Index	RCHSINDX Index	Surveys/Cyclical Indicators	108
39	Trade Balance	USTBTOT Index	Intl Trade & BoP	108
40	U. of Mich. Sentiment	CONSSENT Index	Surveys/Cyclical Indicators	216
41	USDA Corn Prospective Planting	PPLNCORN Index	Agriculture	9
	USDA Soybean Prosp. Plantings	PPLNSOYB Index	Agriculture	9
42	USDA Quarterly All Wheat Stock	UGRSAWTO Index	Agriculture	36
	USDA Quarterly Corn Stocks	UGRSCNTO Index	Agriculture	36
	USDA Quarterly Soybean Stocks	UGRSSBTO Index	Agriculture	36
43	WASDE Corn End Stocks	CUSEENDS Index	Agriculture	98
	WASDE Soybean End Stocks	SUSEENDS Index	Agriculture	98
	WASDE Total Wheat End Stocks	WUSETWES Index	Agriculture	105
44	Wards Total Vehicle Sales	SAARTOTL Index	Retail & Wholesale Sector	108
45	Wholesale Inventories MoM	MWINCHNG Index	Retail & Wholesale Sector	108

Table 5: List of EA News Announcements

	Event	**Ticker**	**Sector**	**No. Obs.**
1	Euro Area: Business Confidence	BEBCI Index	Surveys/Cyclical Indicators	108
2	Euro Area: CPI YoY	ECCPEMUY Index	Prices	108
3	Euro Area: M3 Money Supply YoY	ECMAM3YY Index	Monetary Sector	108
4	Euro Area: GfK Consumer Confidence	ECO1GFKC Index	Surveys/Cyclical Indicators	108
5	Euro Area: Business Climate Indicator	EUBCI Index	Surveys/Cyclical Indicators	108
	Euro Area: Economic Confidence	EUESEMU Index	Surveys/Cyclical Indicators	108
	Euro Area: Industrial Confidence	EUICEMU Index	Surveys/Cyclical Indicators	108
	Euro Area: Services Confidence	EUSCEMU Index	Surveys/Cyclical Indicators	108
6	Euro Area: Consumer Confidence	EUCCEMU Index	Surveys/Cyclical Indicators	203

7	Euro Area: GDP SA QoQ	EUGNEMUQ Index	National Accounts (GDP)	89
8	Euro Area: Household Cons QoQ	EUHNEMUQ Index	National Accounts (GDP)	46
9	Euro Area: Industrial Production SA MoM	EUITEMUM Index	Industrial Sector	108
10	Euro Area: PPI MoM	EUPPEMUM Index	Prices	108
11	Euro Area: ECB Main Refinancing Rate	EURR002W Index	Monetary Sector	96
12	France: Industrial Production MoM	FPIPMOM Index	Industrial Sector	108
	France: Manufacturing Production MoM	FRMPMOM Index	Industrial Sector	108
13	France: CPI YoY	FRCPIYOY Index	Prices	131
14	France: GDP QoQ	FRGEGDPQ Index	National Accounts (GDP)	80
15	France: Trade Balance	FRTEBAL Index	Intl Trade & BoP	108
16	France: ILO Unemployment Rate	FRUEREUO Index	Labor Market	36
17	Germany: Industrial Production WDA YoY	GEINYY Index	Industrial Sector	108
	Germany: Industrial Production SA MoM	GRIPIMOM Index	Industrial Sector	108
18	Germany: Factory Orders WDA YoY	GEIOYY Index	Industrial Sector	108
	Germany: Factory Orders MoM	GRIORTMM Index	Industrial Sector	108
19	Germany: Exports SA MoM	GRBTEXMM Index	Intl Trade & BoP	108
	Germany: Imports SA MoM	GRBTIMMM Index	Intl Trade & BoP	108
	Germany: Current Account Balance	GRCAEU Index	Intl Trade & BoP	108
	Germany: Trade Balance	GRTBALE Index	Intl Trade & BoP	108
20	Germany: CPI YoY	GRCP20YY Index	Prices	216
21	Germany: Retail Sales MoM	GRFRIAMM Index	Retail & Wholesale Sector	108
22	Germany: Exports QoQ	GRGDEXQ Index	National Accounts (GDP)	36
	Germany: Imports QoQ	GRGDIMQ Index	National Accounts (GDP)	36
	Germany: Private Consumption QoQ	GRGDPCQ Index	National Accounts (GDP)	36
23	Germany: GDP SA QoQ	GRGDPPGQ Index	National Accounts (GDP)	72
24	Germany: IFO Business Climate	GRIFPBUS Index	Surveys/Cyclical Indicators	108
	Germany: IFO Current Assessment	GRIFPCA Index	Surveys/Cyclical Indicators	108
	Germany: IFO Expectations	GRIFPEX Index	Surveys/Cyclical Indicators	108
25	Germany: Import Price Index MoM	GRIMP95M Index	Intl Trade & BoP	108
26	Germany: PPI MoM	GRPFIMOM Index	Prices	108

27	Germany: Trade Balance	GRTBALE Index	Intl Trade & BoP	108
28	Germany: Unemployment Change (000's)	GRUECHNG Index	Labor Market	108
	Germany: Unemployment Claims Rate SA	GRUEPR Index	Labor Market	108
29	Germany: ZEW Survey Current Situation	GRZECURR Index	Surveys/Cyclical Indicators	108
	Germany: ZEW Survey Expectations	GRZEWI Index	Surveys/Cyclical Indicators	108
30	France: Manufacturing Confidence	INSESYNT Index	Surveys/Cyclical Indicators	105
31	Italy: Manufacturing Confidence	ITBCI Index	Surveys/Cyclical Indicators	108
32	Italy: CPI EU Harmonized YoY	ITCPEY Index	Prices	215
33	Italy: Unemployment Rate Quarterly	ITEMUNES Index	Labor Market	36
34	Italy: GDP WDA QoQ	ITPIRLQS Index	National Accounts (GDP)	71
35	Italy: Industrial Production MoM	ITPRSANM Index	Industrial Sector	108
36	Italy: Consumer Confidence Index	ITPSSA Index	Surveys/Cyclical Indicators	108
37	Portugal: GDP QoQ	PTGDPQOQ Index	National Accounts (GDP)	72
38	Euro Area: Retail Sales MoM	RSSAEMUM Index	Retail & Wholesale Sector	108
39	Euro Area: Sentix Investor Confidence	SNTEEUGX Index	Surveys/Cyclical Indicators	108
40	Spain: CPI EU Harmonised YoY	SPCPEUYY Index	Prices	217
41	Spain: GDP QoQ	SPNAGDPQ Index	National Accounts (GDP)	72
42	Spain: Unemployment Rate	SPUNEMPR Index	Labor Market	36
43	Euro Area: Unemployment Rate	UMRTEMU Index	Labor Market	108

Table 6: List of UK News Announcements

	Event	**Ticker**	**Sector**	**No. Obs.**
1	Average Weekly Earnings 3M/YoY	UKAWMWHO Index	Labor Market	95
	Weekly Earnings ex Bonus 3M/YoY	UKAWXTOM Index	Labor Market	95
	ILO Unemployment Rate 3Mths	UKUEILOR Index	Labor Market	108
	Jobless Claims Change	UKUEMOM Index	Labor Market	108
2	BOE Asset Purchase Target	UKAPTARG Index	Monetary Sector	95
	Bank of England Bank Rate	UKBRBASE Index	Monetary Sector	103
3	CPI YoY	UKRPCJYR Index	Prices	108

	RPI Ex Mort Int.Payments (YoY)	UKRPXYOY Index	Prices	108
	RPI YoY	UKRPYOY Index	Prices	108
	Retail Price Index	UKRPI Index	Retail & Wholesale Sector	108
4	Current Account Balance	UKCA Index	Intl Trade & BoP	36
5	GDP QoQ	UKGRABIQ Index	National Accounts (GDP)	108
6	GfK Consumer Confidence	UKCCI Index	Surveys/Cyclical Indicators	108
7	Government Spending QoQ	UKGENMYQ Index	National Accounts (GDP)	35
	Gross Fixed Capital Formation QoQ	UKGENPTQ Index	National Accounts (GDP)	35
	Imports QoQ	UKGEIKLQ Index	National Accounts (GDP)	35
	Private Consumption QoQ	UKGEABRQ Index	National Accounts (GDP)	35
8	Halifax House Prices MoM	UKHBSAMM Index	Housing and Real Estate	108
9	Manufacturing Production MoM	UKMPIMOM Index	Industrial Sector	108
	Industrial Production MoM	UKIPIMOM Index	Industrial Sector	108
10	Mortgage Approvals	UKMSVTVX Index	Housing and Real Estate	108
	Net Consumer Credit	UKMSB3PS Index	Personal/Household Sector	107
	Net Lending Sec. on Dwellings	UKMSVTVJ Index	Housing and Real Estate	108
11	Nationwide House PX MoM	UKNBAAMM Index	Housing and Real Estate	108
12	PPI Input NSA YoY	UKPPIINY Index	Prices	108
	PPI Output NSA MoM	UKPPIOC Index	Prices	108
13	PSNB ex Banking Groups	UKPSJ5II Index	National Accounts (GDP)	87
14	Public Sector Net Borrowing	UKPSNB Index	National Accounts (GDP)	108
15	RICS House Price Balance	UKRXPBAL Index	Housing and Real Estate	108
16	Retail Sales Ex Auto Fuel MoM	UKRVAMOM Index	Retail & Wholesale Sector	108
	Retail Sales Inc Auto Fuel MoM	UKRVINFM Index	Retail & Wholesale Sector	94
17	Trade Balance	UKTBTTBA Index	Intl Trade & BoP	108

Table 7: List of JP News Announcements

	Event	**Ticker**	**Sector**	**No. Obs.**
1	All Industry Activity Index MoM	JNTIAIAM Index	Industrial Sector	106
2	Annualized Housing Starts	JNHSAN Index	Housing and Real Estate	108

	Imports YoY	JNTBIMPY Index	Intl Trade & BoP	106
3	BoP Current Account Balance	JNBPAB Index	Intl Trade & BoP	108
4	Capital Spending YoY	JNVNYOYS Index	Industrial Sector	36
5	Core Machine Orders MoM	JNMOCHNG Index	Industrial Sector	108
6	Exports YoY	JNTBEXPY Index	Intl Trade & BoP	106
	Retail Trade YoY	JNNETYOY Index	Retail & Wholesale Sector	108
	Tankan Large Mfg Index	JNTSMFG Index	Surveys/Cyclical Indicators	36
7	GDP SA QoQ	JGDPQGDP Index	National Accounts (GDP)	72
	Industrial Production MoM	JNIPMOM Index	Industrial Sector	216
8	Housing Starts YoY	JNHSYOY Index	Housing and Real Estate	108
9	Job-To-Applicant Ratio	JBTARATE Index	Labor Market	108
10	Jobless Rate	JNUE Index	Labor Market	108
	Leading Index CI	JNCICLEI Index	Surveys/Cyclical Indicators	109
	PPI YoY	JNWSDYOY Index	Prices	108
11	Money Stock M2 YoY	JMNSM2Y Index	Monetary Sector	108
12	Money Stock M3 YoY	JMNSM3Y Index	Monetary Sector	108
	Natl CPI YoY	JNCPIYOY Index	Prices	108
13	Retail Sales MoM	JNRETMOM Index	Retail & Wholesale Sector	108
14	Tertiary Industry Index MoM	JNTIAMOM Index	Industrial Sector	108
15	Tokyo CPI Ex-Fresh Food YoY	JNCPTXFF Index	Prices	108
16	Overall Household Spending YoY	JHHSLERY Index	Personal/Household Sector	108
	Tokyo CPI YoY	JNCPT Index	Prices	108
	Trade Balance	JNTBAL Index	Intl Trade & BoP	108

Appendix B

Table 8: EUR/USD Individual News Variables - Top 5 Most Selected News

Sel.%	Full name	Sector
	Dependent variable: RV - HAR model	
100	**EA:** ECB Main Refinancing Rate	Monetary Sector
100	**US:** EIA Natural Gas Storage Change, Continuing Claim, Initial Jobless Claim	Energy, Labor Market
100	**US:** FOMC Rate Decision (Upper Bound)	Monetary Sector
100	**US:** Wholesale Inventories MoM	Retail & Wholesale Sector
99.2	**US:** Empire Manufacturing	Surveys/Cyclical Indicator
	Dependent variable: RV - HAR-CJ model	
100	**EA:** ECB Main Refinancing Rate	Monetary Sector
100	**US:** EIA Natural Gas Storage Change, Continuing Claim, Initial Jobless Claim	Energy, Labor Market
100	**US:** FOMC Rate Decision (Upper Bound)	Monetary Sector
99.5	**US:** DOE U.S. Refinery Utilization, DOE U.S. Crude Oil Inventories, DOE U.S. Distillate Inventory, DOE U.S. Gasoline Inventories	Energy
98.9	**US:** Change in Private Payrolls, Change in Manufact. Payrolls, Change in Nonfarm Payrolls, Average Hourly Earnings MoM, Average Weekly Hours All Employees, Unemployment Rate	Labor market
	Dependent variable: CC - HAR-CJ model	
100	**EA:** ECB Main Refinancing Rate	Monetary Sector
100	**US:** FOMC Rate Decision (Upper Bound)	Monetary Sector
98.9	**US:** Change in Private Payrolls, Change in Manufact. Payrolls, Change in Nonfarm Payrolls, Average Hourly Earnings MoM, Average Weekly Hours All Employees, Unemployment Rate	Labor market
99.9	**US:** Empire Manufacturing	Surveys/Cyclical Indicator
99.7	**US:** DOE U.S. Refinery Utilization, DOE U.S. Crude Oil Inventories, DOE U.S. Distillate Inventory, DOE U.S. Gasoline Inventories	Energy
	Dependent variable: JC - HAR-CJ model	
83.6	**US:** DOE U.S. Refinery Utilization, DOE U.S. Crude Oil Inventories, DOE U.S. Distillate Inventory, DOE U.S. Gasoline Inventories	Energy
67.1	**EA:** ECB Main Refinancing Rate	Monetary Sector
66.7	**US:** EIA Natural Gas Storage Change, Continuing Claim, Initial Jobless Claim	Energy, Labor Market
61.9	**US:** Change in Private Payrolls, Change in Manufact. Payrolls, Change in Nonfarm Payrolls, Average Hourly Earnings MoM, Average Weekly Hours All Employees, Unemployment Rate	Labor market
61.3	**US:** Trade Balance	Intl Trade & BoP

Table 9: GBP/USD Most Selected News According to BMA

Sel.%	Full name	Sector
	Dependent variable: RV - HAR model	
100	US: FOMC Rate Decision (Upper Bound)	Monetary Sector
99.8	US: Chicago Purchasing Manager	Surveys/Cyclical Indicators
98.7	UK: BOE Asset Purchase Target, Bank of England Bank Rate	Monetary Sector
98.6	US: Construction Spending MoM, ISM Manufacturing, ISM Prices Paid	Housing and Real Estate, Surveys/Cyclical Indicators, Prices
75.9	UK: Average Weekly Earnings 3M/YoY, ILO Unemployment Rate 3Mths, Jobless Claims Change, Weekly Earnings ex Bonus 3M/YoY	Labor Market
47.3	UK: PPI Input NSA YoY, PPI Output NSA MoM	Prices
35.8	US: Average Hourly Earnings MoM, Average Weekly Hours All Employees, Change in Manufact. Payrolls, Change in Nonfarm Payrolls, Change in Private Payrolls, Unemployment Rate	Labor Market
24.4	UK: CPI YoY, RPI Ex Mort Int.Payments (YoY), RPI YoY, Retail Price Index	Prices
	Dependent variable: RV - HAR-CJ model	
100	US: FOMC Rate Decision (Upper Bound)	Prices
99.9	US: Dallas Fed Manf. Activity	Surveys/Cyclical Indicators
99.5	UK: BOE Asset Purchase Target, Bank of England Bank Rate	Prices
52.4	US: U. of Mich. Sentiment	Surveys/Cyclical Indicators
45.5	US: Cap Goods Orders Nondef Ex Air	Industrial Sector
23.9	UK: PPI Input NSA YoY, PPI Output NSA MoM	Prices
	Dependent variable: RV - HAR-CJ model	
54.4	US: Wholesale Inventories MoM	Retail & Wholesale Sector
17.5	UK: BOE Asset Purchase Target, Bank of England Bank Rate	Prices

Table 10: USD/JPY Most Selected News According to BMA

Sel.%	Full name	Sector
	Dependent variable: RV - HAR model	
100	US: Average Hourly Earnings MoM, Average Weekly Hours All Employees, Change in Manufact. Payrolls, Change in Nonfarm Payrolls, Change in Private Payrolls, Unemployment Rate	Labor Market
100	US: FOMC Rate Decision (Upper Bound)	Prices
88.3	JP: Annualized Housing Starts, Imports YoY	Housing and Real Estate, Intl Trade & BoP
42.5	US: Business Inventories, Retail Sales Advance MoM, Retail Sales Ex Auto MoM	Retail & Wholesale Sector
14.8	US: Dallas Fed Manf. Activity	Surveys/Cyclical Indicators
11.8	US: EIA Natural Gas Storage Change, Continuing Claims, Initial Jobless Claims	Energy, Labor Market
	Dependent variable: RV - HAR-CJ model	
95.9	US: Dallas Fed Manf. Activity	Retail & Wholesale Sector
76.2	US: FOMC Rate Decision (Upper Bound)	Prices
75.4	US: Average Hourly Earnings MoM, Average Weekly Hours All Employees, Change in Manufact. Payrolls, Change in Nonfarm Payrolls, Change in Private Payrolls, Unemployment Rate	Labor Market
31.1	US: EIA Natural Gas Storage Change, Continuing Claims, Initial Jobless Claims	Energy, Labor Market
	Dependent variable: RV - HAR-CJ model	
53.3	JP: PPI YoY, Capital Spending YoY, Core Machine Orders MoM	Prices, Industrial Sector
36.2	US: Business Inventories, Retail Sales Advance MoM, Retail Sales Ex Auto MoM	Retail & Wholesale Sector
30.1	US: Average Hourly Earnings MoM, Average Weekly Hours All Employees, Change in Manufact. Payrolls, Change in Nonfarm Payrolls, Change in Private Payrolls, Unemployment Rate	Labor Market
19.6	JP: Annualized Housing Starts, Imports YoY	Housing and Real Estate, Intl Trade & BoP