Mémoire

Multi-Asset Approach to Realized Variance Forecasting: Empirical Evidence from the Canadian Dollar

en vue de l'obtention du grade de

Maîtrise ès sciences
(M. Sc.) Ingénierie Financière

aux

HEC Montréal

supervisé par

David Ardia

proposé par

Alexandre Turenne

Avril 2022

Remerciements

Je suis reconnaissant envers monsieur David Ardia pour le support qu'il m'a apporté lorsque nécessaire tout au long de ce travail ainsi que l'autonomie qu'il m'a laissée, me permettant d'avancer à mon rythme et de concilier travail et études. Je tiens aussi à remercier les professeurs du programme, spécialement monsieur Hatem Ben Ameur pour ses méthodes d'apprentissage uniques et sa fougue contagieuse, ainsi que madame Geneviève Gauthier pour m'avoir inspiré à faire ce programme suite au cours *modèles probabilistes et stochastiques* du baccalauréat. Finalement, je remercie mes parents, ma famille et mes amis pour leur support inconditionnel durant mes études.

Résumé

La transmission de volatilité entre différents actifs est un sujet de recherche important en finance. Cette thèse vérifie si l'inclusion de la volatilité des principales classes d'actifs (i.e., les actions, les obligations, les matières premières et les devises) améliore la précision des modèles de prévision de volatilité sur le dollar canadien. Pour ce faire, nous utilisons la variance réalisée sur différents contrats à terme comme mesure observable de la variance. Nous introduisons des modèles multi-actifs permettant à la variance réalisée d'autres actifs d'influencer nos prévisions. Nous établissons empiriquement que le modèle HAR classique est difficile à battre avec les extensions HAR de la littérature. De plus, nous montrons que la volatilité d'autres actifs influence la volatilité des contrats à terme sur le dollar canadien. L'ajout de ces volatilités dans les modèles de prévision augmente leur performance.

Abstract

Volatility spillovers are important research topics in finance. This thesis evaluates whether including the volatilities of major asset classes (i.e., equities, bonds, commodities, and currencies) improves the accuracy of volatility forecasting models on the Canadian dollar. To do so, we use realized variance as an observable measure of the variance across different futures contracts. We introduce multi-asset models allowing the realized variance of other assets to influence our forecasts. We provide compelling empirical evidence that the classic HAR model is hard to beat with HAR extensions from the literature. Furthermore, we show that volatilities from other assets influence the CAD/USD futures volatility. Adding these volatilities to volatility forecasting models increases their performance.

Contents

List of Figures

List of Tables

Acronyms

ADF	Augmented Dickey-Fuller
BTS	Business Time Sampling
CTS	Calendar Time Sampling
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
HAR	Heterogeneous Autoregressive
JB	Jarque-Bera
RMSE	Root Mean Squared Error
MAE	Mean Absolute Error
MCS	Model Confidence Set
OF	Overvaluation Frequency
OLS	Ordinary Least Squares
ORV	Overnight Realized Variance
RSV	Realized Semi-Variance
RV	Realized Variance
TTS	Transaction Time Sampling

1 Introduction

Having accurate volatility forecasts is of crucial importance in finance. Every significant market participant needs the most precise volatility estimates, from traders to policymakers and researchers. They use volatility or variance as an input in multiple applications. These applications vary from decisions on which assets to actively trade for a given day to volatility trading, option pricing, risk management, portfolio allocation, and model calibration.

Traditional models used by practitioners and academics to measure volatility, such as the family of GARCH models and stochastic volatility models, consider volatility as a function of observables or as a latent variable. This makes the measurement of volatility dependent on the model parameters and assumptions. Most of these models are not able to properly describe essential relations in the process of prices of financial assets. A growing amount of literature empirically demonstrates that returns are not normally distributed and instead have heavy-tails (see, Bates (1996); Begin, Dorion, and Gauthier (2020); Ornthanalai (2014)). For this category of models, correctly describing the dynamics of the returns is of the utmost importance to accurately forecast volatility.

With the advent of greater computing power, we can now fairly easily process high-frequency data. This allows to directly measure volatility, giving life to Merton (1980) idea of using the sum of squared returns at the high-frequency level for an accurate estimation of the variance on any financial asset. Researchers from the start of the millennium have formalized this idea into the current financial theory giving birth to the realized variance as a proxy for the variance of an asset.[1]Volatility now becomes observable, which allows for the use of econometric models more aligned with empirical evidence of price movements.

This thesis aims to study one-day-ahead log-realized variance forecasts for the Canadian dollar futures using high-frequency data. We base our empirical study on the HAR model of Corsi (2009), which is now the standard model for forecasting volatility when high-frequency data is available. This model outperforms traditional models such as GARCH and is rivaling with more advanced models while being easy to implement (see, Aganin (2017); L. Y. Liu, Patton, and Sheppard (2015); Ma, Wei, Huang, and Chen (2014); Vortelinos (2017)).

[1]For a review of the first uses of realized variance see the following, Andersen, Bollerslev, Diebold, and Labys (2003); Barndorff-Nielsen and Shephard (2002); Meddahi (2002); Poon and Granger (2003)

The US dollar is the most traded currency against the Canadian dollar in the Canadian economy. Being our biggest trading partner, the United States of America explains this situation. The US dollar is also the world's reserve currency, and a large number of investments are made with it. For these reasons, the CAD/USD exchange rate is of high importance to a large group of stakeholders, such as traders, investors, business owners, and policymakers. Correctly forecasting its volatility would help most actors in the Canadian financial market. We are the first to study the forecast of the log-realized variance on the CAD/USD exchange rate [2] with a wide range of different specifications of HAR models using high-frequency data and incorporating a diversified intermarket analysis.

Intermarket analysis is a method of analyzing markets by examining the correlations between different asset classes. Instead of looking at financial markets or asset classes individually, intermarket analysis looks at several asset classes, such as stocks, bonds, currencies, and commodities. This type of analysis expands on simply looking at each market or asset in isolation by looking at other markets or assets that have a strong relationship to the market or asset being considered. Traders have been using intermaket analysis for a long time, using what academics call market spillovers.

Two main reasons stand out for using different asset classes to forecast the Canadian dollar futures volatility. Firstly, the globalization of finance has created important interrelations between asset classes (see, Greenspan (1997); Hausler (2002); Piffaut and Miro (2019)). Secondly, other studies have shown the effects of yields, inflation, and other macroeconomics data on the market (see, Fama (1981); Gjerde and Saettem (1999); Hashim, Ramlan, Razali, and Nordin (2017); Naifar and Al Dohaiman (2013); Ratanapakorn and Sharma (2007)). We use different assets such as bonds and the dollar index as the proxy of these variables, which gives us what we might consider high-frequency data on macroeconomic variables.

We provide compelling evidence that applying intermarket analysis using multi-asset models increases the forecast accuracy of volatility. Furthermore, an expanding-window scheme for training the models results in more accurate forecasts than the traditionally-used rolling-window method. The main findings of this thesis should be of interest to the practitioners and applied researchers who will be able to increase the accuracy of their models.

The remainder of this thesis is organized as follows. Section 2 presents the literature review, Section 3 presents the data, Section 4 describes the models used for the forecast study, Section 5 details the empirical framework, Section 6 presents the results, and Section 7 concludes.

[2]We use the Canadian dollar futures, trading on the CME, as a good proxy of the CAD/USD spot exchange rate

2 Literature Review

2.1 Volatility Forecasting in Forex

Most research studies on volatility forecasting of exchange rates use models with daily data instead of high-frequency data. So, Lam, and Li (1999) find that the stochastic volatility models provide a notable improvement in volatility forecasting of different exchange rates. The exchange rates with a performance improvement include the Canadian dollar against the US dollar. Vilasuso (2002) forecasts volatility using the FIGARCH for different exchange rates such as CAD/USD and observes gains in forecasting accuracy compared to the GARCH and IGARCH models. Sadorsky (2005) compares the forecasting performance of different low-frequency models on the CAD/USD foreign exchange rate. More recently, Pilbeam and Langeland (2015) look at the implied volatility derived from currency options. They find that it outperforms a series of different specifications of univariate GARCH models for forecasting the volatility of exchange rates.

The literature on volatility forecasting with high-frequency data is rich for the equity and commodity markets but almost non-existent for the foreign exchange market. Current evidence suggests that high-frequency data increases accuracy for forecasting applications over low-frequency data. Chortareas, Jiang, and Nankervis (2011) observe that high-frequency data significantly enhances the performance of volatility forecasts on the euro against the Swiss franc, UK pound, Japanese yen, and US dollar. Plíhal (2016) reviews the literature and concludes that the HAR model of Corsi (2009) provides better results than the GARCH(1,1). Lyócsa, Molnár, and Fedorko (2016) study various models for forecasting one-day-ahead volatility of the Czech koruna, Hungarian forint, and Polish zloty against the euro. Their findings are in agreement with other studies. The HAR model is rarely outperformed even when considering its extensions. Vortelinos (2017) finds that the simple HAR model outperforms a wide array of models such as GARCH and neural networks for forecasting realized volatility on different financial markets, including foreign exchange rates.

2.2 Realized Variance - RV

For context, we review the realized variance starting with the theoretical foundation, followed by the different sampling schemes, and conclude with the different types of estimators proposed in the literature.

2.2.1 Theoretical foundation

Building on the literature arguing for non-normal market returns, we expose the theoretical foundation of RV using the market model with jumps. There is already plenty of literature reviewing the theoretical and practical foundation of the classical diffusion model without jumps (see, Andersen et al. (2003); Barndorff-Nielsen and Shephard (2002); Bucci (2017); McAleer and Medeiros (2008); Meddahi (2002); Poon and Granger (2003)). Let P_t be the price of an asset and $S_t = \log(P_t)$ its log-transformation. We suppose that the following stochastic differential equation characterizes the log-price dynamics of this asset:

$$dS_t = \mu_t dt + \sigma_t dW_t + dJ_t \tag{2.1}$$

where W_t is a Brownian motion, J_t is a jump process. μ_t is the drift and σ_t the instantaneous stochastic volatility, strictly positive and square integrable.

The quadratic variation of a stochastic process as shown in Protter (2004) is of the following:

$$[X]_T = QV_T = \lim_{n \to \infty} \sum_{i=1}^{n} (X_{s_i} - X_{s_i - 1})^2 \tag{2.2}$$

In the context of the market model with jumps, the quadratic variation is equivalent to the following formulation for the whole duration of the process:

$$[S]_T = QV_T = \int_0^T \sigma_s^2 ds + \sum_{i=1}^{N_T} Y_i^2 \tag{2.3}$$

where Y_i corresponds to the size of jump i, and N_T to the total number of jumps on the whole trajectory.

Let an intraday return between t and $t + \tau$ be:

$$r_t = S_{t+i\frac{\tau}{n}} - S_{t+(i-1)\frac{\tau}{n}}, \quad i \in [1, n] \tag{2.4}$$

and the realized variance, which is a good proxy for the quadratic variation:

$$RV_{t,t+\tau} = \sum_{i=1}^{n} r_{t+i\frac{\tau}{n}}^2 \tag{2.5}$$

It follows that the daily quadratic variance from t to $t+\tau$ satisfies:

$$\begin{aligned} QV_{t,t+\tau} &= [S]_{t+\tau} - [S]_t \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} (S_{t+i\frac{\tau}{n}} - S_{t+(i-1)\frac{\tau}{n}})^2 \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} r^2_{t+i\frac{\tau}{n}} \\ &\simeq RV_{t,t+\tau} \end{aligned} \tag{2.6}$$

Therefore the daily quadratic variation for the jump process is:

$$QV_{t,t+\tau} = \int_t^{t+\tau} \sigma_s^2 ds + \sum_{i=N_t}^{N_{t+\tau}} Y_i^2 \tag{2.7}$$

Barndorff-Nielsen and Shephard (2004) show that the quadratic variation attributable to the continuous part of the process can be estimated by the bipower variation as follows:

$$BV_{t,t+\tau} = \frac{\pi}{2} \sum_{i=2}^{n} |r_{t+i\frac{\tau}{n}} r_{t+(i-2)\frac{\tau}{n}}| \tag{2.8}$$

This formulation allows for the isolation of the continuous part and the jump part of the quadratic variation.

2.2.2 Sampling

The use of realized variance for the modelization of the variance allows us to be independent of models by directly sampling from an observable variable. While this is an advantage, there are some problems regarding the practical implementation of the method. Some literature shows that microstructure noise is present in the markets coming from different sources such as the bid-ask spread and the non-continuous process of price discovery (see, Amihud and Mendelson (1987); Harris (1991); Madhavan (2000)). These effects can introduce a bias in the estimation of RV.

To reduce the impact of microstructure effects on the estimation of RV, such as the bid-ask spread, infrequent trading, price discreteness, and others (Bandi & Russell, 2008), the literature has developed different sampling methods of intraday returns. We present the main sampling schemes in the presence of a process with jumps as reported in Oomen (2005)

- Calendar time sampling (CTS) or sparse sampling: returns are selected at regular spaced time intervals, such as every 15 minutes. Oomen (2005) points out that a vast amount of the literature uses intervals between 5 and 15 minutes when using CTS.

- Business time sampling (BTS): calendar time is scaled in a way that the resulting jump intensity is constant. This scheme is infeasible if the intensity of the jump process is latent. Therefore, in practice, one can estimate the intensity process with standard non-parametric smoothing methods using transactions times. The BTS scheme is based on the expected number of transactions.
- Transaction time sampling (TTS): returns are selected at each n transactions
- Tick time sampling: returns are selected at each price movement

While Oomen (2005) shows that TTS is a better sampling scheme than CTS, especially on low volume trading days, on average, using TTS on highly liquid stocks does not increase the estimator's accuracy markedly. While it can be useful for illiquid stocks or assets, major futures are highly liquid products with relatively low noise. The practitioner should prefer CTS since the precision gain using TTS would most likely not be worth the hassle. Research seems clear that returns of liquid assets such as major commodities futures and financial futures are noise-free above five-minute sampling. Couleau, Serra, and Garcia (2019) find that even if market microstructure noise on live cattle futures increases observed price variance, its effect is not large and do not last for more than four minutes coming to the similar conclusion of Ait-Sahalia and Xiu (2019) for major stocks on the US market. Hu, Serra, and Garcia (2020) also find an increased pricing efficiency, mitigation of short-term volatility, and reduced spreads for most agricultural commodity futures with the increase of algorithmic quoting in the limit order book. Ait-Sahalia and Xiu (2019) identify that the common practice of treating five-minute returns as noise-free is a reasonably safe choice for data sampled only after 2009 for the stocks in the Dow Jones and S&P 100, but not for all the S&P 500 stocks. These results again show that market efficiency has improved over the past decade and that realized variance estimators can be employed on liquid stocks. Hansen and Lunde (2006b) show that the magnitude of the noise component has dramatically decreased in recent years and is small relative to the daily return variation of liquid stocks.

A lot of research has been made to find the optimal sampling frequency of returns (see, for instance, Ait-Sahalia, Mykland, and Zhang (2005); Bandi and Russell (2008)). However, multiple practical issues arise from these approaches. First of all, the optimal sampling frequency approaches are always based on the underlying assumptions of the noise process. Secondly, the noise process is most likely not the same in time, as shown by Gatheral and Oomen (2010). A simple ad hoc rule is much more robust because it does not imply any model to the price and noise process.

2.2.3 Specific Methods

The estimator to compute the realized variance is also an important topic of study in the literature. The microstructure noise can have a significant effect and create bias in the estimator of the RV depending on the methodology used. Initially, sparse sampling was used for the RV estimator, but more recently, numerous alternative estimators have been created. These alternative estimators make more efficient use of the data and reduce the bias created by the microstructure noise (see, Andersen, Bollerslev, and Meddahi (2011); Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008); Hansen and Lunde (2006a); Shephard, Barndorff-Nielsen, and Lunde (2006); Zhang, Mykland, and Aït-Sahalia (2005)).

Although sparse sampling reduces the bias, the variance increases due to discretization. To minimize this bias, Zhang et al. (2005) propose a method to use a more significant portion of the data by subsampling, therefore reducing the bias and variance. The methodology consists of dividing the trading day into different subsamples of x-minute returns. Doing the average of the subsamples, we can compute the subsampled estimator as follows:

$$RV_t^{(s)} = \frac{1}{k} \sum_{i=1}^{k} RV_t^{(i)} \tag{2.9}$$

where $RV_t^{(s)}$ corresponds to the subsampled realized variance on day t, $RV_t^{(i)}$ to the realized variance on day t of subsample i and k to the number of subsamples.

The practitioner is left confused choosing between numerous amount of estimators. Most studies on the merits of the new estimators mainly focus on the speed of convergence to their asymptotic distributions, which is not a reliable guide to finite sample performance. Gatheral and Oomen (2010) show that, in practice, the best variance estimator is not always the one suggested by theory. This can partly be explained by assumptions of the price process, which could be incorrect. They find that the best estimators of RV based on efficiency, implementation, practicality, and robustness are ad hoc implementation of a subsampling estimator, realized kernel, or maximum likelihood realized variance.

3 Data

For this study, we use high-frequency prices at the one-minute interval from FirstRate Data[1] on seven futures contracts to forecast the one-day-ahead log-realized variance of the Canadian dollar futures contract. Log-transformation of the realized variance is used for two main reasons. Firstly, log-realized variance is much less skewed and displays significantly reduced kurtosis making it easier to forecast with linear models (see, Bekaert and Hoerova (2014); Oomen (2001)). Furthermore, it ensures that the values of the forecasts remain on the whole real line (Wilms, Rombouts, & Croux, 2021). Table 3.1 gives information about the different futures contracts used for this empirical application. For the rest of this study, we will reference each futures contract by their product or ticker as presented in the following table.

Table 3.1 Futures contracts in the study

Product	Ticker	Asset Class	Exchange
Canadian Dollar (CAD/USD)	6C	Currencies	CME
E-mini S&P 500	ES	Equities	CME
Euro (EUR/USD)	6E	Currencies	CME
Gold	GC	Commodities	NYMEX
Japanese Yen (JPY/USD)	6J	Currencies	CME
10-Year T-Note	ZN	Bonds	CBOT
WTI Crude Oil	CL	Commodities	NYMEX

These assets are selected because of their supposed influence on the CAD/USD exchange rate returns or volatility. All the chosen assets are part of the four main asset classes: equities, commodities, currencies, and bonds. We limit ourselves to highly liquid futures to have an

[1] FirstRate Data is an organization providing high-resolution intraday stock market, crypto, futures, and FX data used by traders, hedge funds, and academic institutions.

unbiased estimation of the log-realized variance for each product. We also are limited to US futures for reasons of data availability.[2]

For the equities, we choose only the S&P 500 futures contract as it is the most significant financial market index in the world, and it can have an impact on the US dollar, which is often considered a safe-haven asset in times of financial turmoil (Maggiori, 2011).

Regarding the currencies, the euro (EUR/USD) and the Japanese yen (JPY/USD) have both been shown to be net givers of volatility to other currencies in a large array of studies (see, Inagaki (2007); Kitamura (2010); Nekhili, Mensi, and Vo (2021); Rajhans and Jain (2015); Singh, Nishant, and Kumar (2018)). To both of them, they make approximately 70% of the DXY index. The DXY index measures the value of the US dollar relative to the value of a basket of currencies. Other currencies such as AUD/USD and CHF/USD are less traded and do not have the characteristic of being net givers of volatility during long periods. We only select EUR/USD and JPY/USD for these reasons.

Concerning the commodities, we use crude oil and gold. Crude oil has a direct effect on crude exporting countries' currencies such as the Canadian dollar, therefore, impacting the CAD/USD exchange rate (Kemati (2018); Lizardo and Mollick (2010); Yang, Cai, and Hamori (2017)). Crude oil is also the most traded commodity with extremely liquid futures contracts. Gold has long been described as a safe-haven asset in times of financial turmoil and as a hedge against the devaluation of fiat currencies (Baur and McDermott (2010, 2016); Capie, Mills, and Wood (2005); Nguyen, Bedoui, Majdoub, Guesmi, and Chevallier (2020)). These make crude oil and gold great commodities to add to our study for forecasting the CAD/USD log-realized variance.

Finally, we use the U.S. 10 Year Treasury Note for the bonds. Yield differentials affect the exchange rate of two countries. Using bonds, we are able to exploit this relation with high-frequency data.

Our sample period ranges from 2008-04-07 to 2021-11-12. To remove inconsistencies in the data, we apply the following cleaning procedure:

- Remove holidays for which the equities cash market is closed, otherwise known as NYSE or Nasdaq holidays, since it leads to low volume on the futures market and reduced trading hours. Realized variance on these days may not give us a representative value of the actual variance.
- Remove days with more than 25% missing values or more than 5% of missing consecutive data

[2] High-frequency data from the Canadian markets such as the S&P/TSX and the Canada 10 Year Government Bond would have been a great addition to our list of futures contracts since we aim to forecast the log-realized variance of CAD/USD.

– Impute missing values using linear interpolation

– After the preceding steps, keep only the days for which we have data for the whole set of futures contracts

We start with 3,550 days in our sample. Following the cleaning procedure, we remove 244 days, 120 are holidays, and 124 are days where at least one futures have more than 25% of missing data or 5% of missing consecutive data. Therefore, we have 3,306 full trading days for our empirical study.

Calculation of the realized variance[3] for each futures is done on five-minute subsampled returns. As indicated earlier, a five-minute sampling frequency offers a good balance between utilizing finely sampled high-frequency returns and reducing bias coming from microstructure effects. By using subsampling, we are able to reduce the variance of our estimator compared to a standard five-minute sampling frequency.

We calculate two realized variance estimators: one for the day-trading session (9:30 am to 4 pm EST) and the other for the overnight session (6 pm to 9:30 am EST). The day-trading session realized variance shall be called RV, and the overnight realized variance ORV. The choice of segmenting the realized variance in two is made in agreement with the literature. Martens (2002) shows that the best daily volatility forecast is produced by modeling overnight volatility differently from intraday volatility in the stock market. Dutta and Sharma (2012) find that the trading day session behaves like a separate market from the overnight sessions for the S&P 500 futures. We assume that this segmentation between the day session and the overnight session is also proper for the other assets in our empirical application, such as the CAD/USD for which the majority of volume comes during the New York session, similarly to the S&P 500 futures.

[3]For convenience, we annualize the realized variance by multiplying by 252

4 Models

4.1 Benchmarks

4.1.1 Naive

To make sure our models perform better than a naive approach. We consider a simple model with no parameters, for which the best forecast for day t is the previous day value:

$$\log(RV_t) = \log(RV_{t-1}) + \epsilon_t \tag{4.1}$$

where RV_t is the realized variance for day t, RV_{t-1} is the realized variance from the previous trading day and ϵ_t is the error term.

4.1.2 HAR

Despite its simplicity, the HAR model proposed by Corsi (2009) can produce rich dynamics for the variance that closely resemble the empirical data. The HAR is based on the heterogeneous market hypothesis and the asymmetric propagation of variance between long and short-time horizons. To be more precise, the HAR contains three components of variance. The short-time component represents one-day variance, the medium time component corresponding to the average variance over a week, and finally, the long-term component for the monthly average variance. The heterogeneous market hypothesis claims that the three components represent the behavior of different types of market participants with different investment time horizons.

We use this model as our benchmark model, which is now standard in the literature. The following equation defines the model for the forecasting of the log-realized variance:

$$\log(RV_t) = \beta_0 + \beta_1 \log(RV_{t-1}) + \beta_2 \log(\overline{RV}_{t-5}) + \beta_3 \log(\overline{RV}_{t-22}) + \epsilon_t \tag{4.2}$$

where RV_t is the realized variance for day t, RV_{t-1} is the realized variance from the previous trading day, $\overline{RV}_{t-5}$ is the average realized variance from the previous five trading days, $\overline{RV}_{t-22}$ is the average realized variance from the previous 22 trading days.

The average realized variance over the past n days is defined as follows:

$$\overline{RV}_{t-n} = \frac{1}{n}\sum_{i=1}^{n} RV_{t-i} \tag{4.3}$$

In the HAR model, it is standard to estimate the betas by OLS.

4.2 Simple models

4.2.1 HAR-RSV

Patton and Sheppard (2015) find that the volatility associated with negative returns for the S&P 500 and 105 individual stocks have better predictive information about the future volatility than the volatility associated with positive returns. Realized semi-variances used in the model are defined as follows :

$$\begin{aligned} RV_t^{+} &= \sum_{i=1}^{n} r_i^2 I_{[r_i>0]} \\ RV_t^{-} &= \sum_{i=1}^{n} r_i^2 I_{[r_i<0]} \end{aligned} \tag{4.4}$$

To see if the relation between past negative realized variance and future realized variance is still intact, we add the previous day's negative realized variance as a regressor to the classical HAR:

$$\log(RV_t) = \beta_0 + \beta_1 \log(RV_{t-1}) + \beta_2 \log(\overline{RV}_{t-5}) + \beta_3 \log(\overline{RV}_{t-22}) + \gamma \log(RV_{t-1}^{-}) + \epsilon_t \tag{4.5}$$

where RV_{t-1}^{-} represents the previous day negative realized variance, and the other predictive variables correspond to the benchmark model. Patton and Sheppard (2015) more complex models did not show significantly better performance in out-of-sample forecasts. To keep the model parsimonious and less prone to noise, we only add the previous day's negative realized variance.

4.2.2 HAR-CJ

With the growing literature on jumps in asset prices, Andersen, Bollerslev, and Diebold (2007) and Corsi, Pirino, and Reno (2010) decompose an asset's price process into two components: the continuous component C and the jump component J. In their study, Andersen et al. (2007) rely on the bipower variation proposed by Barndorff-Nielsen and Shephard (2004, 2006):

$$BV_{t,t+\tau} = \frac{\pi}{2}\sum_{i=2}^{n} |r_{t+i\frac{\tau}{n}} r_{t+(i-2)\frac{\tau}{n}}| \tag{4.6}$$

The jump component on day t, J_t, is estimated as the difference between the realized variance and the realized bipower variation. In theory, the realized variance should always be greater than the bipower variation, but since it is not always the case in practice, J_t should be limited to positive numbers:

$$J_t = \max(RV_t - BV_t, 0) \tag{4.7}$$

The continuous part of the process is the difference between the quadratic variation, estimated by the realized variance, and the jump component:

$$C_t = RV_t - J_t \tag{4.8}$$

By replacing the realized variance in our benchmark model with the continuous and jump components, we get the equivalent of the HAR-CJ used by Andersen et al. (2007) for forecasting log-realized variance:

$$\begin{aligned} \log(RV_t) = &\beta_0 + \beta_1 \log(C_{t-1}) + \beta_2 \log(J_{t-1}) + \beta_3 \log(\overline{C}_{t-5}) \\ &+ \beta_4 \log(\overline{J}_{t-5}) + \beta_5 \log(\overline{C}_{t-22}) + \beta_6 \log(J_{t-22}) + \epsilon_t \end{aligned} \tag{4.9}$$

where J_{t-1} is defined by equation (4.7) and C_{t-1} by equation (4.7). $\overline{C}_{t-5}$, $\overline{J}_{t-5}$, $\overline{C}_{t-22}$ and $\overline{J}_{t-22}$ are the average of the continuous and jump components on their respective period:

$$\overline{C}_{t-n} = \frac{1}{n} \sum_{i=1}^{n} C_{t-i} \tag{4.10}$$

$$\overline{J}_{t-n} = \frac{1}{n} \sum_{i=1}^{n} J_{t-i} \tag{4.11}$$

4.2.3 HAR-ORV

Following the reasons exposed in the data section to separate the day-trading realized variance and the overnight realized variance in two different estimators instead of one, we propose a small variation to the benchmark HAR model to have a better forecast of variance. We argue that the overnight variance might still have important information for the day-trading session variance. We add the log-overnight realized variance as a regressor in the model to test this assumption:

$$\begin{aligned} \log(RV_t) = &\beta_0 + \beta_1 \log(RV_{t-1}) + \beta_2 \log(\overline{RV}_{t-5}) + \beta_3 \log(\overline{RV}_{t-22}) + \\ &\gamma \log(ORV_{t-1}) + \epsilon_t \end{aligned} \tag{4.12}$$

ORV_{t-1} corresponds to the overnight realized variance preceding the day-trading session t, which corresponds to the realized variance from the futures open at 6:00 pm EST until the New York session at 9:30 am EST on the following day.

4.2.4 HAR-X

With the globalization of markets, the interdependence of asset prices has increased in the past years (see, Beirne and Gieck (2014); Gravelle, Kichian, and Morley (2003); Lam and Ang (2006)). The financialization of commodity markets has made them highly linked to equity markets and currency markets. Bond yields directly affect stock prices because of the discounting of future cash flows and currencies because of interest rate spreads between different countries. In addition, the US dollar is considered a safe-haven asset in periods of global financial turmoil. For these reasons, we use various futures contracts representing the main asset classes (i.e., equities, commodities, currencies, and bonds) to exploit the effects of volatility transmission on the CAD/USD exchange rate. To be more precise, we create a HAR-X model for each futures contract in our data giving us six different models. The general specification is of the following:

$$\begin{aligned} \log(RV_t) =& \beta_0 + \beta_1 \log(RV_{t-1}) + \beta_2 \log(\overline{RV}_{t-5}) + \beta_3 \log(\overline{RV}_{t-22}) + \\ & \gamma \log(RVX_{t-1}) + \epsilon_t \end{aligned} \tag{4.13}$$

where RVX_{t-1} corresponds to the previous day's realized variance of one of the futures contracts in this study (i.e., 6E, ES, GC, 6J, ZN, CL). As can be noted, we only use the previous day's realized variance of the additional futures contracts because we expect the longer-term variance trends to be similar across assets.

4.3 Multi-asset models

4.3.1 HAR-X-Boosted

To consider more than one asset in forecasting the log-realized variance on the CAD/USD futures, we propose a model called HAR-X-Boosted. We use a lasso regression to remove selection and look-ahead biases, which would allow us to choose the best regressors. This type of regression is useful when having multiple regressors because it encourages simple and sparse models by selecting a subsample of variables. This is especially important in finance because of the small number of observations and high noise to signal ratio present in the data. Lasso regression also makes it possible to interpret a model better since "low-value" variables are deleted from the model. Although this gives us a dynamic model that can change for each training period, the general form of the model remains the same:

$$\begin{aligned} \log(RV_t) =& \beta_0 + \beta_1 \log(RV_{t-1}) + \beta_2 \log(\overline{RV}_{t-5}) + \beta_3 \log(\overline{RV}_{t-22}) \\ & + \sum_{i=1}^{6} \gamma_i \log(RVX^{(i)}_{t-1}) + \epsilon_t \end{aligned} \tag{4.14}$$

where $RVX^{(i)}_{t-1}$ corresponds to the previous day's realized variance of one of the futures contracts in this study (i.e., 6E, ES, GC, 6J, ZN, CL).

The betas and the gammas are estimated on each training set through quadratic programming. In a context of standard regression with betas as coefficients, we can find the coefficients by minimizing the following function:

$$\sum_{i=1}^{n}(y_i - \sum_j x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \tag{4.15}$$

where p is the number of predictors, n the number of observations, and λ a tuning parameter that controls the penalty's strength associated with high betas. The optimal lambda (i.e., λ^*) is usually found using k-fold cross-validation on the training set. The model is then recalibrated on the whole training set using λ^*. In our case, we keep the serial dependence of observations in the cross-validation process, which is crucial when working with time series.[1] To do this, we manually assign the observations in each fold in such a way that the serial dependence is always respected.

4.3.2 HAR-X-Averaging

Another approach to consider the information coming from the different assets consists of producing the average forecast of all HAR-X models. Using this approach, we can capture the information coming from each asset. Degiannakis and Filis (2017), C. Liu and Maheu (2009), Lyócsa and Molnár (2016), and Wang and Nishiyama (2015) have all shown an increase in performance when using a model averaging approach for forecasting realized variance. For a review of techniques in model averaging for econometrics, see Steel (2020) and Nonejad (2021). One of the features of model averaging is that it reduces model specification uncertainty. The HAR-X-Averaging averages the forecasts coming from each basic HAR-X model. The reason for excluding the HAR-X-Boosted is that we want to compare a complex model (i.e., HAR-X-Boosted) model with a combination of simpler models (i.e., HAR-X-Averaging).

[1]When working with time series, an alternative to cross-validation is using a BIC-type criterion to set the parameter λ. For an example of application, see Ardia, Bluteau, and Boudt (2019)

5 Empirical framework

5.1 Forecasting procedure

For this experiment, we use a rolling and an expanding window to forecast the one-day-ahead log-realized variance. We train the models at the end of each day before predicting the next day's log-realized variance. A vast amount of studies using the rolling-window estimation method use windows between 125 and 1,500 days (Cech and Baruník (2017); Corsi (2009); Lyócsa and Molnár (2016); Wilms et al. (2021)). Changing the estimation length affects the forecasting ability of the models presented. We consider windows of length 500 (two years) and 1,000 (four years) days. We start with 1,000 days for the first training set for the expanding-window method. The longest out-of-sample period possible with an estimation window of 1,000 days is from 2012-07-12 to 2021-11-12, approximately 2,284 trading days. Having the same out-of-sample period for all window sizes is important to compare results. Therefore, we use this period for every window length and training scheme. All models except the Naive model[1] and the HAR-X-Boosted[2] are estimated by OLS on each training set. Fig. 5.1 shows the rolling window scheme and the expanding window scheme used in our study.

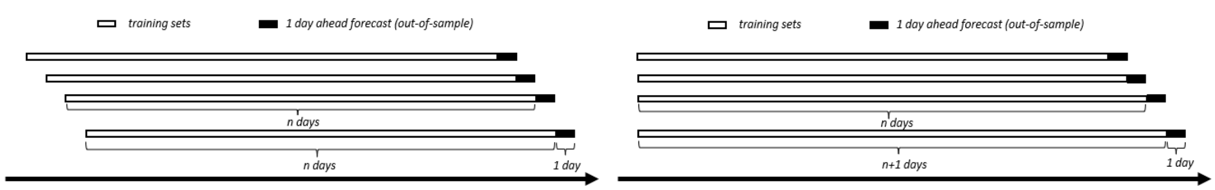

Fig. 5.1 Rolling-window (left) and expanding-window (right) schemes

The HAR-X-Boosted has to some degree, a more complicated estimation process. To determine the optimal lambda (i.e., λ^*) associated with the L1-penalty, we use five-fold cross-validation robust to the serial dependence of observations on each training set. The lasso regression is then calibrated on the whole training set before doing the forecast for the next day.

[1]The Naive model has no parameter to estimate

[2]The HAR-X-Boosted is estimated by minimizing equation (4.15) through quadratic programming

5.2 Forecasting accuracy

5.2.1 Loss functions

We compare the forecast performance of the simple models and the multi-asset models on how they can forecast the log-realized variance through the out-of-sample period. Comparing such performance is usually of more interest to the practitioner than in-sample analysis because it better represents models forecasting ability. The forecast performance of each model is evaluated with the root mean squared error (RMSE) and mean absolute error (MAE):

$$RMSE_{t,T} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\widehat{\log(RV_t)} - \log(RV_t))^2} \tag{5.1}$$

$$MAE_{t,T} = \frac{1}{T}\sum_{t=1}^{T}|\widehat{\log(RV_t)} - \log(RV_t)| \tag{5.2}$$

where $\widehat{\log(RV_t)}$ is the log-realized variance forecast at day t, $\log(RV_t)$ is the true value and T is the total number of days in the out-of-sample period. The MAE is used to make sure that a small amount of extreme errors does not drive up the RMSE, which could give wrong information about a model performance in the case where the RMSE was the only used performance measure.

5.2.2 Alternative evaluation

As an alternative to traditional performance measures and loss functions, we compute the overvaluation frequency of the forecast value with the actual value OF, giving us the proportion of forecasts above the real value:

$$OF_{t,T} = \frac{1}{T}\sum_{t=1}^{T} I_{[(\widehat{\log(RV_t)} > \log(RV_t)]} \tag{5.3}$$

This metric reveals if a model tends to be biased towards over-forecasting or under-forecasting the log-realized variance. Ideally, a model should have a value of 50% making the model without a directional bias.

5.2.3 Model Confidence Set

To test if the proposed models significantly outperform the standard HAR model in regards to the MAE and RMSE, we apply Hansen, Lunde, and Nason (2011) model confidence set (MCS) procedure. Given a set of models $\mathcal{M}_0$, we identify the model confidence set $\widehat{\mathcal{M}}^*_{1-\alpha} \subset \mathcal{M}_0$, which corresponds to the set of models that contains the significantly better forecasting models for a given loss function L and a given level of confidence α. The MCS procedure repeatedly tests the null hypothesis of equal forecasting accuracy:

$$H_0 : E[L_{i,t} - L_{j,t}] = 0 \tag{5.4}$$

where $L_{i,t} - L_{j,t}$ is the loss differential between model i and j. Starting with the full set of models $\mathcal{M} = \mathcal{M}_0$, this procedure sequentially eliminates the worst performing model from $\mathcal{M}$ until the hypothesis cannot be rejected anymore. The set of final models $\mathcal{M}_{1-a}$ is then called the MCS.

Using this procedure, we compare each model with the benchmark with respect to the MAE and RMSE. Therefore, we do pairwise tests for each loss function using the sets $\mathcal{M}_0 : \{benchmark, model_i\}$ where $model_i$ is any model other than the benchmark. In addition, to determine which models have the highest significantly better forecasting ability in their category, we also compute the model confidence set (MCS) on the group of simple models and multi-asset models.[3]

[3] All the MCS tests are performed using the Model Confidence Set Procedure package for R from Bernardi and Catania (2018)

6 Results

6.1 Realized variance

Fig. 6.1 plots the log-realized variance of every futures contract from April 2008 to November 2021. We observe periods of large fluctuations common to all assets, such as during the financial crisis of 2008 and the March 2020 Covid-19 episode. We also observe the increase of the log-realized variance for CAD/USD and crude oil around 2015. A significant collapse in oil prices caused this increase in volatility. Prices fell by almost 75% in less than two years.

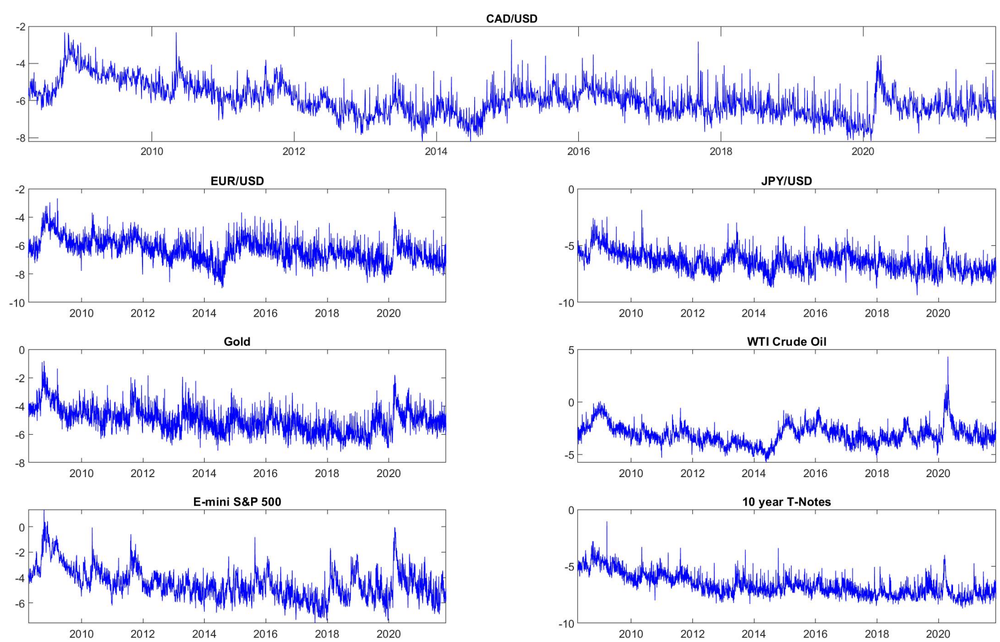

Fig. 6.1 Log-realized variances

Descriptive statistics are computed on each product and reported in Table 6.1. Results show that crude oil, gold, and the S&P 500 are the most volatile products with a respective mean log-realized variance of -3.101, -4.988, and -4.444. On the other hand, bonds and currencies are the least volatile. These results are in agreement with common knowledge in financial theory. At the 99% confidence level, the Augmented Dickey-Fuller test confirms the stationarity of the log-realized variances while the Jarque-Bera statistic rejects the null hypothesis of gaussian distribution. Even though the Jarque-Bera test rejects the null hypothesis, the log-realized variances show a kurtosis close to 3 with a slightly positive skewness. The distributions of log-realized variance are closer to normal distributions than the realized variance distributions. This is in agreement with other studies such as Oomen (2001) and Bekaert and Hoerova (2014). Skewness and kurtosis for the realized variance series are exposed in Table 6.2.

Table 6.1 Descriptive statistics of the log-realized variance estimates

Statistics	CAD	CL	EUR	ES	GC	JPY	ZN
Mean	-5.960	-3.101	-6.303	-4.444	-4.988	-6.405	-6.632
Median	-6.051	-3.186	-6.344	-4.604	-5.041	-6.460	-6.846
Std. Dev	0.897	0.980	0.873	1.269	0.937	0.934	0.996
Skewness	0.529	0.822	0.282	0.722	0.513	0.475	0.812
Kurtosis	3.307	5.183	3.208	3.794	3.423	3.601	3.368
ADF	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
JB	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

Notes: The descriptive statistics correspond to the mean, median, standard deviation, skewness, and kurtosis. The ADF and Jarque-Bera tests p-values are reported at less than 1%, therefore rejecting their null hypothesis at the 99% confidence level.

Table 6.2 Skewness and kurtosis of the realized variance estimates

Statistics	CAD	CL	EUR	ES	GC	JPY	ZN
Skewness	6.69	54.90	5.95	14.40	9.31	12.03	34.97
Kurtosis	74.38	3109.77	65.13	351.59	140.42	250.90	1654.46

6.2 Forecasting experiment

Fig. 6.2 plots the log-realized variance and the corresponding forecast of the best model overall (i.e., HAR-X-Averaging). We can observe that our model is successful in predicting the general trend of the log-realized variance but has difficulty predicting sudden large increases or decreases to their real extent.

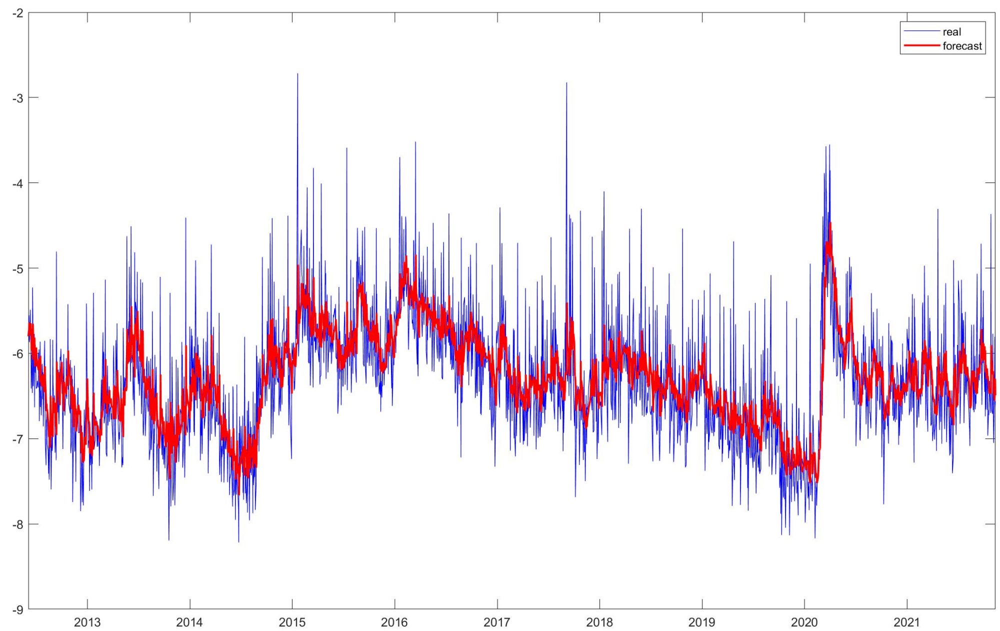

Fig. 6.2 Forecasts of the log-realized variance on CAD/USD futures

The MAE and RMSE of the forecasting experiment for every model and window size are reported in Table 6.5. The three models with the lowest value for every loss function and window size have shaded cells in Table 6.5. Bold values denote the smallest loss function value for each model. Asterisks indicate the statistical significance of the pairwise MCS procedure between the HAR and the corresponding model. The overvaluation frequency (OF) is also reported for additional information on the performance of the models.

6.2.1 Simple models

Simple models regroup all models, excluding the HAR-X-Boosted and the HAR-X-Averaging, which are both part of multi-asset models. In our study, the vast majority of simple models outperform the HAR on all training schemes in terms of MAE and RMSE. Still, the MCS pairwise procedure mostly fails to reject the null hypothesis of equal predictive ability between the benchmark and the corresponding model. This tells us that the outperformance might only come from chance and sample choice. It is worth noting that specifically, for the expanding-window scheme, most models' outperformance is statistically confirmed by the MCS. The HAR-X (ES) is the most consistent model in outperforming the HAR for all training schemes and often ranks in the top 3 of all models, even when including the multi-asset models. With a p-value of 1 for all training schemes and loss functions, Table 6.3 confirms that the HAR-X (ES) is the model of choice when considering only the simple models. Table 6.3 also validates that, excluding the expanding-window scheme, the HAR is hard to statistically outperform since it is part of the confidence set of the group best-performing models at a 75% confidence level.

Table 6.3 Log-Realized Variance: MCS p-values for the set of simple models

	500 days window		1,000 days window		expanding window	
Model	MAE	RMSE	MAE	RMSE	MAE	RMSE
Naive	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HAR	**0.7934**	**0.9986**	**0.8416**	**0.9994**	0.0806	0.0746
HAR-RSV	**0.7012**	**0.9934**	**0.9918**	**0.9998**	**0.8204**	**0.2508**
HAR-CJ	**0.9108**	**0.9682**	**1.0000**	**1.0000**	**0.9770**	**1.0000**
HAR-ORV	**0.7654**	**0.9986**	**0.9918**	**0.9998**	**0.9186**	0.0858
HAR-X (ZN)	**0.4672**	**0.9884**	0.1276	**1.0000**	**0.9606**	**0.5482**
HAR-X (CL)	**1.0000**	**1.0000**	**1.0000**	**1.0000**	0.0768	**0.8686**
HAR-X (GC)	**0.8858**	**0.9890**	**1.0000**	**1.0000**	**0.9988**	**0.9684**
HAR-X (6E)	**0.9858**	**1.0000**	**0.8418**	**0.9984**	**0.8160**	0.2036
HAR-X (6J)	**0.9998**	**1.0000**	**1.0000**	**1.0000**	**0.8920**	**0.5320**
HAR-X (ES)	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**

Notes: Values in bold denote that the model is part of the confidence set of best models at a confidence level of 75%.

6.2.2 Multi-asset models

For the multi-asset models, the evidence is clear. Multi-asset models outperform the HAR for all training schemes. This finding is confirmed by Table 6.4 which shows that the HAR is never part of the confidence set of best-performing models. Furthermore, the simpler model that averages the forecast of each HAR-X model tends to outperform the more complex one based on a lasso regression. One reason for this result could be the high noise to signal ratio in financial data as reported by Israel, Kelly, and Moskowitz (2020). Another reason given by the same authors is the small amount of data available when training the model on a daily timeframe. This second reason could explain why the HAR-X-Boosted competes better with the HAR-X-Averaging when using the expanding window scheme, which includes the most data in the training set. The HAR-X-Averaging consistently ranks in the top 3 models for the lowest MAE/RMSE in every training scheme. There is one exception, and that is for the RMSE in the expanding window model where the model is 4th with an RMSE of 0.5389 behind the HAR-X-Boosted, HAR-X (ES) and the HAR-CJ all of which have an RMSE of 0.5386. The results are also statistically significant at a significance level of at least 10%. The HAR-X-Averaging is the only model to rejects the null hypothesis of equal predictive ability with the benchmark HAR for all loss functions and training schemes. For these reasons, we claim that the HAR-X-Averaging is the best model in our study. Adding information about the variance of other assets can help increase the forecasting performance of models.

Table 6.4 Log-Realized Variance: MCS p-values for the set of multi-asset models

	500 days window		1,000 days window		expanding window	
Model	MAE	RMSE	MAE	RMSE	MAE	RMSE
Naive	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HAR	0.0382	0.1744	0.1254	0.1930	0.0002	0.0230
HAR-X-Boosted	**0.9582**	**1.0000**	**0.2800**	**0.8454**	**0.8278**	**1.0000**
HAR-X-Averaging	**1.0000**	**0.9954**	**1.0000**	**1.0000**	**1.0000**	**0.9212**

Notes: Values in bold denote that the model is part of the confidence set of best models at a confidence level of 75%.

6.2.3 Window size

Bold values in Table 6.5 clearly show that for all models except the HAR-X (CL), the best performing training scheme corresponds to the expanding window. This result is not surprising since log-realized variances tend to be close to gaussian. Having more data helps us reduce the generalization error of our models since the larger sample gives us a more accurate representation of the underlying distribution of the log-realized variance. Another important result is that we can see an overall decrease in the forecasting error for both the MAE and RMSE when increasing the window size for the rolling-window scheme. The 500 days window has a larger error than the 1,000 days window for every model. The overvaluation of forecasts also decreases with window size, which also points toward more precise forecasts when increasing the calibration set size. In Table 6.5 The pairwise MCS procedure confirms that the expanding window scheme is the best scheme for estimating the models. A good amount of models statistically outperforms the HAR at a significance level of at least 15% with that scheme. In contrast, other schemes cannot reject the null hypothesis of equal predictive ability as strongly.

Table 6.5 Log-Realized Variance: Forecasting Performance

	500 days window			1,000 days window			expanding window		
Model	MAE	RMSE	OF	MAE	RMSE	OF	MAE	RMSE	OF
Naive	0.5047	0.6861	0.5213						
HAR	0.4031***	0.5424***	0.5821	0.3992***	**0.5412*****	0.5735	**0.3990*****	0.5413***	0.5748
HAR-RSV	0.4034	0.5426	0.5817	0.3990†	0.5411	0.5709	**0.3985**†	**0.5410**†	0.5709
HAR-CJ	0.4061	0.5521	0.5843	0.3989	0.5406	0.5761	**0.3975*****	**0.5386*****	0.5774
HAR-ORV	0.4033	0.5426	0.5860	0.3989*	0.5412	0.5730	**0.3980*****	**0.5411**	0.5687
HAR-X (ZN)	0.4037	0.5419	0.5786	0.4002	0.5406	0.5769	**0.3975*****	**0.5402***	0.5696
HAR-X (CL)	0.4008*	0.5412†	0.5799	**0.3991**	**0.5401**	0.5860	0.4006	0.5408	0.5937
HAR-X (GC)	0.4030	0.5420	0.5847	0.3988	0.5406	0.5752	**0.3966*****	**0.5400**†	0.5666
HAR-X (6E)	0.4025	0.5411	0.5812	0.3990	0.5408	0.5636	**0.3982*****	**0.5405****	0.5709
HAR-X (6J)	0.4016†	0.5409†	0.5765	0.3980†	0.5403	0.5636	**0.3980**†	**0.5403***	0.5761
HAR-X (ES)	0.4010†	0.5405†	0.5808	0.3984	0.5403	0.5752	**0.3955*****	**0.5386***	0.5713
HAR-X-Boosted	0.4009†	0.5393†	0.5752	0.3999	0.5398†	0.5756	**0.3973**†	**0.5386**†	0.5812
HAR-X-Averaging	0.4004***	0.5394**	0.5821	0.3977***	0.5390**	0.5765	**0.3967*****	**0.5389*****	0.5748

Notes: *, **, ***, †, denote statistical significance of the pairwise MCS procedure between the benchmark model (HAR) and the corresponding model at a significance level of 5%, 10%, 15%, and 50%. Take note that the benchmark model corresponds to the Naive model for the HAR results. Shaded cells show the three models with the lowest value for every loss function and window size. Bold values denote the smallest loss function value for each model. RMSE, MAE, and OF correspondingly stand for root mean squared error, mean absolute error, and overvaluation frequency.

6.2.4 Overvaluation frequency

One result not discussed yet is the over-forecasting frequency metric. As shown in Table 6.5, the naive model does best in this measure. This means that our models tend to forecast more often above the actual value than the naive model. Fig. 6.3 shows the squared errors of forecasts of the HAR-X-Averaging model. The red points correspond to central banks' rate statements, and the green points to special market events. It is interesting to observe that the vast majority of large errors come from days where there is a central bank rate statement. These statements are scheduled eight times per year for the Federal Reserve and the Bank of Canada (BOC). Our models over-forecasting tendency is undoubtedly affected by these important days for the foreign exchange market. The trained models are biased towards higher forecasts because of the considerable variance on these days. In other words, the distribution of variance differs greatly on days of central banks' rate statements from other days. Therefore, removing these days from our training sets and modelizing these days differently, such as adding a fixed amount to the forecast of log-realized variance, could correct the over-forecasting tendency and greatly increase the overall forecasting performance of our models. Another alternative would be to use robust regression estimation methods which would reduce sensitivity to outliers.

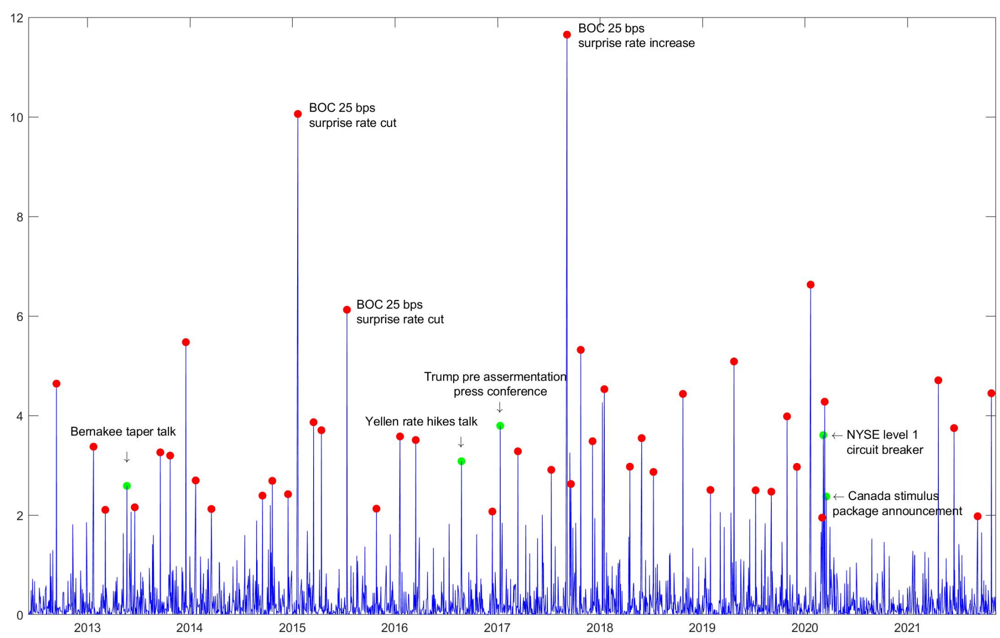

Fig. 6.3 Squared errors of the log-realized variance forecasts on CAD/USD futures

6.2.5 Robustness checks

Even though log-realized variance has more desirable properties for forecasting purposes, traders and investors work with volatility. Therefore, we validate our results by transforming the log-realized variance forecasts into realized volatility forecasts. We transform the log-realized variance into realized volatility as follows:

$$\widehat{RV_t^{0.5}} = e^{\widehat{\log(RV_t)}/2} \tag{6.1}$$

where $\widehat{\log(RV_t)}$ is the log-realized variance forecast at day t and $\widehat{RV_t^{0.5}}$ is the transformation of the forecast to realized volatility. We then, do the same analysis as on the log-realized variance. Table A.1, A.2, A.3 with the results are available in the Appendix.

Findings remain broadly the same when evaluating forecasts of the realized volatility. Results are also robust to different window sizes and loss functions as it is the case in section 6.2. Overall the standard HAR is hard to beat with other models from the literature, such as the HAR-RSV or HAR-CJ when using the rolling-window training method. Models with one exogenous variable corresponding to another asset often outperform the HAR but not enough to pass the pairwise MCS test. Multi-asset models outperform the HAR in terms of MAE and RMSE. MCS procedures confirm these results, but to a lesser degree than when forecasting log-realized variances.

6.2.6 Variable analysis

To determine what variables have the most effect on the log-realized variance forecasts, we look at the evolution of variables chosen through time in the HAR-X-Boosted trained on an expanding window. Fig. 6.4 shows the variables selected through days in the forecasting sample. The color gray represents inclusion of the variable in the model, while the color white represents a rejection of the variable (i.e., a beta of 0). The variables are displayed in the same order as Table 6.6, which presents the mean and median value of the beta associated with each variable for the whole forecasting sample.

We observe that not all variables are selected through time and that some variables have a much larger effect than others on the following day log-realized variance. The three variables with the most influence correspond to the one proposed by Corsi (2009) which are the daily, weekly, and monthly log-realized variance of the Canadian dollar futures. The previous day log-realized variance of the S&P 500 is the only variable selected through the whole sample. With a median beta value of 0.0781, it has an important effect on the next day log-realized variance of Canadian dollar futures. This could explain why the HAR-X (ES) has a performance much closer to the multi-asset models than all the other simple models. Crude oil and gold are also present in the HAR-X-Boosted for most of the sample, but their effect of the next day log-realized variance is more limited with a corresponding beta of 0.0258 and 0.0145. The ten-year bonds, euro, and Japanese yen have almost no effect on the log-realized variance of the Canadian dollar futures. Furthermore, their effect is not constant through time, as shown by their difference in mean and median.

Table 6.6 Mean and median value of absolute betas in the HAR-X-Boosted

Variable	Mean	Median
CAD - daily	0.1159	0.1118
CAD - weekly	0.3049	0.3007
CAD - monthly	0.2927	0.2979
ZN - daily	0.0085	0.0000
CL - daily	0.0226	0.0258
GC - daily	0.0148	0.0145
EUR - daily	0.0025	0.0000
JPY - daily	0.0098	0.0066
ES - daily	0.0791	0.0781

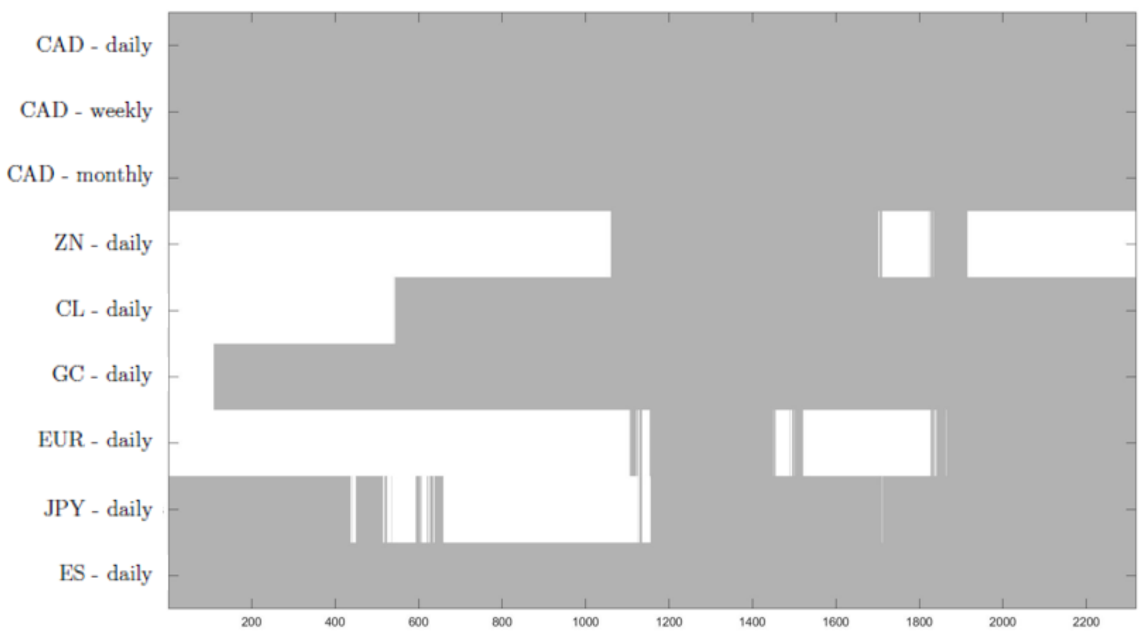

Fig. 6.4 Variable selection through time in the HAR-X-Boosted
Note: Gray represents inclusion/White rejection. The x-axis corresponds to the number of days in our forecasting sample. The y-axis corresponds to each variable in the model

6.2.7 Recommendations

Overall there is weak evidence that traditional models can outperform the standard HAR model using a fixed number of days rolling window for training(i.e. 500, 1,000 days) but the evidence is stronger with an expanding-window scheme. The HAR-X (ES) is the best model across training schemes when considering only the simple models. When considering all models, the HAR-X-Averaging stands out above others. All the models have a higher overvaluation frequency of forecasts compared to the naive model.

Following our results, we do the following useful recommendations to the practitioner:

- Using an expanding window training scheme which is not often considered in the literature on high-frequency volatility forecasting.
- Incorporating in models the information about the volatility of different assets for forecasting the log-realized variance of the CAD/USD exchange rate.
- Using the HAR-X (ES) as an alternative to the HAR if data for ES is available as it is significantly outperforming the HAR model.
- Removing the central banks' rate statements from the training procedure to reduce bias in the estimation of the models or using a robust approach for linear regression estimation.

7 Conclusion

This thesis investigates different models to outperform the classical HAR for one-day-ahead forecasts of the log-realized variance of Canadian dollar futures. More specifically, we are interested in knowing if adding information on the variance of different assets can improve such forecasts. To do so, we use two types of models. First, we use models from the literature such as the HAR-RSV, HAR-CJ, HAR-X. Second, we create models for modelizing the log-realized variance of the Canadian dollar futures with different assets (i.e., multi-asset models).

This thesis contributes to the literature in three important ways. First, we show that the HAR is hard to beat when using other models from the current literature. Second, our out-of-sample results demonstrate that multi-asset models consistently outperform the HAR over different training schemes and loss functions. In addition, other assets' past variance affects current variance on the Canadian dollar futures. Therefore, adding this information in forecasting models is crucial to have more accurate forecasts. Third and finally, we show that including more data in model training results in better performance and that an expanding-window scheme is preferred over a rolling-window scheme.

A significant proportion of the errors in our models' forecasts comes from days of central bank rate statements. Including these days in models' estimation also inevitably leads to an over-forecasting bias for "normal" days. Finding a way to correct this bias, such as removing these days from the estimation process and correcting forecasts on these days or using models robust to outliers (e.g., robust regressions), may increase the performance of the proposed models. Another interesting subject to consider would be to extend the multi-asset models approach to different assets hence confirming or refuting our hypothesis on the utility of these models for the whole sphere of tradable assets. We leave these subjects open for future research.

References

Aganin, A. (2017). Forecast comparison of volatility models on Russian stock market. *Russian Presidential Academy of National Economy and Public Administration (RANEPA), Applied Econometrics*, *48*, 63-84.

Ait-Sahalia, Y., Mykland, P. A., & Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies*, *18*(2), 351-416.

Ait-Sahalia, Y., & Xiu, D. (2019). A Hausman test for the presence of market microstructure noise in high frequency data. *Journal of Econometrics*, *211*(1), 176-205.

Amihud, Y., & Mendelson, H. (1987). Trading mechanisms and stock returns: An empirical investigation. *Journal of Econometrics*, *42*(3), 533-553.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *Review of Economics and Statistics*, *89*(4), 701-720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579-625.

Andersen, T. G., Bollerslev, T., & Meddahi, N. (2011). Realized volatility forecasting and market microstructure noise. *Journal of Econometrics*, *160*(1), 220-234.

Ardia, D., Bluteau, K., & Boudt, K. (2019). Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting*, *35*(4), 1370-1386.

Bandi, F. M., & Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies*, *75*(2), 339-369.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., & Shephard, N. (2008). Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, *76*(6), 1481-1536.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, *17*(5), 457-477.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1-37.

Barndorff-Nielsen, O. E., & Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, *4*(1), 1-30.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. *Review of Financial Studies*, *9*(1), 69-107.

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance*, *34*(8), 1886-1898.

Baur, D. G., & McDermott, T. K. (2016). Why is gold a safe haven? *Journal of Behavioral and Experimental Finance*, *10*, 63-71.

Begin, J. F., Dorion, C., & Gauthier, G. (2020). Idiosyncratic jump risk matters: Evidence from equity returns and options. *Review of Financial Studies*, *33*(1), 155-211.

Beirne, J., & Gieck, J. (2014). Interdependence and contagion in global asset markets. *Review of International Economics*, *22*(4), 639-659.

Bekaert, G., & Hoerova, M. (2014). The VIX, the variance premium and stock market volatility. *Journal of Econometrics*, *183*(2), 181-192.

Bernardi, M., & Catania, L. (2018). The model confidence set package for R. *International Journal of Computational Economics and Econometrics*, *8*(2), 144-158.

Bucci, A. (2017). Forecasting realized volatility: A review. *Journal of Advanced Studies in Finance*, *8*(16), 94-138.

Capie, F., Mills, T. C., & Wood, G. (2005). Gold as a hedge against the dollar. *Journal of International Financial Markets, Institutions and Money*, *15*(4), 343-352.

Cech, F., & Baruník, J. (2017). On the modelling and forecasting of multivariate realized volatility: Generalized heterogeneous autoregressive (ghar) model. *Journal of Forecasting*, *36*(2), 181-206.

Chortareas, G., Jiang, Y., & Nankervis, J. C. (2011). Forecasting exchange rate volatility using high-frequency data: Is the euro different? *International Journal of Forecasting*, *27*(4), 1089-1107.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, *7*(2), 174-196.

Corsi, F., Pirino, D., & Reno, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, *159*(2), 276-288.

Couleau, A., Serra, T., & Garcia, P. (2019). Microstructure noise and realized variance in the live cattle futures market. *American Journal of Agricultural Economics*, *101*(2), 563-578.

Degiannakis, S., & Filis, G. (2017). Forecasting oil price realized volatility using information channels from other asset classes. *Journal of International Money and Finance*, *76*, 28-49.

Dutta, S., & Sharma, S. C. (2012). Daytime vs. overnight trading in equity index futures markets. *Accounting and Finance Research*, *1*(2), 12-34.

Fama, E. F. (1981). Stock returns, real activity, inflation, and money. *American Economic Review*, *71*(4), 545-565.

Gatheral, J., & Oomen, R. C. (2010). Zero-intelligence realized variance estimation. *Finance and Stochastics*, *14*(2), 249-283.

Gauthier, G. (2020). *MATH 80631 Méthodes d'apprentissage appliquées aux données financières* (Realized Variance [PDF]). HEC Montréal.

Gjerde, O., & Saettem, F. (1999). Causal relations among stock returns and macroeconomic variables in a small, open economy. *Journal of International Financial Markets, Institutions & Money*, *9*(1), 61-74.

Gravelle, T., Kichian, M., & Morley, J. (2003). *Shift contagion in asset markets* (Working Paper No. 2003-5). Bank of Canada.

Greenspan, A. (1997). *Globalization of finance.* (Remarks at the 15th Annual Monetary Conference of the Cato Institute, October 14, Washington, D.C.)

Hansen, P. R., & Lunde, A. (2006a). Consistent ranking of volatility models. *Journal of Econometrics*, *131*(1-2), 97-121.

Hansen, P. R., & Lunde, A. (2006b). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, *24*(2), 127-161.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453-497.

Harris, L. (1991). Stock price clustering and discreteness. *Review of Financial Studies*, *4*(3), 389-415.

Hashim, S. L., Ramlan, H., Razali, N. H., & Nordin, N. Z. (2017). Macroeconomic variables affecting the volatility of gold price. *Journal of Global Business and Social Entrepreneurship*, *3*(5), 97-106.

Hausler, G. (2002). The globalization of finance. *Finance & Development*, *39*(1), 10-12.

Hu, Z., Serra, T., & Garcia, P. (2020). Algorithmic quoting, trading, and market quality in agricultural commodity futures markets. *Applied Economics*, *52*(58), 6277-6291.

Inagaki, K. (2007). Testing for volatility spillover between the British pound and the euro. *Research in International Business and Finance*, *21*(2), 161-174.

Israel, R., Kelly, B. T., & Moskowitz, T. J. (2020). Can machines "learn" finance? *Journal of Investment Management*, *18*(2), 23-36.

Kemati, R. (2018). Empirical analysis of crude oil price effects on exchange rate volatility. *Empirical Economic Review*, *1*(2), 17-48.

Kitamura, Y. (2010). Testing for intraday interdependence and volatility spillover among the euro, the pound and the Swiss franc markets. *Research in International Business and Finance*, *24*(2), 158-171.

Lam, S. S., & Ang, W. W. L. (2006). Globalization and stock market returns. *Global Economy Journal*, *6*(1), 1850082.

Liu, C., & Maheu, J. M. (2009). Forecasting realized volatility: A bayesian model-averaging approach. *Journal of Applied Econometrics*, *24*(5), 709-733.

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*(1), 293-311.

Lizardo, R. A., & Mollick, A. V. (2010). Oil price fluctuations and US dollar exchange rates. *Energy Economics*, *32*(2), 399-408.

Lyócsa, S., & Molnár, P. (2016). Volatility forecasting of strategically linked commodity ETFs: Gold-silver. *Quantitative Finance*, *16*(12), 1809-1822.

Lyócsa, S., Molnár, P., & Fedorko, I. (2016). Forecasting exchange rate volatility: The case of the Czech Republic, Hungary and Poland. *Finance a Uver*, *66*(5), 453-475.

Ma, F., Wei, Y., Huang, D., & Chen, Y. (2014). Which is the better forecasting model? A comparison between HAR-RV and multifractality volatility. *Physica A: Statistical Mechanics and its Applications*, *405*, 171-180.

Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets*, *3*(3), 205-258.

Maggiori, M. (2011). *The US dollar safety premium.* (AFA 2013 San Diego Meetings Paper)

Martens, M. (2002). Measuring and forecasting S&P 500 index-futures volatility using high-frequency data. *Journal of Futures Markets*, *22*(6), 497-518.

McAleer, M., & Medeiros, M. C. (2008). Realized volatility: A review. *Econometric Reviews*, *27*(1-3), 10-45.

Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, *17*(5), 479-508.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, *8*(4), 323-361.

Naifar, N., & Al Dohaiman, M. S. (2013). Nonlinear analysis among crude oil prices, stock markets' return and macroeconomic variables. *International Review of Economics & Finance*, *27*, 416-431.

Nekhili, R., Mensi, W., & Vo, X. V. (2021). Multiscale spillovers and connectedness between gold, copper, oil, wheat and currency markets. *Resources Policy*, *74*, 102263.

Nguyen, Q. N., Bedoui, R., Majdoub, N., Guesmi, K., & Chevallier, J. (2020). Hedging and safe-haven characteristics of gold against currencies: An investigation based on multivariate dynamic copula theory. *Resources Policy*, *68*, 101766.

Nonejad, N. (2021). An overview of dynamic model averaging techniques in time-series econometrics. *Journal of Economic Surveys*, *35*(2), 566-614.

Oomen, R. C. (2001). *Using high frequency stock market index data to calculate, model & forecast realized return variance* (Economics Discussion Paper No. 2001-6). European University Institute.

Oomen, R. C. (2005). Properties of bias-corrected realized variance under alternative sampling schemes. *Journal of Financial Econometrics*, *3*(4), 555-577.

Ornthanalai, C. (2014). Levy jump risk: Evidence from options and returns. *Journal of Financial Economics*, *112*(1), 69-90.

Patton, A. J., & Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics*, *97*(3), 683-697.

Piffaut, P. V., & Miro, D. R. (2019). Financial market volatility thresholds and its interrelation with the currency. *Journal of Emerging Issues in Economics, Finance and Banking*, *8*(1), 2782-2801.

Pilbeam, K., & Langeland, K. N. (2015). Forecasting exchange rate volatility: GARCH models versus implied volatility forecasts. *International Economics and Economic Policy*, *12*(1), 127-142.

Plíhal, T. (2016). Forecasting exchange rate volatility: Suggestions for further research. In *European financial systems 2016: Proceedings of the 13th international scientific conference* (p. 609-613).

Poon, S. H., & Granger, C. W. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, *41*(2), 478-539.

Protter, P. (2004). *Stochastic integration and differential equations.* Springer.

Rajhans, R. K., & Jain, A. (2015). Volatility spillover in foreign exchange markets. *Paradigm*, *19*(2), 137-151.

Ratanapakorn, O., & Sharma, S. C. (2007). Dynamic analysis between the us stock returns and the macroeconomic variables. *Applied Financial Economics*, *17*(5), 369-377.

Sadorsky, P. (2005). Stochastic volatility forecasting and risk management. *Applied Financial Economics*, *15*(2), 121-135.

Shephard, N., Barndorff-Nielsen, O. E., & Lunde, A. (2006). *Subsampling realised kernels* (Economics Series Working Papers No. 278). University of Oxford.

Singh, V. K., Nishant, S., & Kumar, P. (2018). Dynamic and directional network connectedness of crude oil and currencies: Evidence from implied volatility. *Energy Economics*, *76*, 48-36.

So, M. K., Lam, K., & Li, W. K. (1999). Forecasting exchange rate volatility using autoregressive random variance model. *Applied Financial Economics*, *9*(6), 583-591.

Steel, M. F. (2020). Model averaging and its use in economics. *Journal of Economic Literature*, *58*(3), 644-719.

Vilasuso, J. (2002). Forecasting exchange rate volatility. *Economics Letters*, *76*(1), 59-64.

Vortelinos, D. I. (2017). Forecasting realized volatility: HAR against principal components combining, neural networks and GARCH. *Research in International Business and Finance*, *39*, 824-839.

Wang, C., & Nishiyama, Y. (2015). Volatility forecast of stock indices by model averaging using high-frequency data. *International Review of Economics & Finance*, *40*, 324-337.

Wilms, I., Rombouts, J., & Croux, C. (2021). Multivariate volatility forecasts for stock market indices. *International Journal of Forecasting*, *37*(2), 484-499.

Yang, L., Cai, X. J., & Hamori, S. (2017). Does the crude oil price influence the exchange rates of oil-importing and oil-exporting countries differently? A wavelet coherence analysis. *International Review of Economics & Finance*, *49*, 536-547.

Yue, Q. I. U., & Tian, X. I. E. (2018). Forecasting foreign exchange realized volatility: A least square model averaging approach. *Journal of Systems Science and Mathematical Sciences*, *38*(6), 725.

Zhang, L., Mykland, P. A., & Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, *100*(472), 1394-1411.

A APPENDIX

Table A.1 Realized Volatility: MCS p-values for the set of simple models

	500 days window		1,000 days window		opening window	
Model	MAE	RMSE	MAE	RMSE	MAE	RMSE
Naive	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HAR	**0.8908**	**0.9634**	**0.5516**	**0.8460**	0.0084	**0.4074**
HAR-RSV	**0.8436**	**0.9636**	**0.9740**	**0.9772**	**0.4734**	**0.7634**
HAR-CJ	**0.7806**	**0.8758**	**0.9998**	**1.0000**	**0.7368**	**1.0000**
HAR-ORV	**0.8802**	**0.9712**	**0.9760**	**0.9646**	**0.6418**	**0.4396**
HAR-X (ZN)	**0.7378**	**1.0000**	0.1244	**1.0000**	**0.8118**	**0.6518**
HAR-X (CL)	**1.0000**	**1.0000**	**0.9912**	**1.0000**	0.0168	**1.0000**
HAR-X (GC)	**0.8728**	**0.9582**	**0.9998**	**0.9048**	**0.9400**	**0.8240**
HAR-X (6E)	**0.9564**	**0.9642**	**0.7560**	**0.7658**	**0.4490**	**0.4152**
HAR-X (6J)	**0.9982**	**0.8942**	**1.0000**	**0.7324**	**0.6086**	**0.6022**
HAR-X (ES)	**1.0000**	**1.0000**	**1.0000**	**0.9996**	**1.0000**	**1.0000**

Notes: Values in bold denote that the model is part of the confidence set of best models at a confidence level of 75%.

Table A.2 Realized Volatility: MCS p-values for the set of multi-asset models

	500 days window		1,000 days window		opening window	
Model	MAE	RMSE	MAE	RMSE	MAE	RMSE
Naive	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
HAR	0.0360	**0.2982**	0.1346	0.2418	0.0006	0.1230
HAR-X-Boosted	**1.0000**	**1.0000**	**0.5412**	**1.0000**	**1.0000**	**1.0000**
HAR-X-Averaging	**0.9712**	**0.9230**	**1.0000**	**0.6074**	**0.8904**	**0.5812**

Notes: Values in bold denote that the model is part of the confidence set of best models at a confidence level of 75%.

Table A.3 Realized Volatility: Forecasting Performance

	500 days window			1,000 days window			opening window		
Model	MAE	RMSE	OF	MAE	RMSE	OF	MAE	RMSE	OF
Naive	1.2578	2.0683	0.5213						
HAR	0.9666***	1.5926***	0.5821	**0.9585*****	1.5956***	0.5735	0.9603***	**1.5918*****	0.5748
HAR-RSV	0.9671	1.5924	0.5817	**0.9577**†	1.5945*	0.5709	0.9587†	**1.5898***	0.5709
HAR-CJ	0.9719	1.5989	0.5843	0.9574	1.5884†	0.5761	**0.9564***	**1.5845****	0.5774
HAR-ORV	0.9667	1.5923	0.5860	0.9575**	1.5949†	0.5730	**0.9573*****	**1.5909**†	0.5687
HAR-X (ZN)	0.9679	**1.5855**†	0.5786	0.9607	1.5885†	0.5769	**0.9559*****	1.5895†	0.5696
HAR-X (CL)	0.9610†	1.5873†	0.5799	**0.9586**	1.5883†	0.5860	0.9646	**1.5860**†	0.5937
HAR-X (GC)	0.9666	**1.5895**	0.5847	0.9570	1.5923	0.5752	**0.9537*****	1.5899	0.5666
HAR-X (6E)	0.9654	**1.5887**	0.5812	**0.9580**	1.5928	0.5683	0.9582***	1.5901†	0.5709
HAR-X (6J)	0.9637†	1.5899	0.5765	**0.9553***	1.5941	0.5636	0.9576*	**1.5897**†	0.5761
HAR-X (ES)	0.9600*	1.5851†	0.5808	0.9553†	1.5901	0.5752	**0.9489*****	**1.5849**†	0.5713
HAR-X-Boosted	0.9593*	1.5820†	0.5752	0.9579	1.5851†	0.5756	**0.9527***	**1.5813**†	0.5812
HAR-X-Averaging	0.9602***	**1.5834***	0.5821	0.9546**	1.5879*	0.5765	**0.9540*****	1.5858**	0.5748

Notes: *, **, ***, †, denote statistical significance of the pairwise MCS procedure between the benchmark model (HAR) and the corresponding model at a significance level of 5%, 10%, 15%, and 50%. Take note that the benchmark model corresponds to the Naive model for the HAR results. Shaded cells show the three models with the lowest value for every loss function and window size. Bold values denote the smallest loss function value for each model. RMSE, MAE, and OF correspondingly stand for root mean squared error, mean absolute error, and overvaluation frequency.