Erasmus University Rotterdam

Erasmus School of Economics

Bachelor Thesis: Econometrie en Operationele Research

Joint Lasso For Volatility Modelling

Author	*Student Number*	*Supervisor*	*Second Assessor*
Esther van de Velde	481023	A. Tetereva	W. Wang

Date Final Version: July 4, 2020

Abstract

In this paper, the *Heterogeneous Autoregressive Model of Realized Volatility* (HAR-RV) is considered. This model is usually estimated by ordinary least squares, but more recently, a body of literature documented the use of neural networks to forecast volatility. Neural networks estimate a model without making assumptions about the underlying form of the model. However, many days of intra-day data are needed for both estimation methods, which can be very bothersome to work with. Therefore, the joint lasso by Dondelinger et al. (2020) is applied to these models, as this operator helps to reduce the required sample sizes, by jointly estimating stock-specific (or subgroup of stocks-specific) regression coefficients. This can be useful as stocks tend to have similar volatility patterns over time: high volatility during crisis, but low and stable volatility during normal times. Empirical results, based on 10 stocks of the Dow Jones Industrial Average, show that by using the joint lasso, the sample size can be reduced while still maintaining accuracy. In fact, the HAR-RV model with such a joint lasso restriction, estimated using a small sample, outperforms a standard HAR-RV equation by equation model, estimated with a larger sample. The neural network with joint lasso restriction does not outperform the HAR-RV equation by equation significantly, but shows potential for larger sample sizes.

Contents

1 Introduction

Predicting the correct volatility structure of different stock returns, is of great interest for many financial institutions. It leads to efficient portfolios, risk management and option pricing models. However, these stock returns exhibit the well-known stylized facts, which are often challenging to model. The stylized facts of asset returns are listed by Cont (2001). One of these features is slowly decreasing autocorrelation in squared and absolute returns. Even though models exist that capture this feature to some extent, such as the GARCH model, Corsi (2009) states that for these models this decrease in autocorrelation is not slow enough. In addition, daily volatility is a nuisance parameter: it is not available without a model for returns. Luckily, high frequency data seems to be a promising solution to this problem. One can now predict the daily volatility of a stock based on 5, 10 minute or even millisecond return, the so called 'realized volatility'. Corsi (2009) uses this realized volatility to estimate a *Heterogeneous Autoregressive Model of Realized Volatility* (HAR-RV). The model predicts the realized volatility tomorrow based on the realized volatility today, this week and this month. The HAR-RV model has proven to be useful, as it is parsimonious and still is able to replicate the stylized facts.

However, the HAR-RV model still relies on some assumptions, such as normally distributed error terms and linearity. These assumptions do not always hold in practise. For this reason, neural networks have been used to forecast volatility. They suit this field particularly well, because these models are often able to recognize complex patterns in data without making any assumptions about the underlying data generating process. Bucci (2019) shows that these architectures can outperform standard econometric models of volatility.

A problematic part when working with both neural networks and the HAR-RV model is that they require a lot of data. The HAR-RV is estimated by standard linear regression (Heij, de Boer, Franses, Kloek, & van Dijk (2004)), which requires at least 20 observations per covariate and often performs better when more data is available (given that the relation between the independent and dependent variables is relatively stable). The neural networks often require around 70% of the data for training. When working with intra-day data, this can be bothersome. In addition, the volatility of stocks often shows similar patterns. In normal times, the volatility is low and relatively stable. During the financial crisis, stock returns fluctuated a lot, which resulted in high volatility. Most stocks are similar in this behaviour, but not identical. Using this information to model the volatility could be useful. Nevertheless, naive implementations of neural networks and the HAR-RV model ignore this information.

A recently assembled model solves these problems. The joint least absolute shrinkage and selection operator (joint lasso) could be used to receive estimates of the coefficients of a certain model. This operator is implemented for data that contains subgroups, which might differ in terms of underlying regression models, but still could be informative to each other. This kind of lasso allows information sharing between subgroups, by encouraging similarity between subgroup-specific coefficients, such that there is less data needed within one subgroup. The joint lasso has been used for a variety of biomedical topics, such as is done by Dondelinger, Mukherjee, & Alzheimer's Disease Neuroimaging Initiative (2020). The authors assume that one subgroup might have some useful information for another subgroup of the same illness, in this case Alzheimer. The same idea can be applied to stocks within and across industries, assuming that different stocks have similar volatility structures. The joint lasso implemented for a HAR-RV or neural network should be able to solve the problem of working with many days of intra-day data. To check for this, this model could be estimated using different rolling window sizes. This model should still be able to make flexible forecasts even for a small rolling window. At the same time, it tests the hypothesis that subgroups of stocks (e.g. stocks from the same industry) might have a similar volatility structure to other subgroups of stocks (e.g. from a closely related industry).

All in all, in this paper, the volatility forecasts of a regular HAR-RV equation by equation model will be compared to volatility forecasts of a HAR-RV model with joint lasso restrictions and those of a neural network model with HAR-RV inputs with a joint lasso restriction. This finally leads to the central question of this paper, which is: "To what extent it is useful to jointly determine the volatility structure of stocks using a joint lasso restriction?" To answer this question, data of the ten most liquid stocks of the Dow Jones Industrial Average (DJIA) for the period from February 2001 until December 2009 from the Oxford-Man Institute of Quantitative Finance (2009) will be used.

The results show that it is indeed useful to apply a joint lasso to reduce the sample size. In fact, a joint lasso HAR-RV model for small rolling window sizes outperforms a standard HAR-RV model for larger rolling window sizes. Nonetheless, similar results do not apply to the neural network, for which the joint lasso performs better for larger rolling window sizes.

The remainder of the paper is constructed as follows. In Section 2, a literature review is given. An introduction to the data set can be found in Section 3. Next, the methods are presented in Section 4. Section 5 contains the results. Section 6 concludes the paper.

2 Literature

In this paper, the volatility of different stocks is predicted. Stock returns often exhibit the stylized facts as described by Cont (2001). Therefore, there is need for sophisticated models that capture these features. Such models include stochastic volatility models, with as example the ARCH model by Engle (1982) and the GARCH model by Bollerslev (1986a). Especially the GARCH model does indeed seem to incorporate most of the stylized facts, which is directly deductible from how Bollerslev (1986a) describes his model. However, Corsi (2009), states that the autocorrelations of the squared and absolute returns often persist over a long time. Standard GARCH models and many other comparable models described by Bollerslev (1986b) are not able to reproduce this long term memory and converge to Gaussian noise over long periods of time (Gopikrishnan et al. (2000)). Therefore, Corsi (2009) introduces a *Heterogeneous Autoregressive model of Realized Volatility* (HAR-RV) model for long-memory volatility. The realized volatility is a proxy of the real volatility based on intra-day returns. This model forecasts the realized volatility the next day, by regressing on a constant, the realized volatility today, this week, and this month. Even though this model does not belong to the class of long-memory models, because of the way it is constructed (the form is similar to an AR model) the model has been showing promising results so far. As stated before, empirical asset data often shows volatility persistence, the HAR-RV model is able to capture this behaviour. In addition, it can also model the main stylized facts of financial data.

Even so, the HAR-RV model is still lacking as it assumes a few characteristics of the underlying data generating process, such as a linear relation between the dependent and independent variables. For this reason, the use of neural networks is upcoming for financial data. These algorithms do not make assumptions about the data generating process. Liu (2019) and Bucci (2019) show that neural networks can be used for volatility forecasts. In particular, the long short-term memory (LSTM) networks perform well, introduced by Hochreiter & Schmidhuber (1997). LSTMs are able to 'remember' relevant information of the inputs and overlook unimportant bits without having to specify the number of (time-)lags beforehand, making it a useful architecture for time series data. Similarly to the HAR-RV model, it has good performance when the independent variable is significantly correlated with lags of higher order.

However, studies show that some stocks might influence each other. A study by Karolyi (1995) confirms that different stock markets have effect on each other. Rosenow, Gopikrishnan, Plerou, & Stanley (2003) extend this and shows that there is interaction between the price change of different stocks. It is expected that stocks in the same industry are impacted similarly by economic shocks. This is also shown by Heston & Rouwenhorst (1995) and Drummen & Zimmermann (1992). The latter paper shows that an industry factor explains approximately 9% of the stock variance. So, it is plausible that stocks in similar industries have a comparable volatility structure and it would be interesting to incorporate this feature in the model. Ordinary least squares (OLS) (Heij et al. (2004)) estimation for the HAR-RV model, does not capture the similarity in stock behaviour. Čech & Baruník (2017) proposes a multivariate version of the HAR-RV model that makes use of the covariance structure of different stocks. Nevertheless, this model has two drawbacks: Firstly, Čech & Baruník (2017) have to use a rolling window with fixed length of 750 days. As stated before, it could be problematic to work with so many days of intra-day data. Secondly, it does not allow information sharing between the coefficients of stocks that are similar. It might be wasteful to ignore this information.

A more recently implemented model might be useful for this purpose, the joint lasso by Dondelinger et al. (2020). Their estimation algorithm is useful when looking at subgroups of data that might have similarities but are not identical. Subgroup specific estimation of a model might be challenging due to small sample size. Their solution is to jointly estimate subgroup-specific regression coefficients using a penalized framework in form of a lasso that also minimizes the difference between coefficients of different subgroups, the 'joint lasso' (JL). This allows for information-sharing between the subgroups. Audrino & Knaus (2016) show that a simple ℓ_1 norm penalty on the size of the coefficients, might be useful when estimating the HAR-RV model, if it is the true model. The JL model has so far only been used for biomedicial data and it would be interesting to see if it is implementable for financial time series data in order to reduce the data necessary to make accurate forecasts.

Thus, this paper contributes to this literature, because it applies the JL to financial data. The JL is rewritten to fit time series data. Additionally, it is examined whether the JL is able to outperform the equation by equation methods and neural networks, and still is able to do so for small data sets.

3 Data

The data is obtained from Oxford-Man Institute of Quantitative Finance (2009) and includes ten stocks over the period February 2001 until December 2009. During this time frame, there are two periods that can be classified as 'crisis', namely the stock market downturn in the period of 2001-2002 and the financial crisis in 2008-2009 (with 2007 already showing increased volatility). The stocks include the ten most liquid stocks from the Dow Jones Industrial Average (DJIA), Alcoa (AA), American Express (AXP), Bank of America (BAC), Coca Cola (KO), Du Pont (DD), General Electric (GE), International Business Machines (IBM), JP Morgan (JPM), Microsoft (MSFT), and Exxon Mobil (XOM). The data is originally from Noureldin et al. (2012) and is therefore cleaned accordingly. The data contains the daily realized volatility, $RV_t^{(day)}$, based on 5-minute intra-day returns, and the daily returns, r_t. In Figure 1, the realized volatility over time is shown.

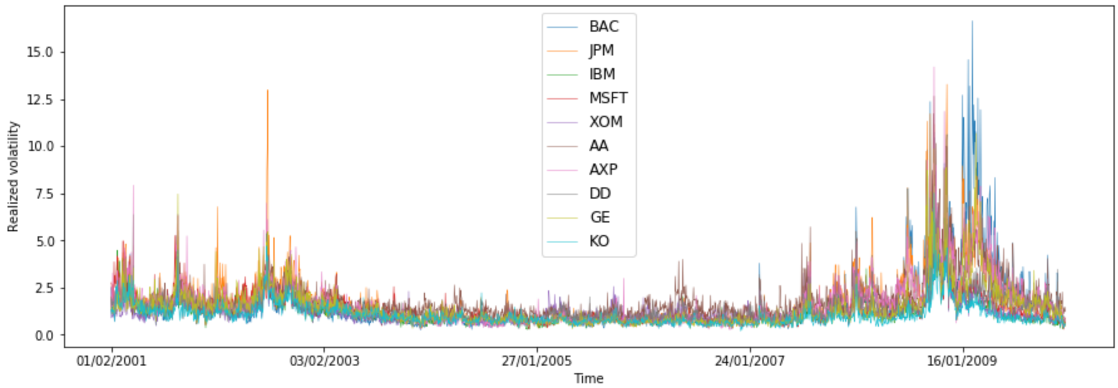


Figure 1: Realized volatility of the ten stocks over the period February 2001 until December 2009.

As can be seen from the figure, the volatility of the stocks shows similar movements. They are on comparable low level during non-crisis years and during crisis years their volatility increases significantly. However, not all in a similar way. During the financial crisis, BAC was definitely the most volatile. This is of course not unexpected as the crisis hit banks the hardest. The other two financial services stocks, AXP and JPM, show similar volatility to the BAC. Both IBM and MSFT have the lowest volatility during this period. They both belong to the computer technology industry. This again could mean that stocks belonging to a similar industry, do have similar volatility. Additional insights can be found in Table 1.

Table 1: Summary Statistics Realized Volatility

	AA	AXP	BAC	KO	DD	GE	IBM	JPM	MSFT	XOM
Maximum	160.24	201.88	277.31	56.51	63.87	114.26	57.54	176.48	43.11	115.38
Minimum	0.29	0.077	0.07	0.04	0.16	0.10	0.08	0.11	0.08	0.13
Average	4.94	4.42	5.46	1.41	2.53	3.20	1.93	5.06	2.45	2.07
Volatility	8.93	9.15	16.81	2.25	3.73	7.11	3.36	11.09	3.41	4.15

This table confirms that the financial stocks often have the highest volatility and that this volatility varies the most. Quiet stocks are KO, DD, IBM and MSFT. Especially KO has a low minimum. The more volatile stocks are the AA, AXP, BAC and JPM.

4 Methodology

4.1 Realized Volatility

The following notation is similar to Barndorff-Nielsen & Shephard (2004). Daily returns are equal to $r_t = y_t - y_{t-1}$, where y is the log price. Suppose now, that for a certain stock, there is data on the prices at M intervals with equal size on the t'th day. The high frequency return on day t in the jth intra-day interval is then equal to: $r_{j,t} = y_{t-1+\frac{j}{M}} - y_{t-1+\frac{(j-1)}{M}}$. The daily returns can then be obtained as $r_t = \sum_{j=1}^{M} r_{j,t}$.

Intra-day returns can be used to get an accurate proxy of the real volatility that day, the realized volatility (RV). There are a few ways to calculate the RV. Noureldin et al. (2012), who assembled the data set used in this paper, use the following simple formulation.

$$RV_t^{(day)} = \sqrt{\sum_{j=1}^{M} r_{j,t}^2}$$

Besides, Noureldin et al. (2012) use M = 288, because 5 minute intervals of returns are used. The weekly realized volatility $RV_t^{(week)}$ can be obtained from $RV_t^{(day)}$ in the following way: $RV_t^{(week)} = \frac{1}{5}(RV_t^{(day)} + RV_{t-1}^{(day)} + ... + RV_{t-4}^{(day)})$. Monthly realized volatility $RV_t^{(month)}$ is obtained from $RV_t^{(day)}$ in the following way: $RV_t^{(month)} = \frac{1}{22}(RV_t^{(day)} + RV_{t-1}^{(day)} + ... + RV_{t-21}^{(day)})$. Note that the number of trading days per week and month is 5 and 22 respectively.

4.2 Notation

Each stock $i \in \{1, \dots, N\}$ has the same p different features, namely a constant, daily, weekly and monthly realized volatility. The stocks are classified in a subgroup $k \in \{1, \dots, K\}$ according to a beforehand specified feature. The number of stocks in one subgroup k is denoted by n_k. Each stock could have a different time range, equal to T_i. This leads to a total sample of $T = \sum_{i=1}^{n} T_i$, and a subgroup sample of $T_k = \sum_{i \in k} T_i$. For each stock i, $RV_{i,t} = [1, RV_{i,t}^{(day)}, RV_{i,t}^{(week)}, RV_{i,t}^{(month)}]$ is the $1 \times p$ variable vector at time t and $RV_{i,t+1}^{(day)}$ is the corresponding daily realized volatility but a day later. $RV_{k,t}$ is the $n_k \times p$ variable matrix containing all the variables of the stocks in subgroup k. Similarly $RV_{k,t+1}^{(day)}$ $(k \times 1)$ is obtained.

4.3 HAR-RV Model

As stated before, the volatility is modelled using a HAR-RV (henceforth denoted by HAR) model by Corsi (2009), which has the following form.

$$RV_{t+1}^{(day)} = \beta^{(constant)} + \beta^{(day)} RV_t^{(day)} + \beta^{(week)} RV_t^{(week)} + \beta^{(month)} RV_t^{(month)} + \varepsilon_{t+1} \quad (1)$$

Here ε_{t+1} is the error term. The regression coefficients are $\beta = [\beta^{(constant)}, \beta^{(day)}, \beta^{(week)}, \beta^{(month)}]$. In case one creates subgroups of stocks, where each subgroup has the same coefficients for each lag of the realized volatility, one subgroup k of size n_k has the following form:

$$\begin{aligned}
RV_{1,t+1}^{(day)} &= \beta_k^{(constant)} + \beta_k^{(day)} RV_{1,t}^{(day)} + \beta_k^{(week)} RV_{1,t}^{(week)} + \beta_k^{(month)} RV_{1,t}^{(month)} + \varepsilon_{1,t+1} \\
RV_{2,t+1}^{(day)} &= \beta_k^{(constant)} + \beta_k^{(day)} RV_{2,t}^{(day)} + \beta_k^{(week)} RV_{2,t}^{(week)} + \beta_k^{(month)} RV_{2,t}^{(month)} + \varepsilon_{2,t+1} \\
&\vdots \\
RV_{n_k,t+1}^{(day)} &= \beta_k^{(constant)} + \beta_k^{(day)} RV_{n_k,t}^{(day)} + \beta_k^{(week)} RV_{n_k,t}^{(week)} + \beta_k^{(month)} RV_{n_k,t}^{(month)} + \varepsilon_{n_k,t+1}
\end{aligned}$$

All there is left now is to estimate $\beta_k = [\beta_k^{(constant)}, \beta_k^{(day)}, \beta_k^{(week)}, \beta_k^{(month)}]'$ per subgroup.

4.4 Joint Lasso

The coefficients in the model described in Section 4.3 are estimated using the JL of Dondelinger et al. (2020). Their estimation model has the following form:

$$\hat{B} = \underset{B=\beta_1 \dots \beta_k}{\arg\min} \sum_{k=1}^{K} \frac{1}{n_k} \| RV_{k,t+1}^{(day)} - RV_{k,t}\beta_k \|_2^2 + \lambda \|\beta_k\|_1 + \gamma \sum_{k'>k} \tau_{k,k'} f_q(\beta_k, \beta_{k'}), \tag{2}$$

where $B = [\beta_1 \dots \beta_K]$ and λ, γ and τ are tuning parameters. $\tau_{k,k'}$ gives the possibility to encourage similarity in coefficients between pairs of subgroups. Moreover, $f_q(\beta_k, \beta_{k'}) = \|\beta_k - \beta_{k'}\|_q^q$ for the ℓ_q norm ($q = 1, 2$). In case of ℓ_2 fusion, the last term is used to encourage similarity between subgroup-specific coefficients. Hence, in this case, coefficients of different subgroups can be similar but not equal. ℓ_1 fusion allows for equal coefficients.

The optimization methods are the same as in Dondelinger et al. (2020). The ℓ_2 formulation is easier to compute as the fusion part, $f_2(\beta_k, \beta_{k'})$, is continuously differentiable, which is why the SKLEARN (2020) package for Python (1995) is used. The ℓ_1 is more computationally exhaustive, so the FUSER (2017) package for R (2014) is used to compute the coefficients. By using the RPY2 (2020) package the FUSER (2017) package can still be worked from in Python (1995). For more details about the different algorithms used to optimize for both ℓ_2 and ℓ_1 fusions, the paper by Dondelinger et al. (2020) should be looked into.

4.5 Groups

To use the JL, described in Section 2, the stocks need to be grouped in some way. The stocks are grouped according agglomerative hierarchical clustering based on the euclidean distance between some common feature of the stocks. By using the euclidean distance, each element of each feature has equal importance. This feature is defined in two different manners.

The first feature is the industry the stock is subject to. A simple and straight-forward way to define this feature, is to use the market beta of a specific stock b_i, using the the the return on the overall market, r_m, and the return of the single stock, r_i. For this paper, the r_m is the return on the DJIA, because all the stocks are all included in this index. Finally, to compute this b_i, historical data of the last 5 months is used. It is expected that stocks from a similar industry have a similar market beta, therefore this method is referred to as 'industry grouping'.

The second feature is the coefficient estimates, $\tilde{\beta}_i$, of a stock when estimated by HAR equation by equation regression by in-sample data. Using agglomerative hierarchical clustering, the stocks are sorted into groups according to the similarity in $\tilde{\beta}_i$. The grouping is therefore done in a completely data driven way, without using economic insights. This method will be referred to as 'data driven grouping'.

In both cases the optimal number of groups is decided by cross-validation. The JLs estimated in these two different ways are compared to each other, after which the most preferred feature is chosen. The JL based on that feature is then compared to the other benchmark models, described later in Section 4.8.

4.6 Setting the Tuning Parameters

Both λ and γ are set using cross-validation. $\tau_{k,k'}$ could also be set using such method, but, Dondelinger et al. (2020) mention that this could be inconvenient in practice, for instance because of computational restrictions. Alternatively, $\tau_{k,k'}$ is set using a distance function $\delta(k,k')$. The distance function indicates the similarity of subgroups and allow for more fusion between similar groups. Because there are two different ways of grouping as described in Section 4.5, two different measures of distance are considered.

For the 'industry grouping', the distance function is $\delta(k,k') = \|b_k - b_{k'}\|$, remember that b_k is the market beta. However, this time the average return per group r_k is used instead of r_i.

For the 'data driven grouping', the distance function has the following form $\delta(k,k') = \|\tilde{\beta}_k - \tilde{\beta}_{k'}\|$. Here $\tilde{\beta}_k$ is the coefficient vector of a simple HAR regression per group - without JL.

For both ways of grouping, the formulation of $\delta(k,k')$ assumes that when features are similar it is reflected by similarity between the underlying regression coefficients, which is what is tested in this paper. Finally for both groupings, the weight parameter has the following form: $\tau_{k,k'} = 1 - \frac{\delta(k,k')}{\delta_{max}}$, such that the weights are higher for similar groups. Here δ_{max} is the largest distance between subgroup pair k,k' .

4.7 Long Short Term Memory Networks

Finally, the JL restrictions are implemented for a Long Short Term Memory (LSTM) network. A LSTM is an extension of the recurrent neural network (RNN) class that, other than a 'latent' variable h_{t-1} (which contains the output the last period), also keeps track of other latent variables, namely the 'input' (i_t), 'forget' (f_t) and 'output' (o_t) variables. Without these latent variables, normal RNNs have exponential decay in their autocorrelations, so they are often unable to explain high order lags, similar to models such as the GARCH model. The LSTMs are capable to deal with long term memory, similar to the HAR model. The implementation of this LSTM is almost equal to Liu (2019), so details about implementation can be found there. Even so, to give some insights, one memory cell of the LSTM has the following functions:

$$
\begin{aligned}
f_t &= \sigma(W_f[C_{t-1}, h_{t-1}, x_t] + b_f) & C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \\
i_t &= \sigma(W_u[C_{t-1}, h_{t-1}, x_t] + b_i) & \tilde{C}_t &= tanh(W_c[h_{t-1}, x_t] + b_C) \\
o_t &= \sigma(W_o[C_t, h_{t-1}, x_t] + b_o) & h_t &= o_t tanh(C_t)
\end{aligned}
$$

Here x_t is the input vector, in this case consisting of the HAR variables. The state of the output units is given by h_t. C_t is the cell state vector, which contains the memory of the LSTM and the information that will be transformed to h_t, which is used to compute the next state variable C_{t+1}. W and b are the weights and constant. σ is a sigmoid function, which gives outputs between 0 and 1, and $tanh$ a specific sigmoid function that rescales between -1 and 1.

The closer the value i_t is to 1, the more the $\tilde{C}_t$ is passed onto the state variable C_t and memorized in future states. $\tilde{C}_t$ is a 'candidate' state variable. f_t determines whether C_t will be remembered for the next period. If it is close to 1, the unit will be memorized for the next period. Finally, similarly to the other variables, the output variable o_t determines the values of the outputs of the LSTM, h_t. Both h_t and C_t are the memory of the LSTM.

Bucci (2019) assumes one hidden layer throughout their research, which is applied here as well. The author actually recommends to use 50 nodes in this layer to train the LSTM, however, this is different for this LSTM considering how the JL restrictions are applied.

In this paper, the LSTM is trained on the HAR inputs, $RV_t = [RV_t^{(day)}, RV_t^{(week)}, RV_t^{(month)}]$, (without the constant) in order to predict the daily realized volatility the next day $RV_{t+1}^{(daily)}$. To predict the realized volatility the next day of a specific stock i, the HAR inputs of all N stocks are used. However, the weights, W, are restricted in such a way that the inputs of different stocks do not interact until the first layer. Each stock i with $P = 3$ inputs is connected its own Z = 50 nodes. This implies that there is $n \times Z$ nodes in total, with $N \times Z \times P$ weights. Stock i has Z nodes where $W_{ipz} \neq 0$ for $p = 1, \ldots, P$ and $z = 1, \ldots, Z$, these weights are set by the LSTM. For the other $(n-1) \times 50$ nodes $W_{jpz} = 0$, where $j \neq i$, $p = 1, \ldots, P$ and $z = 1, \ldots, Z$. After the first layer, however, all nodes are connected to the same output. The weights to the output have an ℓ_1 and ℓ_2 restriction. In the Appendix, Section A.1, a visual representation of this idea can be found. The LSTM is developed with Keras (2017). This is a neural network API, in Python (1995) running together with TensorFLOW (Chollet (2017)). This model will be evaluated separately from the other JL HAR models, as it has a very different structure.

4.8 Benchmark Models

The performance of the group HAR models with JL restriction are compared with three benchmark models. The first is the equation by equation HAR model. This is the HAR model estimated separately for each stock as in Equation 1. In this paper, the main objective is to outperform this model. Furthermore, the HAR model is applied per group. Similarly to the JL, but this time there is no exchange of information between groups. This is done to check whether it is helpful to share information. These two benchmark models are implemented in Python (1995) using the SKLEARN (2020) package. Finally, in order to check whether grouping is useful, the JL is applied without grouping the stocks (each stock is its own group). Implementation is similar to Section 2.

4.9 Rolling window

All the models described before are estimated for a rolling window with $L = 250, 125, 94, 63$ days, in order to have both large and small data sizes. As stated before, it is tested whether the JL still performs well in small data sets. Hence, in order to compare the performance in small data sets to big data sets, different sizes of L are introduced. The rolling window of size L moves over the data set and for each window makes a one step ahead forecast. The groups, as explained in Section 4.5, are re-estimated every rolling window, implying that the groups can differ over time and only the data of the current rolling window is used to group the stocks. Finally, the expectation is that for the HAR equation by equation the performance will worsen with a decrease of L, but the group HAR models with JL restriction will still perform well.

4.10 Model Evaluation

First of all, the performance is examined by the model confidence set (MCS) introduced by Hansen, Lunde, & Nason (2011). When inspecting multiple models, a relevant question is, "Which model is the best?" Especially when comparing many models, this question can be too specific to answer. Statistical tests often do not yield a single model that dominates all other models at a significant level. However, MCS enables one to reduce the set of models. Given a certain confidence level α, the MCS does a sequence of predictive accuracy tests to define a set of superior models (SSM). To account for correlation in the data, it uses a block-bootstrap method to take random samples from the data to apply the prediction test to. For this paper, the block-bootstraps are of length 23 to account for the fact that in the HAR model, the realized volatility the next day depends on the realized volatility this month. The forecast performance of different models are compared to each other in pairs according to the loss difference. The model with worst performance is discarded. Finally, the set of SSM contains a set of models that outperform all the other models. A script created by Gong (2019) is used to run this test.

The models is evaluated in terms of root mean squared error (RMSE) and quasi-likelihood (QLIKE), as proposed by Patton (2011). These loss functions are used for MCS, Diebold Mariano (introduced later) and evaluated in general. Especially the QLIKE is useful for financial data as it is more robust for large outliers, which is typical for volatility data. The RMSE is a common loss function for both in-sample and out-sample performance, but it is more sensitive to outliers. Differences between these two can lead to interesting conclusions about how outliers affect the performance of the models.

$$RMSE = \sqrt{\sum_{t=1}^{T} \frac{(\hat{RV}_t^{(daily)} - RV_t^{(daily)})^2}{T}} \qquad QLIKE = \sum_{t=1}^{T} \frac{\frac{RV_t^{(daily)}}{\hat{RV}_t^{(daily)}} - \log \frac{RV_t^{(daily)}}{\hat{RV}_t^{(daily)}} - 1}{T}$$

However, the RMSE and QLIKE loss functions give no information about the significance of the forecast performance. For this reason, a Diebold Mariano (DM) test statistic is applied. DM tests the difference between loss functions of two models, with the null hypothesis of a difference equal to zero, implying equal performance of the two models. As it is uncertain whether the forecast errors are normally distributed, the DM formulation of Harvey, Leybourne, & Newbold (1997) is used. This formulation is more robust for such models. Harvey et al. (1997) considers the following modification for one step ahead forecasts and a sample of size T:

$$DM_{harvey} = \sqrt{\frac{T-1}{T}} DM \sim t(T-1).$$

5 Results

5.1 Groups

First, a choice is made between the joint lasso estimation based on the 'industry groups' or 'data driven groups'. The 'industry grouping' is based on the market beta, b_t. Figure 2 shows the b_t over time. Especially the b_t of the BAC seems interesting, $b_{t,BAC}$. Before the crisis in 2008, $b_{t,BAC}$ is relatively low and stable, which implies that the returns of this stocks move less than the market. It moves relatively similar to the KO ($b_{t,KO}$), another relatively stable stock. However, the $b_{t,BAC}$ rises a lot when the crisis starts. It is now close to the other financial stocks, JPM and AXP, implying that these stocks are grouped together. From Figure 1, it is known that these stocks have similar level of volatility during that time, so this way of grouping might be beneficial for these stocks. The groups can thus differ a lot over time and are not necessarily grouped together with stocks from the same industry.

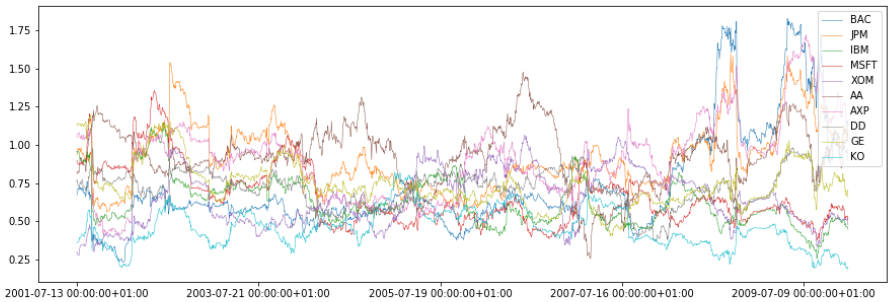
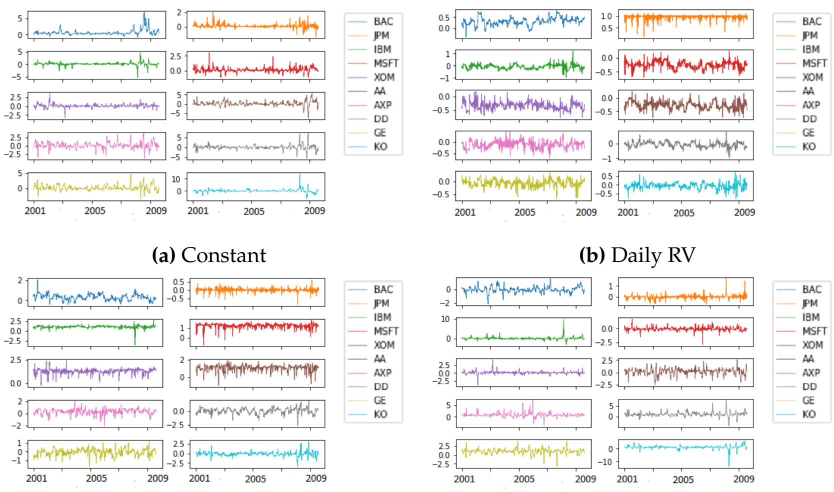


Figure 2: The market betas b_t over time for the different stocks. The market betas were estimated using 5 month historical returns of the market (DJIA) and the specific stocks. b_t lower (higher) than one implies that the stock moves less (more) than the market.



(a) Constant

(b) Daily RV

(c) Weekly RV

(d) Monthly RV

Figure 3: The coefficients of the HAR equation by equation for the different stocks over time, rolling window 63.

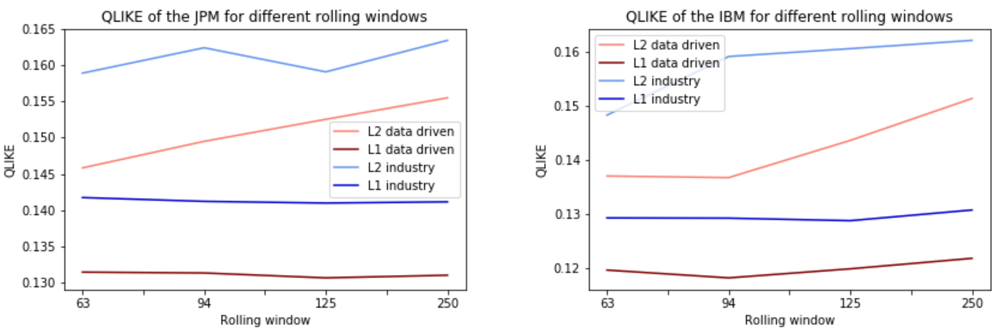
For the 'data driven grouping', stocks are grouped based on the similarity of the coefficients $\tilde{\beta}_k$ when implementing a HAR equation by equation model. Figure 3 shows the coefficients of this model for different stocks over time only for a rolling window of 63, because of a lack of space. Similar plots for the other rolling windows are available upon request.

Different from the 'industry grouping', the BAC, JPM and AXP do not necessarily have comparable coefficients during the crisis. Although their coefficients do seems to fluctuate more during this time. This does not necessarily lead to grouping them together, because that would need the coefficients to fluctuate in the same manner. Thus, Figure 3 shows that stocks from similar industries do not necessarily move the same. If this was the case, BAC and JPM or IBM, MSFT and GE and DD and AA would exhibit similar patterns. This might be an implication that a data driven way of estimation is more suitable for these stocks as the patterns in coefficients cannot be explained from industry alone.

From cross-validation, the optimal number of groups is found to be 5, in both cases. The number of groups is fixed over time, but the number of stocks in each group does change.

The choice between 'data driven' and 'industry grouping' is made based on which grouping method leads to a group HAR model with JL restriction with the most accurate predictions, in terms of RMSE and QLIKE. In Figure 4, the QLIKE is plotted for the two group HAR models with JL restriction over the different rolling windows for two stocks. Due to large similarities between stocks and between the RMSE and QLIKE, plots for the RMSE and the other stocks can be found in the Appendix, Section A.2. For all the stocks, the data driven group HAR model with ℓ_1 restriction has the lowest QLIKE. As can be seen from Figure 4, this conclusion does not depend on rolling window size. The JL ℓ_1 with data driven groups, outperforms the industry groups JL ℓ_1 and ℓ_2 on 5% and 1% level respectively when applying the DM test. This is understandable, because the data driven groups already contain stocks that have similar coefficients when estimated separately, which is useful considering they will all have the same coefficient when estimated as a group. This is not the case for 'industry grouping'.

As stated before for the RMSE similar results are found. However, for AA, KO and XOM, for a rolling window of 250, the 'industry grouping' HAR with ℓ_1 norm has a lower RMSE than 'data driven grouping', but not significantly lower. Besides, even in these cases a lower RMSE is found for the 'data driven groups' at another rolling window size. However, it is good to keep in mind that 'industry grouping' does have potential for larger rolling window sizes.



(a) JPM **(b)** AA

Figure 4: The QLIKE of different stocks over rolling window sizes $L = 63, 94, 125, 250$. Comparing 'industry grouping' with 'data driven grouping'. The light and dark blue (red) lines are the JL RMSE and QLIKE based on 'industry' ('data driven') groups for ℓ_2 and ℓ_1 respectively. The L1 and L2 in the figure are equal to ℓ_1 and ℓ_2 respectively.

From the above results, it seems a fair conclusion that the volatility forecasts of most stocks seems to have a better precision when the stocks are grouped in a data driven way. Therefore, for simplicity, this method is considered for the remainder of this paper and the 'industry grouping' is excluded from further discussion. When referred to ℓ_1 and ℓ_2 or JL, the underlying grouping and/or tuning method is data driven (in case the JL is applied without grouping, only the tuning part is data driven).

5.2 Model Comparison (whole sample)

Here the performance of the group JL HAR models, ℓ_1 and ℓ_2, are compared to the benchmark models, HAR equation by equation, HAR as group and finally the HAR with JL restrictions, ℓ_1 and ℓ_2, but without grouping the stocks. The comparison is based on the model confidence set, loss functions (QLIKE, RMSE) and DM tests.

5.2.1 Model Confidence Set

The predictive performance of the group JL and the benchmark models are compared via MCS. The results are shown in Table 2. On average the models with the highest probability of being included in the SSM, are either the per group or per stock HAR with ℓ_1 fusion, for both loss functions. If the (per group or per stock) HAR with ℓ_1 fusion is not in the SSM with the highest probability, the HAR with ℓ_2 fusion is in the SSM with the highest probability instead. There are no stocks for which the ℓ_1 and ℓ_2 are outperformed by the other benchmark models. Based on these probabilities, the volatility of the stocks BAC, JPM, MSFT, AXP and KO should be predicted as group with ℓ_1 or ℓ_2 fusion. Most of these stocks are the volatile stocks, as described in Section 3. For the other stocks, the volatility should be predicted without grouping, but with fusion between stocks coefficients. Tables for rolling windows of $L = 94, 125, 250$ can be found in the Appendix, Section A.3. From the analysis of these rolling windows, similar findings can be drawn. Besides, from these tables, it becomes clear that with a decrease in sample size, it is more useful to use a JL ℓ_1 and ℓ_2 restriction when estimating the HAR model.

Table 2: Model Confidence Set (MCS), $\alpha = 0.10$, rolling window of 63

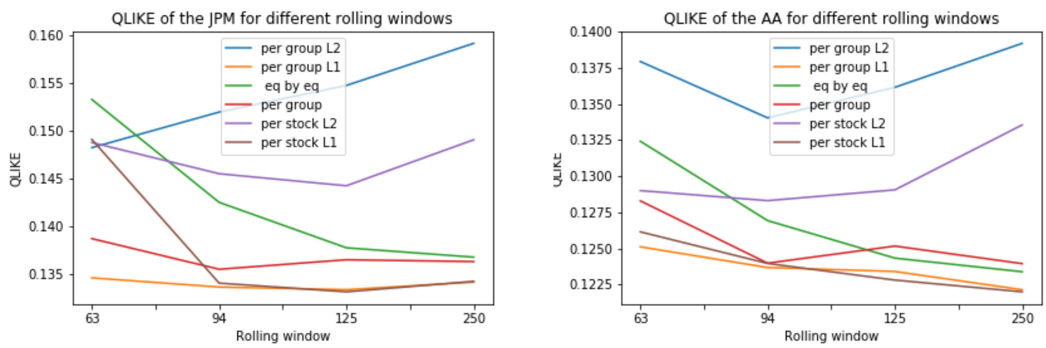
	BAC	JPM	IBM	MSFT	XOM	AA	AXP	DD	GE	KO
QLIKE										
per group L2	0.174*	0.130*	0.217*	0.130*	0.348**	0.130*	**1.000****	0.0	0.043	**1.000****
per group L1	**1.000****	**1.000****	0.087	**1.000****	0.739**	0.348**	0.130*	0.0	0.304**	0.000
eq by eq	0.174*	0.130*	0.217*	0.435**	0.174*	0.217*	0.000	0.0	0.130*	0.391**
per group	0.609**	0.130*	0.304*	0.565**	0.304**	0.087	0.000	0.0	0.130*	0.000
per stock L2	0.174*	0.174*	0.130*	0.565**	**1.000****	0.174*	0.000	**1.0****	0.130*	0.087
per stock L1	0.174*	0.000	**1.000****	0.043	0.739**	**1.000****	0.348**	0.0	**1.000****	0.043
MSE										
per group L2	0.217*	0.087	0.435**	0.130*	0.130*	0.043	**1.000****	0.0	0.000	**1.000****
per group L1	**1.000****	**1.000****	0.261**	**1.000****	0.391**	0.435**	0.304**	0.0	0.304**	0.000
eq by eq	0.304**	0.087	0.435**	0.391**	0.087	0.217**	0.000	0.0	0.043	0.435**
per group	0.435**	0.087	0.435**	0.565**	0.087	0.000	0.217*	0.0	0.130*	0.000
per stock L2	0.348**	0.217*	0.348**	0.652**	**1.000****	0.174*	0.000	**1.0****	0.130*	0.087
per stock L1	0.087	0.000	**1.000****	0.217*	0.522**	**1.000****	0.522**	0.0	1.000**	0.000

The forecasts with MCS p-value larger than 0.1 and 0.25 are indicated by * and ** respectively.

5.2.2 Loss Functions and Diebold Mariano

Now, the predictive accuracy of the different models is examined based on their loss functions and DM tests (as a robustness check of the MCS). The QLIKE and RMSE plots can again be found in the Appendix, Section A.4. The DM tests can also be found there. A notable result is that in terms of MSE, most of the time, no model seems to outperform the others significantly. This is easy to explain by the nature of the MSE function. In general, volatility processes have small number of very large observations, which usually dominates the whole process. The implication of this is that loss functions, such as the MSE usually fail to reject the null hypotheses by the non-robust formulation of this function (in terms of large outliers).

The QLIKE is more robust in this case, which is why this measure shows more potential in terms of significance. In Figure 5, again the QLIKE for only two stocks is plotted. Comparing the prediction performance of the ℓ_1 and ℓ_2, it seems that in general ℓ_1 performed the best for both per group HAR and per stock HAR. The ℓ_2 is outperformed by the ℓ_1 at 1% level for each rolling window. A reason for this could be that ℓ_1 allows different groups to have the same coefficients, while ℓ_2 does not allow for this. For per group JL HAR this could be an indication that 5 groups might be too many.



(a) JPM

(b) AA

Figure 5: The QLIKE of different stocks (JPM left, AA right) over rolling window sizes $L = 63, 94, 125, 250$. Comparing the JL HAR model with the benchmark models over the whole sample period. The L1 and L2 in the figure are equal to ℓ_1 and ℓ_2 respectively.

Nevertheless, the ℓ_2 still seems to have some potential, as the QLIKE seems to decrease with a reduction in the size of the rolling window, while the HAR equation by equation, has an increasing QLIKE as the rolling window gets smaller. Therefore, the performance of the ℓ_2 fusion might actually perform better than the HAR equation by equation for even smaller rolling windows. This already happens for the JPM stock at a rolling window of 63 (the HAR equation by equation and per group/per stock ℓ_2 intersect between $L = 94$ and 63) as can be seen from Figure 5a. However, the performance is not significantly better. Overall, in this setting, the HAR with ℓ_2 fusion is not a beneficial estimation method.

Compared to the other models, the QLIKE of the HAR with ℓ_1 fusion stays relatively stable over the rolling windows (again for both group ℓ_1 and single stock ℓ_1). Because the QLIKE of the HAR equation by equation gets larger with a decrease in rolling window size, the performance of the HAR with ℓ_1 fusion gets more significant compared to HAR equation by equation with a decrease in rolling window size. The ℓ_1 usually is significant on 1% level for rolling window sizes of $L = 63, 94$ and either 5% or 10% for the bigger sizes. More importantly, the group ℓ_1 HAR for smaller rolling window sizes outperforms the HAR equation by equation for bigger rolling window sizes. For instance, the ℓ_1 HAR based on a rolling window of $L = 63$ often outperforms the HAR equation by equation based on $L = 250$ on 5% significance level. This implies that one could reduce the sample size without losing precision when shifting from HAR equation by equation to group HAR with ℓ_1 fusion. There is no significant difference between the per group ℓ_1 and per stock ℓ_1. Often the group ℓ_1 is closer to significantly outperforming the stock ℓ_1. There are two occasions where the group ℓ_1 significantly outperforms per stock ℓ_1, which is for the GE and KO stock and a rolling window of 125 and/or 250.

Furthermore, the HAR per group without fusion has a worse performance than the HAR per group with ℓ_1 fusion. The difference between these two models gets larger as the rolling window size decreases. This implies that it is useful to take into account the coefficients of groups that are similar to the currently estimated group. This is especially important when the rolling window size gets smaller.

In terms of overall performance, the ℓ_1 fusion seems to be a profitable coefficient estimation method. The per group and per stock HAR with ℓ_1 fusion, outperforms all the other models most of the time. The rolling window for which the JL ℓ_1 performed the best changes per stock, but does seem to be either $L = 94$ (based on QLIKE) or $L = 125$ (based on RMSE) on average. Concluding that using a JL ℓ_1 restriction is useful, especially for smaller rolling windows.

Finally, these results are similar to what is seen for the MCS in Section 5.2.1. The JL models with ℓ_1 fusion do indeed seem to outperform the other models most of the time. However, what stands out is that based on the MCS, for the AXP and DD (and rolling window of 63) the ℓ_2 fusions should have outperformed the other models, which is not the case. In fact, the group ℓ_2 is even outperformed by all the other models for DD. This contradiction could be due to three reasons. First, the random component of the MCS somehow always bootstraps the best performing observations for ℓ_2 (highly unlikely). Secondly, the autocorrelation of the volatility is not captured correctly by the block bootstraps of length 23. Lastly, it could be that there are a few observations that really twist the performance of the two ℓ_2 models, and these observations are often not sampled in the bootstrap. In Section 5.3, it is shown that the latter is indeed the case and the performance of each model is time dependent.

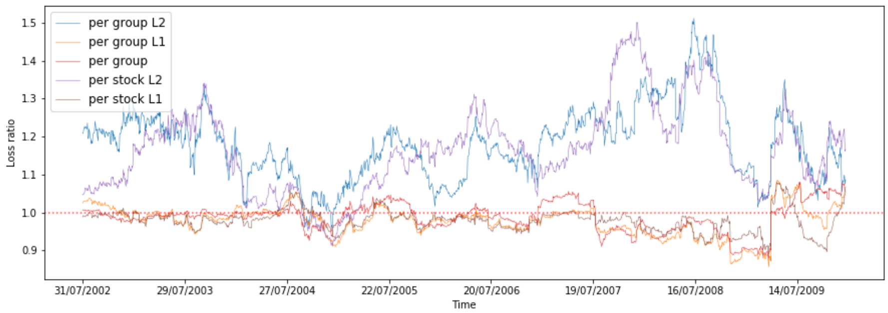
5.3 Moving Window

The fact that the QLIKE loss function leads to different conclusions than the RMSE loss function, might imply that some large outliers significantly affect the overall performance of the models. The QLIKE is more robust for large outliers as it standardizes the errors. The MSE does not do this and performs relatively well when the volatility process is not dominated by a few large observations. Besides, the MCS shows some contradicting results, which could also be explained by outliers. It would thus be interesting to see how the loss functions behave over time.

Based on a moving window of 100, the QLIKE of the different models is computed over time. To make the results more accessible the loss ratio is plotted: $\frac{QLIKE_j}{QLIKE_i}$, comparing model j

to benchmark model i. If the loss ratio is smaller than 1 it means model j outperforms benchmark i. The QLIKE of the HAR equation by equation model is taken as benchmark and the relative performance of the other models is measured: per group ℓ_2, per group ℓ_1, per group (no fusion), per stock ℓ_2 or per stock ℓ_1 HAR. In Figure 6 the results are shown. For simplicity, only $L = 63, 250$ were plotted in Figure 6a and Figure 6b respectively. Again, the QLIKE ratios for the other stocks and the RMSE ratios were also plotted, but led to similar figures, so for simplicity, only the the MSFT is taken as example and only the QLIKE is shown here.

(a) $L = 63$



(b) $L = 250$

Figure 6: Loss ratio: $\frac{QLIKE_j}{QLIKE_i}$ of the MSFT over time for a moving window of 100 to compute the QLIKE for the different models. Here i is the HAR equation by equation model. j is per group ℓ_2, per group ℓ_1, per group (no fusion), per stock ℓ_2 or per stock ℓ_1 HAR. If the loss ratio level is under the red dotted line with value 1, it means model j outperforms the HAR equation by equation model.

Two periods can be distinguished from these figures. Looking at Figure 6a, in general the ℓ_1 models have the lowest QLIKE relative to the HAR equation by equation model. Turning to Figure 6b ($L = 250$), the HAR equation by equation performs relatively well in the period before 2007, because the loss ratio comparing to per group ℓ_1 or per stock ℓ_1 is close to zero. However, after 2007, QLIKE ratio becomes a lot smaller. Thus, the per group or per stock ℓ_1 again seems to be the most accurate. This could be an indication that during crisis, jointly predicting and sharing information between the volatility of different stocks is useful. Hence, the HAR equation by equation performs relatively well before 2007 (based on $L = 250$) and the JL HAR exhibits its best performance after 2007.

The HAR models with ℓ_2 restriction, have their best performance between 2003 and 2007 as well. This is mostly the case for the per stock ℓ_2 based on a rolling window of 63. The MCS showed that especially for the DD stock, this model is preferable (Table 2). A plot similar to

Figure 6 indeed shows that the QLIKE ratio (comparing per stock ℓ_2 and HAR equation by equation) is often below 1, with just a few large outliers especially after 2007. This result can be found in the Appendix, Section A.5. This might explain why only this model is in the SSM, but does not have an overall good performance.

The purpose of this paper is to reduce the data needed to make accurate predictions by using the JL HAR instead of estimating the HAR equation by equation. In fact, it might be even more practical if the JL HAR is still able to outperform the HAR equation by equation model if it is estimated with a smaller sample than the HAR equation by equation. For this purpose, both the per group and per stock ℓ_1 models performance based on a rolling window of size 63 is compared to the performance of the HAR equation by equation for a rolling window of 250, again using the QLIKE ratio. The results are plotted in Figure 7.

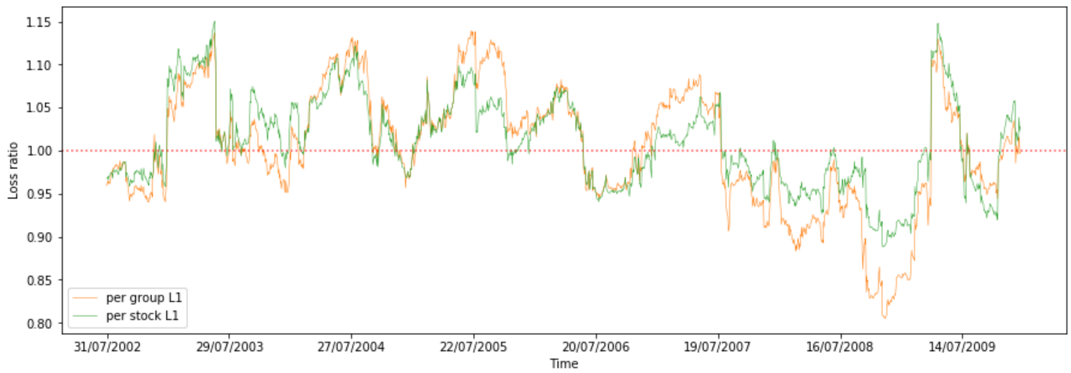


Figure 7: Loss ratio over time using a moving window of 100 to compute the loss functions. Comparing per group and per stock ℓ_1 ($L = 63$) to HAR equation by equation ($L = 250$). The L1 and L2 in the figure are equal to ℓ_1 and ℓ_2 respectively. The stock considered here is the MSFT.

The two periods can again be clearly distinguished. Figure 7 shows that the ℓ_1 JL HAR models gain a lot in accuracy after 2007, while before this period the HAR equation by equation model has the best accuracy.

In Figure 8, the DM statistic of comparing group HAR with ℓ_1 fusion for rolling window of size 63 with HAR equation by equation for rolling window of 250 over time is indicated with a blue line. The dotted red lines indicate the significance levels at 5% (two-sided) and crossing the highest red line means significantly outperforming HAR equation by equation and crossing the lowest red line implies significantly underperforming HAR equation by equation. The group ℓ_1 is only able to outperform the HAR equation by equation for a short period in 2008. It shows the highest statistics during the crisis. The pink lines indicated the DM statistic over time comparing per stock fused ℓ_1 to the HAR equation by equation. This leads to a similar conclusion as for group ℓ_1.

Finally the orange line indicates the DM statistic of comparing the group ℓ_1 to the per stock ℓ_1. During the crisis, this statistic is low, implying that the group ℓ_1 performs better. Before 2007, this statistic is especially high in the advantage of the per stock ℓ_1. There are only two points in 2005 and 2006, where the per stock ℓ_1 significantly outperforms the per group ℓ_1.

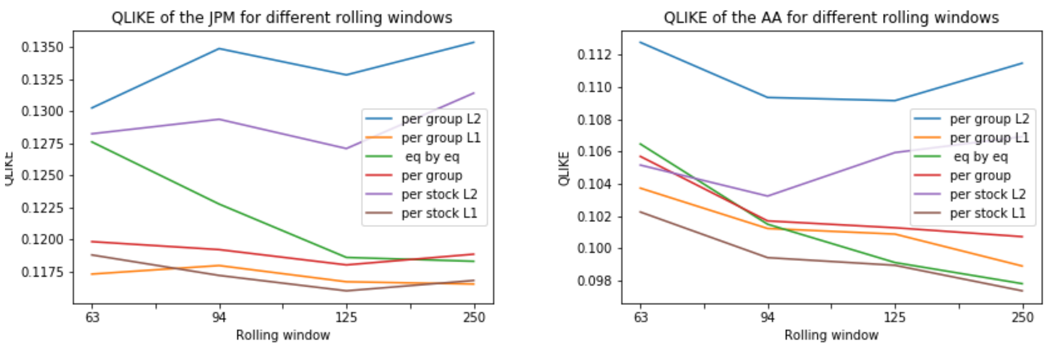
Hence, this figure shows that during the crisis, grouping stocks has potential, while in times of no financial distress (before 2007), HAR equation by equation or per stock ℓ_1 has the most potential.

Figure 8: Diebold Mariano statistic comparing per group and per stock ℓ_1 for a rolling window of 63 to HAR equation by equation for rolling window of 250 (MSFT). The L1 and L2 in the figure are equal to ℓ_1 and ℓ_2 respectively.

In the next two sections, the behaviour of the models is examined more thoroughly during these periods: 2003-2006 and 2007-2009.

5.4 Model Comparison (2003-2006)

Section 5.3 showed different performance of the models in the period of 2003-2006. Therefore, this period is examined more thoroughly. In Figure 9, QLIKE for the different models is plotted, again only for two stocks. The RMSE and plots for other stocks are available upon request. Comparing these models to the overall period, as plotted in Figure 5, the performance for the JPM seems relatively similar. However, for the AA, the per stock HAR model with ℓ_1 fusion now outperforms the per group HAR model with ℓ_1 fusion. In fact, this is significant at 5% level based on a DM statistic. The HAR equation by equation also performs better than the per group ℓ_1 but not significantly. This behaviour is observed for most stocks. Although, sometimes the per group HAR ℓ_1 still has a lower QLIKE than the HAR equation by equation, but not significantly lower. For these stocks, the higher rolling window sizes usually have the lowest loss functions. Only for the AXP and IBM the overall best performing model is the per group ℓ_1 for $L = 94$, which significantly outperforms the HAR equation by equation for $L = 250$ (1% level).



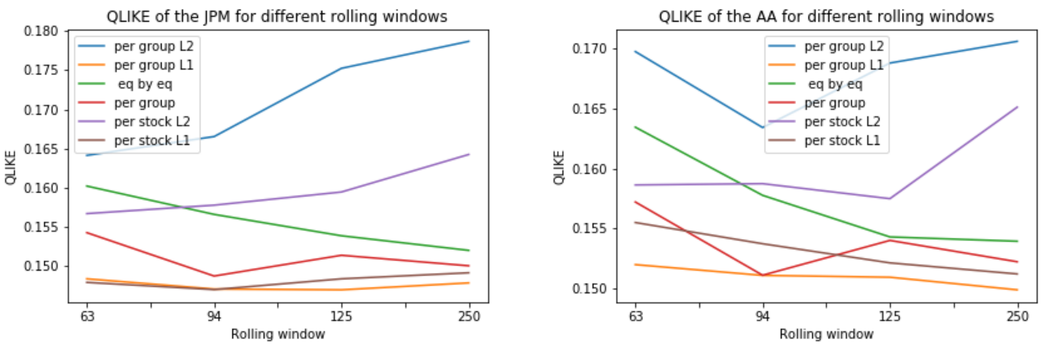
(a) JPM **(b)** AA

Figure 9: The QLIKE of different stocks (JPM left, AA right) over rolling window sizes $L = 63, 94, 125, 250$. Comparing the JL HAR model with the benchmark models 2003-2006. The L1 and L2 in the figure are equal to ℓ_1 and ℓ_2 respectively.

The former indicates two things. First, in times of relatively stable volatility, it is better to use larger rolling windows and thus more data. This is an understandable result, because there are no big observations that distort the volatility forecasts (for larger rolling window sizes these often have long lasting effects) and the volatility coefficients are stable over time (so larger rolling windows lead to more precision). Secondly, in times of low volatility, estimating per stock (whether that is with ℓ_1 fusion or without) leads to more accurate forecasts.

5.5 Model Comparison (2007-2009)

Section 5.3 indicates that during the crisis the group HAR with ℓ_1 restriction should especially perform well. Thus, for the period of 2007 until 2009, the results are shown in Figure 10, again only the QLIKE for only three stocks. The RMSE and plots for other stocks are available upon request. Different from what is observed in Figure 9, the two ℓ_1 models now significantly outperform the HAR equation by equation model. This is also the case for the AA stocks, for which the HAR equation by equation actually outperforms the group ℓ_1 in the period from 2003-2006. For the other stocks, the group ℓ_1 also outperforms the other models. The best rolling window size depends on the stock, but is usually around 94 or 125. For a few stocks, such as the BAC, the group ℓ_1 based on a smaller rolling window (for the BAC, $L = 63$) significantly outperforms the HAR equation by equation for $L = 250$, on 10% or even 5% level. The group HAR without fusion, does actually perform quite good during this period, however even in case it performs well, it is never able to outperform the group HAR with ℓ_1 fusion.



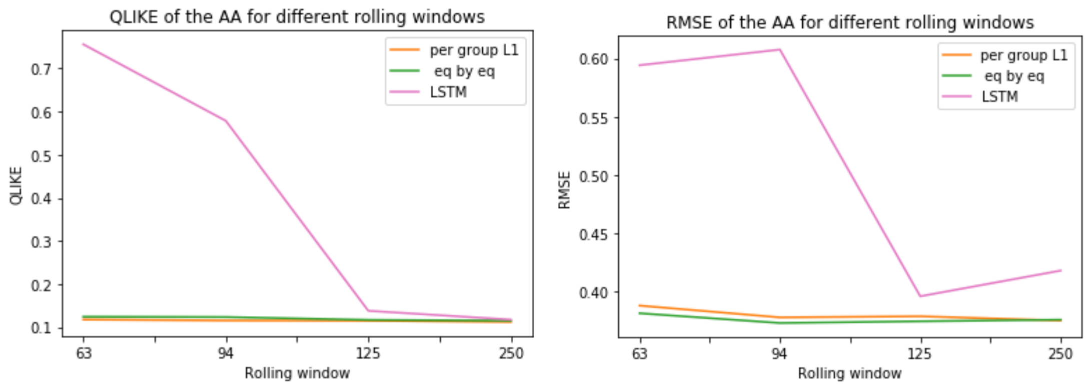
(a) JPM **(b)** AA

Figure 10: The QLIKE of different stocks (JPM left, AA right) over rolling window sizes $L = 63, 94, 125, 250$. Comparing the JL HAR model with the benchmark models in the period 2007-2009. The L1 and L2 in the figure are equal to ℓ_1 and ℓ_2 respectively.

These results indicate that it is useful to jointly determine the coefficients of the HAR model using ℓ_1 fusion by grouping the stocks, during volatile times. A reason for this could be that there are a few stocks that react relatively early to financial shocks, while some other stocks react later. If these two kinds of stocks are grouped together, it is beneficial for the late reacting stocks, because the model is already fitted for future shocks. In addition, it is beneficial to use smaller rolling windows in these times. In this way, the coefficient estimates are less sensitive to outliers from the past. In general, the group HAR with ℓ_1 restriction already outperforms the HAR equation by equation model for smaller rolling windows. As it is especially useful during times of high volatility to use small rolling windows, the ℓ_1 fusion gains a lot in accuracy.

5.6 Long Short Term Memory Network

Finally, an LSTM is estimated with a JL restriction, to capture possible non-linearities in the data. The training of a neural network can be computationally demanding, depending on the number of layers and nodes. As of right now, LSTM models are the most demanding and require efficient computers with high RAM as Liu (2019) states. This is why for this section, only one stock's volatility is trained using an LSTM. This stock is AA. Just like the models discussed before, the LSTM is estimated four times using a rolling window of $L = 63, 94, 125, 250$. 70% of the data is used for training, this leaves 229 observations for evaluation. For all the rolling window sizes, the LSTM failed to converge completely. Nevertheless, the LSTM is compared to the HAR with ℓ_1 and HAR equation by equation as plotted in Figure 11.



(a) QLIKE

(b) RMSE

Figure 11: The QLIKE and RMSE for different rolling windows.

As can be seen from the figure, the LSTM is definitely outmatched by the other models based on the smaller rolling window sizes. Only for the biggest size, $L = 250$ it shows potential. The LSTM's QLIKE (0.118) especially gets close to the other two models (ℓ_1: 0.112 and HAR equation by equation: 0.115). In fact the other two models do not significantly outperform this model for this rolling window. For $L = 63, 94, 125$ the LSTM is significantly outperformed by the other models based on QLIKE, and for $L = 63, 94$ based on RMSE on either 5% and 1% level. This result is not unexpected, because as is shown by Bucci (2019), the LSTM required as many as 10 or even 20 years of data to be trained. Besides, for both Bucci (2019) and Liu (2019) the LSTM is not estimated using a rolling window. Thus, the JL LSTM might have more potential for larger rolling windows, or no rolling window at all.

6 Conclusion

In this paper the joint lasso is applied to stock volatility data. In particular, an effort is made to outperform the standard HAR-RV model by grouping and introducing an ℓ_1 and ℓ_2 restriction on the coefficients between groups. This led to the following question: "To what extent it is useful to jointly determine the volatility structure of stocks using a joint lasso restriction?".

First, a choice is made between grouping the stocks by industry or in a data driven way. The 'data driven' grouping has the best accuracy based on a Diebold Mariano test. This result is not unexpected, because the data driven way groups the stocks according to their similarity in coefficients, after which the groups will have the same coefficients. However, the 'industry grouping' has potential for greater rolling windows.

Secondly, the performance of the 'data driven' group HAR with JL restriction is compared to the benchmark models using MCS. This measure shows that HAR with JL (both ℓ_1 and ℓ_2 restrictions) outperforms the HAR equation by equation model, especially for small data sizes. The DM test confirms this result for ℓ_1, but not for ℓ_2. Due to these contradicting results, the performance of the models is compared over time. This shows that the HAR equation by equation model or per stock HAR with ℓ_1 restriction performs the best in times of low and stable volatility for large data sizes, while the per group HAR with ℓ_1 restriction performs the best for times of high and unstable volatility for small data sizes. In fact, during the crisis, the JL HAR models estimated using small data sizes are even able to significantly outperform the HAR equation by equation model estimated with more data. The reason behind this, is that the volatility is unstable during crisis, such that it is more beneficial to use less data to be less sensitive to outliers from the past. In addition, the JL models are able to make more accurate predictions when estimating with less data. For non-crisis years, the volatility is relatively stable, so more data can be used to estimate the models, making the HAR equation by equation model more beneficial to use.

Finally, an LSTM is trained on the HAR inputs, using a joint lasso restriction. This model did not perform well for small data sizes. Nevertheless, it does show potential for larger moving windows. This is not unexpected, as Bucci (2019) shows that for training an LSTM up to twenty years of data is needed.

The results in this paper are useful to financial institutions, as briefly mentioned in Section 1. For example, reliable volatility models can be used to set up efficient portfolios and make accurate option price predictions, which is useful for hedging risk.

In the future, this research could be repeated for many more stocks, because ten stocks might not be sufficient to form groups. Since the stocks used are the most liquid stocks from the DJIA, these stocks are already similar in some aspect. This might have caused the group HAR with ℓ_1 norm to perform this well. It would be interesting to see how the joint lasso performs for many randomly selected stocks. The number of groups could also be made to change over time and decided by means of some function, such as information loss measures - in this paper the groups were fixed over time and decided by cross validation.

Additionally, there is only one data set on realized volatility available, Oxford-Man Institute of Quantitative Finance (2009). This data set already contained measures of realized volatility, and did not allow the user to set the realized volatility themselves. In the future, it might be useful to also consider measures of realized volatility that are robust for rare jumps in intra-day data, such as is done by Barndorff-Nielsen & Shephard (2004) (power and bipower variation). Finally, LSTMs have been proven to be useful for modelling financial data in the past. The results in this paper might therefore have been caused inefficiency in terms of sample size or just general tuning. It would definitely be worthwhile to look into efficient ways to implement the joint lasso into such networks in the future.

References

Audrino, F., & Knaus, S. D. (2016). Lassoing the har model: A model selection perspective on realized volatility dynamics. *Econometric Reviews*, *35*(8-10), 1485–1521.

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

Bollerslev, T. (1986a). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Bollerslev, T. (1986b). Glossary to arch (garch. In *in volatility and time series econometrics essays in honor of robert engle. markwatson, tim bollerslev and je rey.*

Bucci, A. (2019). Realized volatility forecasting with neural networks.

Čech, 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.

Chollet, F. (2017). Keras: The python deep learning library.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues.

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

Dondelinger, F., Mukherjee, S., & Alzheimer's Disease Neuroimaging Initiative. (2020). The joint lasso: high-dimensional regression for group structured data. *Biostatistics*, *21*(2), 219–235.

Dondelinger, F., & Wilkinson, O. (2017). Fuser: Fused lasso for high-dimensional regression over groups.
(`https://cran.r-project.org/web/packages/fuser/`)

Drummen, M., & Zimmermann, H. (1992). The structure of european stock returns. *Financial Analysts Journal*, *48*(4), 15–26.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Gautier, L. (2020). Python interface to the r language (embedded r).
(`https://pypi.org/project/rpy2/`)

Gong, M. (2019). Model confidence set python implementation.
(`https://michael-gong.com/blogs/model-confidence-set/`)

Gopikrishnan, P., Plerou, V., Liu, Y., Amaral, L. N., Gabaix, X., & Stanley, H. E. (2000). Scaling and correlation in financial time series. *Physica A: Statistical Mechanics and its Applications*, *287*(3-4), 362–373.

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

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, *13*(2), 281–291.

Heij, C., de Boer, P., Franses, P. H., Kloek, T., & van Dijk, H. K. (2004). *Econometric methods with applications in business and economics*. Oxford University Press Inc., New York.

Heston, S. L., & Rouwenhorst, K. G. (1995). Industry and country effects in international stock returns. *Journal of Portfolio Management*, *21*, 53–53.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735–1780.

Karolyi, G. A. (1995). A multivariate garch model of international transmissions of stock returns and volatility: The case of the united states and canada. *Journal of Business & Economic Statistics*, *13*(1), 11–25.

Liu, Y. (2019). Novel volatility forecasting using deep learning–long short term memory recurrent neural networks. *Expert Systems with Applications*, *132*, 99–109.

Noureldin, D., Shephard, N., & Sheppard, K. (2012). Multivariate high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics*, *27*(6), 907–933.

Oxford-Man Institute of Quantitative Finance. (2009). Realized library. (Available at `https://realized.oxford-man.ox.ac.uk/`. Accessed on 13th of May 2020)

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*(1), 246–256.

R Core Team. (2014). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org/`

Rosenow, B., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003). Dynamics of cross-correlations in the stock market. *Physica A: Statistical Mechanics and its Applications*, *324*(1-2), 241–246.

Scikit-learn machine learning in python. (2020). (`https://scikit-learn.org/stable/`)

Van Rossum, G., & Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.

A Appendix

A.1 LSTM Representation

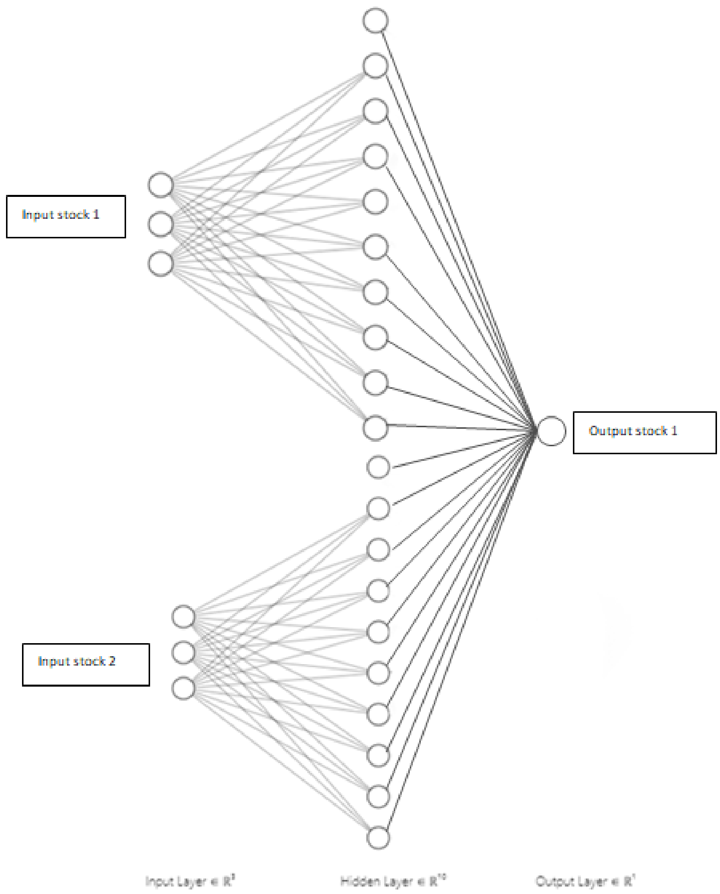
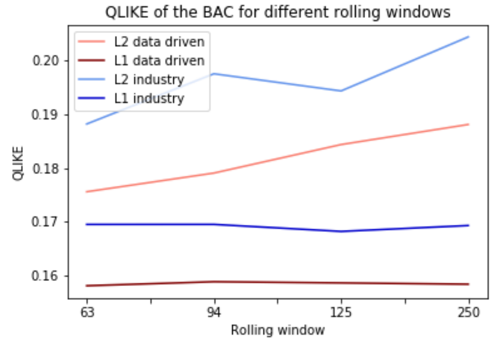
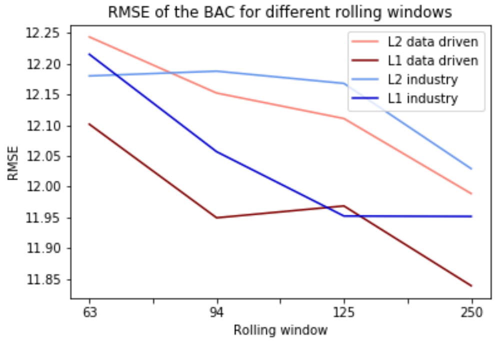


Figure 12: Visual representation of the LSTM for two stocks, $i = 1, 2$, with each three inputs, $p = 1, 2, 3$ and each 9 nodes, $z = 1, ..., 9$

A.2 Industry versus Data Driven Grouping

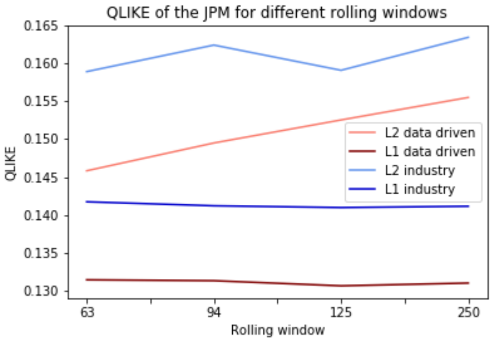


(a) QLIKE

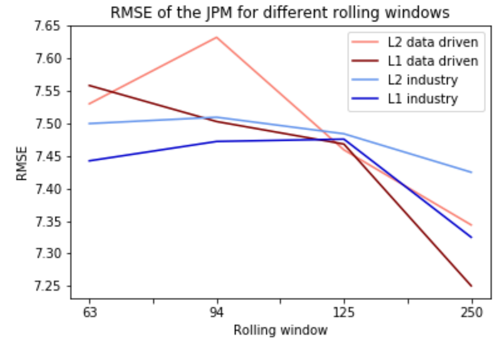


(b) RMSE

Figure 13: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the BAC stock).

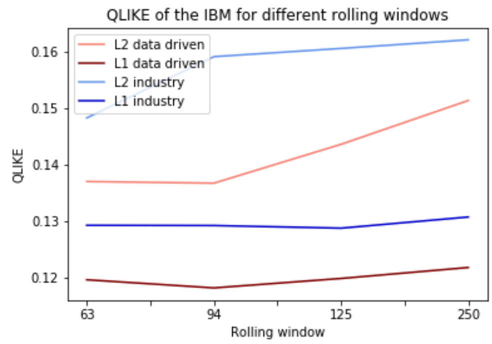


(a) QLIKE

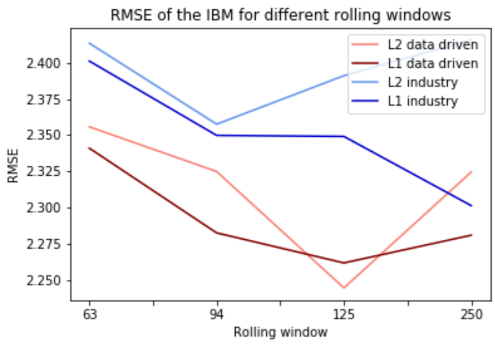


(b) RMSE

Figure 14: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the JPM stock).

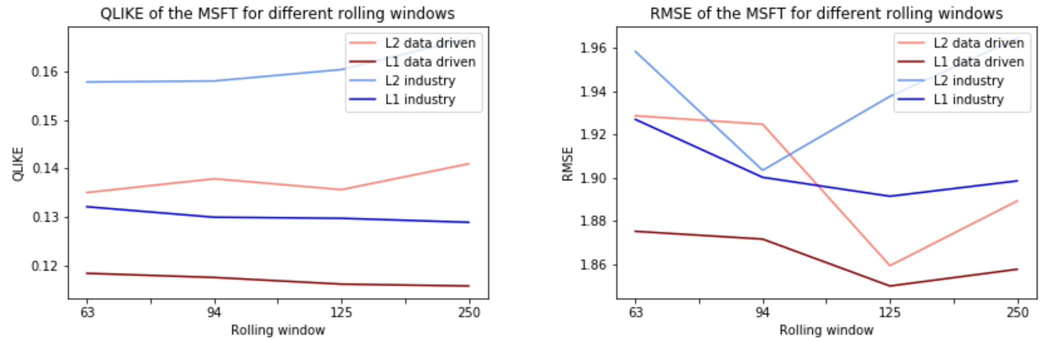


(a) QLIKE



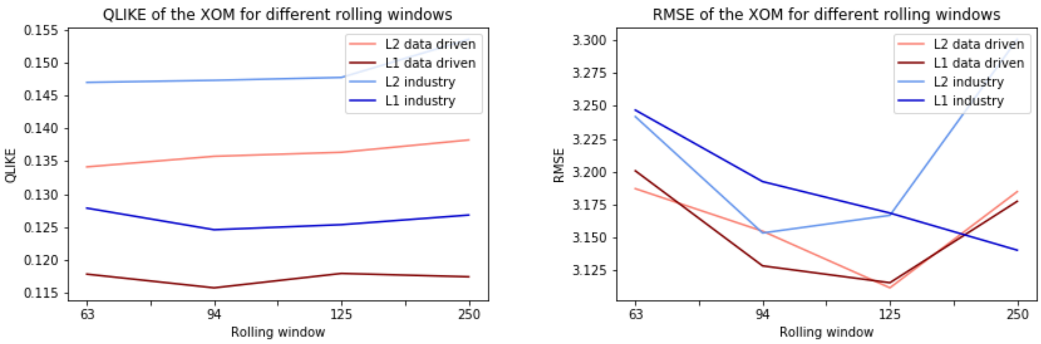
(b) RMSE

Figure 15: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the IBM stock).



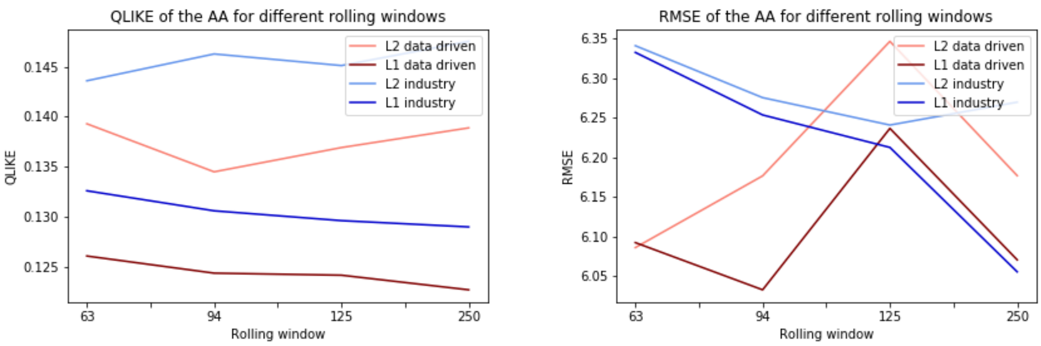
(a) QLIKE **(b)** RMSE

Figure 16: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the MSFTstock).



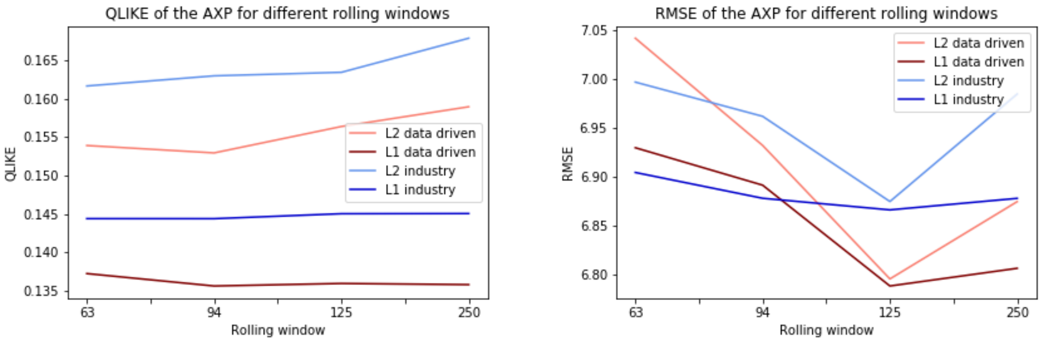
(a) QLIKE **(b)** RMSE

Figure 17: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the XOM stock).



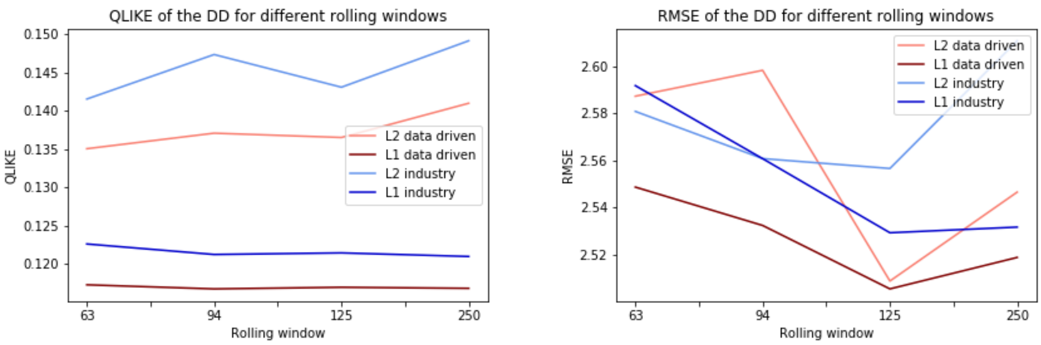
(a) QLIKE **(b)** RMSE

Figure 18: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the AA stock).



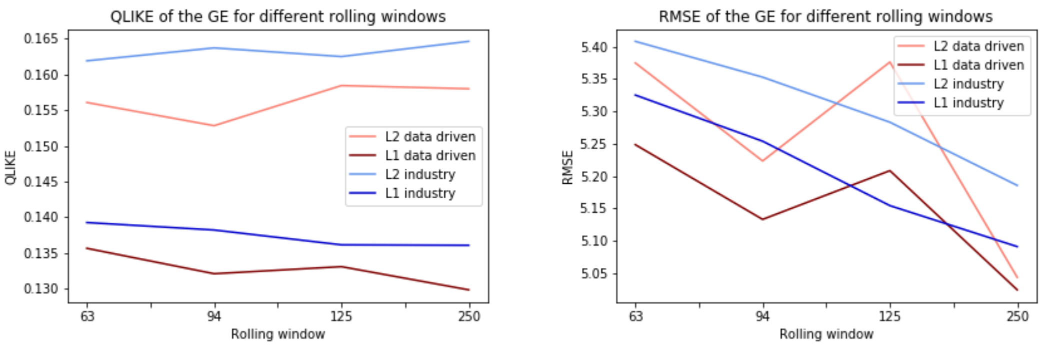
(a) QLIKE **(b)** RMSE

Figure 19: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the AXP stock).



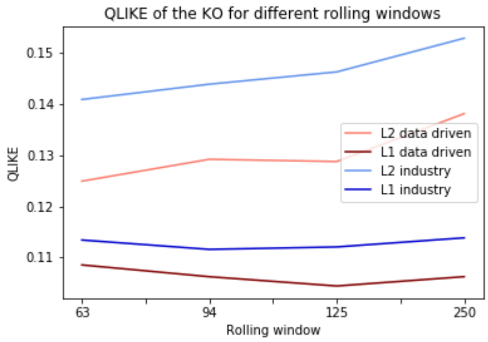
(a) QLIKE **(b)** RMSE

Figure 20: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the DD stock).

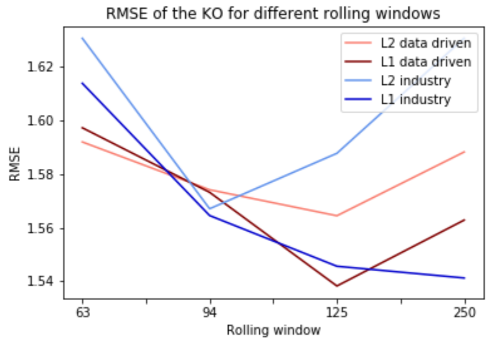


(a) QLIKE **(b)** RMSE

Figure 21: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the GE stock).



(a) QLIKE



(b) RMSE

Figure 22: Loss functions of the group HAR with JL restrictions with industry or data driven groups for different rolling windows (for the GE stock).

A.3 Model Confidence Set (MCS) (whole sample)

Table 3: MCS $\alpha = 0.10$ Rolling window of 94

	BAC	JPM	IBM	MSFT	XOM	AA	AXP	DD	GE	KO
QLIKE										
per group L2	0.043	0.130*	0.043	0.087	0.348**	0.652**	0.609**	0.522**	0.000	0.174*
per group L1	1.000**	0.130*	1.000**	0.043	1.000**	1.000**	0.087	0.348**	1.000**	0.174*
eq by eq	0.043	0.261**	0.087	1.000**	0.130*	0.043	0.609**	0.522**	0.043	0.174*
per group	0.087	0.087	0.087	0.087	0.130*	0.435**	1.000**	0.522**	0.565**	0.174*
per stock L2	0.087	0.174*	0.000	0.043	0.348**	0.652**	0.739**	1.000**	0.696**	1.000**
per stock L1	0.609**	1.000**	0.087	0.087	0.174*	0.652**	0.087	0.522**	0.000	0.174*
RMSE										
per group L2	0.043	0.174*	0.087	0.130*	0.304**	0.130*	0.565**	0.565**	0.087	0.087
per group L1	1.000**	0.174*	1.000**	0.087	1.000**	1.000**	0.087	0.391**	1.000**	0.087
eq by eq	0.043	0.261**	0.217*	1.000**	0.043	0.000	0.478**	0.565**	0.130*	0.087
per group	0.043	0.130*	0.174*	0.130*	0.043	0.130*	1.000**	0.391**	0.391**	0.087
per stock L2	0.304**	0.174*	0.087	0.130*	0.348**	0.130*	0.609**	1.000**	0.435**	1.000**
per stock L1	0.609**	1.000**	0.174*	0.304**	0.174*	0.130*	0.087	0.565*	0.043	0.087

The forecasts with MCS p-value larger than 0.1 and 0.25 are indicated by * and ** respectively.

Table 4: MCS $\alpha = 0.10$ Rolling window of 125

	BAC	JPM	IBM	MSFT	XOM	AA	AXP	DD	GE	KO
QLIKE										
per group L2	1.000**	0.087	0.130*	0.391**	0.130*	1.000**	0.043	0.261**	0.000	0.174*
per group L1	0.043	1.000**	0.130*	0.391**	0.130*	0.348**	0.043	0.391	0.000	0.174*
eq by eq	0.087	0.304**	0.130*	0.000	0.130*	0.348**	0.043	0.522**	0.043	1.000**
per group	0.130*	0.087	0.130*	1.000**	0.130*	0.348**	0.043	0.261**	1.000**	0.174*
per stock L2	0.043	0.087	1.000**	0.000	1.000**	0.043	1.000**	1.000**	0.000	0.391**
per stock L1	0.130*	0.435**	0.217*	0.043	0.696**	0.348**	0.304**	0.522**	0.043	0.174*
RMSE										
per group L2	1.000**	0.174*	0.304**	0.391**	0.043	1.000**	0.087	0.217*	0.043	0.043
per group L1	0.087	1.000**	0.304**	0.739**	0.043	0.348**	0.087	0.261**	0.043	0.043
eq by eq	0.130*	0.304**	0.261**	0.043	0.043	0.130*	0.087	0.522**	0.043	1.000**
per group	0.130*	0.130*	0.304**	1.000**	0.043	0.130*	0.087	0.217*	1.000**	0.043
per stock L2	0.087	0.130*	1.000**	0.043	1.000**	0.087	1.000**	1.000**	0.043	0.435**
per stock L1	0.217*	0.348**	0.435**	0.087	0.652**	0.087	0.348**	0.522**	0.043	0.043

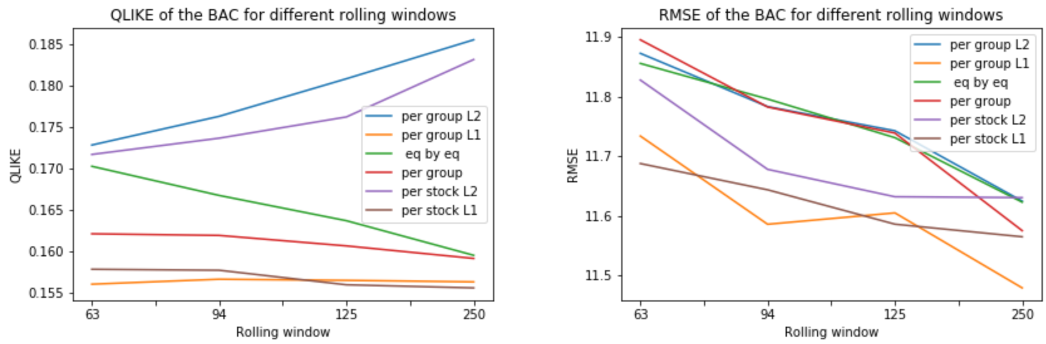
The forecasts with MCS p-value larger than 0.1 and 0.25 are indicated by * and ** respectively.

Table 5: MCS $\alpha = 0.10$ Rolling window of 250

	BAC	JPM	IBM	MSFT	XOM	AA	AXP	DD	GE	KO
QLIKE										
per group L2	0.087	0.087	0.130	1.000**	0.304**	0.087	1.000**	0.130*	0.087	0.348**
per group L1	0.087	1.000**	0.087	0.348**	1.000**	0.087	0.348**	0.087	0.087	0.609**
eq by eq	0.087	0.087	0.130*	0.348**	0.217*	0.087	0.130*	1.000**	0.130*	1.000**
per group	1.000**	0.087	0.130*	0.348**	0.304**	0.087	0.130*	0.000	0.130*	0.522**
per stock L2	0.000	0.087	0.130*	0.174*	0.261**	1.000**	0.000	0.000	1.000**	0.478**
per stock L1	0.087	0.087	1.000**	0.348**	0.261**	0.087	0.000	0.130*	0.130*	0.391**
RMSE										
per group L2	0.087	0.043	0.478**	1.000**	0.304**	0.000	1.000**	0.174*	0.087	0.348**
per group L1	0.087	1.000**	0.217*	0.217*	1.000**	0.000	0.478**	0.130*	0.130*	0.652**
eq by eq	0.087	0.043	0.304**	0.217*	0.217*	0.000	0.478**	1.000**	0.130*	1.000**
per group	1.000**	0.043	0.304**	0.217*	0.304**	0.000	0.478**	0.043	0.130*	0.522**
per stock L2	0.000	0.043	0.391**	0.087	0.261**	1.000**	0.043	0.043	1.000**	0.435**
per stock L1	0.087	0.043	1.000**	0.217*	0.261**	0.000	0.043	0.174*	0.130*	0.348**

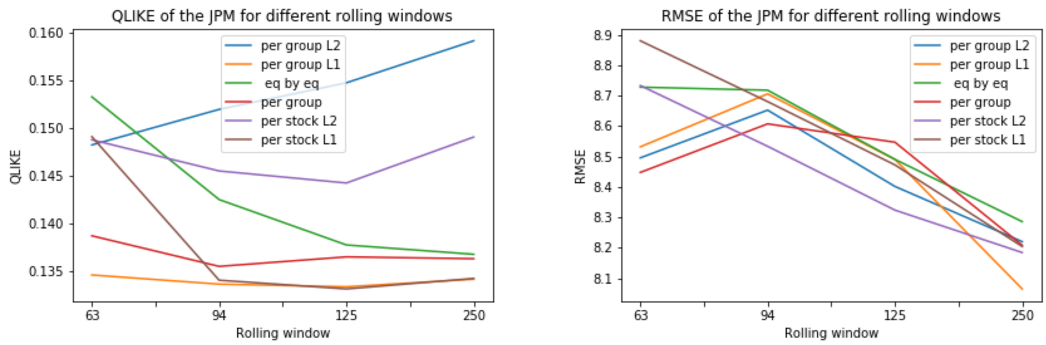
The forecasts with MCS p-value larger than 0.1 and 0.25 are indicated by * and ** respectively.

A.4 Performance measures (whole sample)



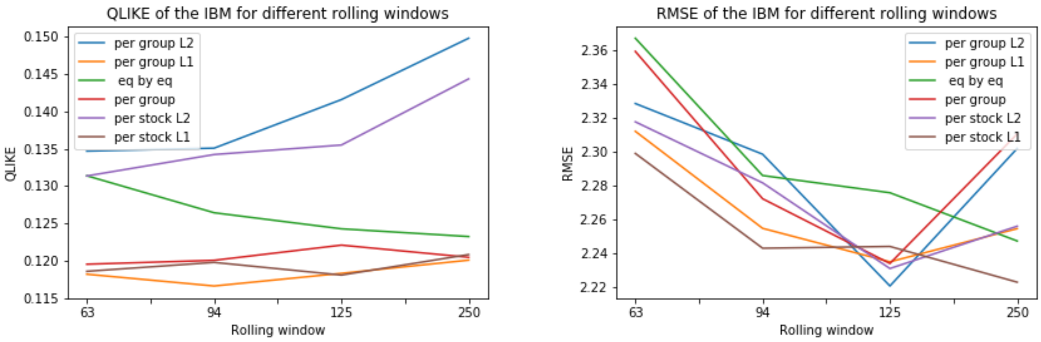
(a) QLIKE **(b)** RSME

Figure 23: Loss functions of the BAC for different rolling windows.



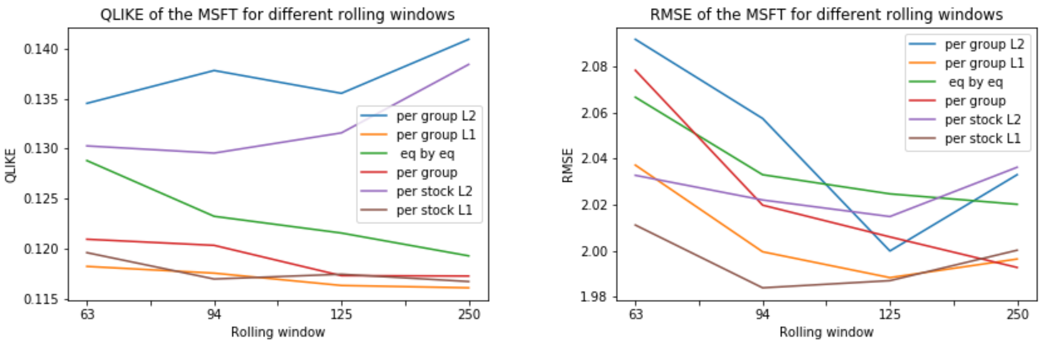
(a) QLIKE **(b)** RSME

Figure 24: Loss functions of the JPM for different rolling windows.



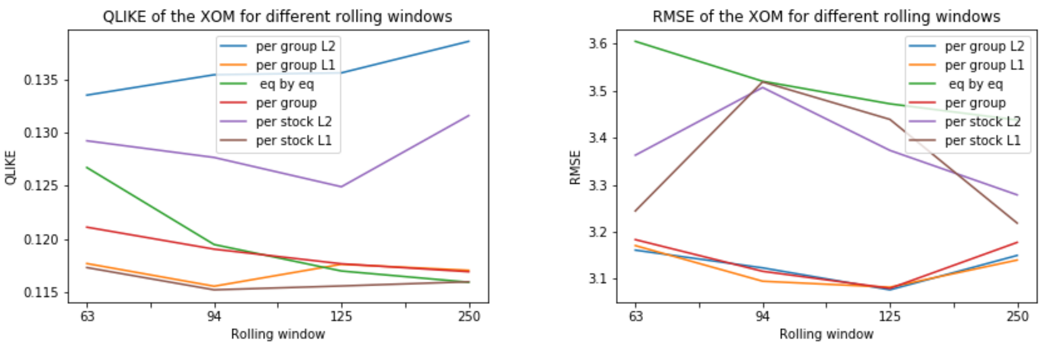
(a) QLIKE **(b)** RSME

Figure 25: Loss functions of the IBM for different rolling windows.



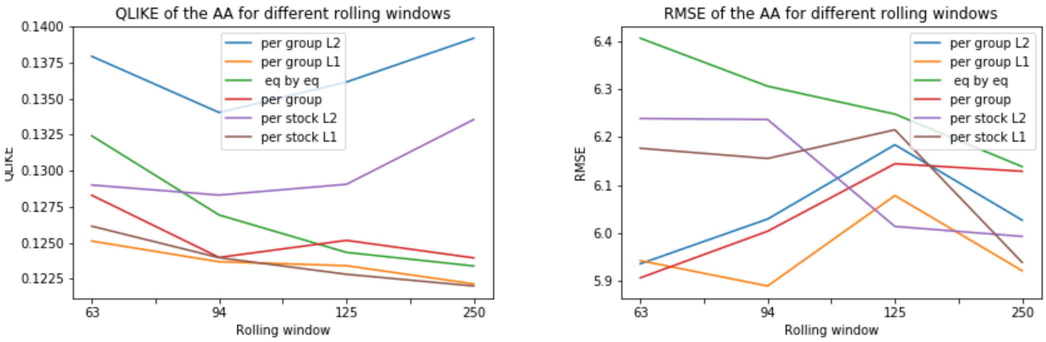
(a) QLIKE **(b)** RSME

Figure 26: Loss functions of the MSFT for different rolling windows.



(a) QLIKE **(b)** RSME

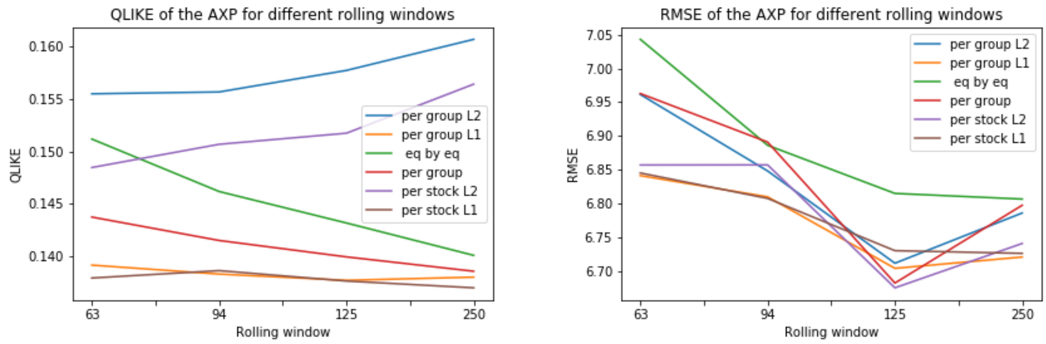
Figure 27: Loss functions of the XOM for different rolling windows.



(a) QLIKE

(b) RSME

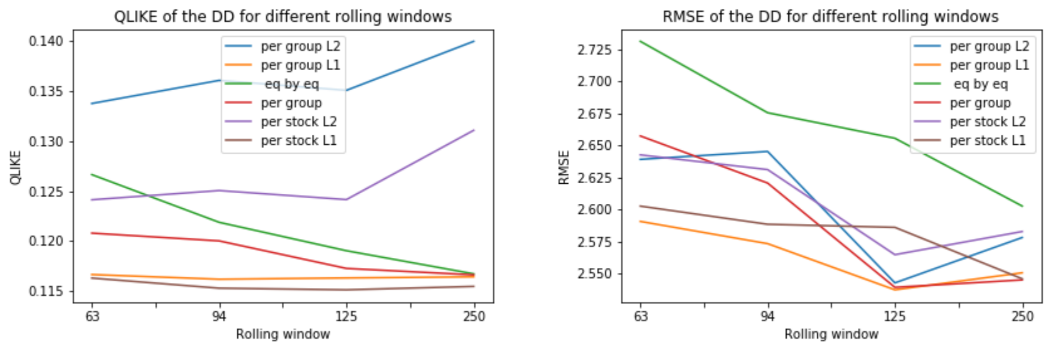
Figure 28: Loss functions of the AA for different rolling windows.



(a) QLIKE

(b) RSME

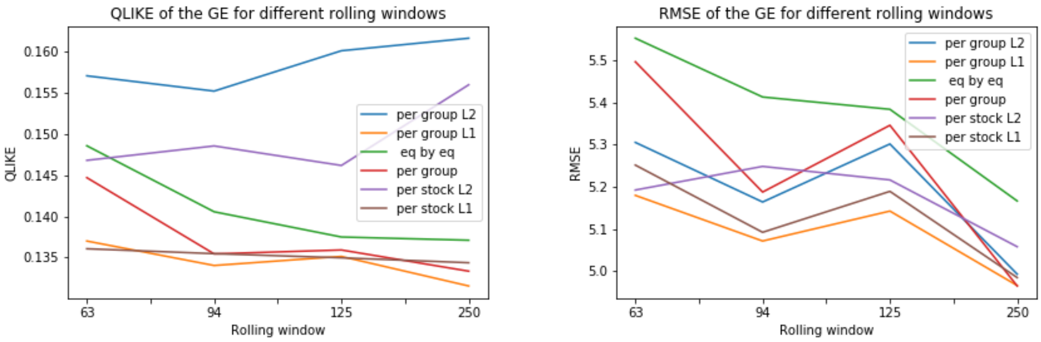
Figure 29: Loss functions of the AXP for different rolling windows.



(a) QLIKE

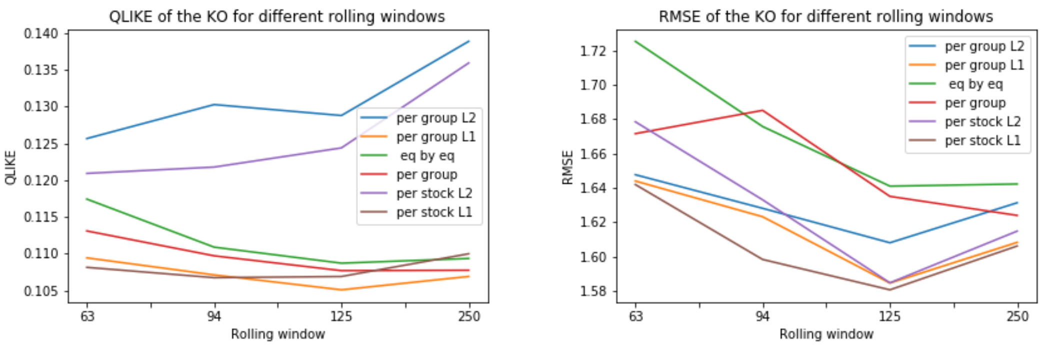
(b) RSME

Figure 30: Loss functions of the DD for different rolling windows.



(a) QLIKE **(b)** RSME

Figure 31: Loss functions of the GE for different rolling windows.



(a) QLIKE **(b)** RSME

Figure 32: Loss functions of the KO for different rolling windows.

Table 6: Loss functions

	L = 250		L = 125		L = 94		L = 63	
Forecast model	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE
BAC								
per group L2	11.872705	0.172854	11.783479	0.176336	11.742963	0.180892	11.625212	0.185614
per group L1	11.734174	0.155990	11.586367	0.156590	11.605431	0.156455	11.479756	0.156268
eq by eq	11.855833	0.170279	11.796169	0.166749	11.731453	0.163683	11.623384	0.159490
per group	11.895426	0.162095	11.782843	0.161901	11.739073	0.160630	11.575666	0.159111
per stock L2	11.827994	0.171715	11.678361	0.173687	11.632451	0.176262	11.630994	0.183223
per stock L1	11.688106	0.157797	11.644095	0.157679	11.586502	0.155916	11.565405	0.155538
JPM								
per group L2	8.495443	0.148210	8.652363	0.151947	8.401452	0.154742	8.219941	0.159160
per group L1	8.531684	0.134539	8.705772	0.133575	8.490261	0.133298	8.064509	0.134097
eq by eq	8.727912	0.153266	8.717886	0.142468	8.490523	0.137695	8.285719	0.136714
per group	8.447564	0.138658	8.607135	0.135444	8.546896	0.136441	8.208029	0.136250
per stock L2	8.733193	0.148748	8.533417	0.145473	8.323577	0.144208	8.184065	0.149039
per stock L1	8.880389	0.149075	8.680363	0.133988	8.472197	0.133075	8.203967	0.134175
IBM								
per group L2	2.328370	0.134676	2.298518	0.135071	2.220899	0.141576	2.301910	0.149794
per group L1	2.312013	0.118190	2.254855	0.116612	2.235137	0.118307	2.254681	0.120075
eq by eq	2.366832	0.131374	2.285972	0.126422	2.275828	0.124270	2.247410	0.123235
per group	2.359065	0.119527	2.272285	0.120049	2.234161	0.122085	2.310441	0.120454
per stock L2	2.317585	0.131349	2.281651	0.134218	2.231192	0.135491	2.256120	0.144351
per stock L1	2.298972	0.118577	2.243144	0.119771	2.244195	0.118077	2.223151	0.120822
MSFT								
per group L2	2.091845	0.134534	2.057484	0.137828	1.999926	0.135541	2.033070	0.140948
per group L1	2.037226	0.118201	1.999590	0.117526	1.988360	0.116292	1.996435	0.116056
eq by eq	2.066715	0.128804	2.033049	0.123221	2.024736	0.121545	2.020202	0.119257
per group	2.078410	0.120923	2.019830	0.120311	2.006043	0.117274	1.992766	0.117236
per stock L2	2.032789	0.130273	2.022075	0.129553	2.014879	0.131578	2.036302	0.138436
per stock L1	2.011195	0.119581	1.983930	0.116939	1.987040	0.117417	2.000265	0.116685
XOM								
per group L2	3.160852	0.133507	3.123148	0.135421	3.076389	0.135597	3.149600	0.138558
per group L1	3.170520	0.117693	3.094692	0.115562	3.081966	0.117624	3.139754	0.117058
eq by eq	3.605203	0.126708	3.520220	0.119479	3.472258	0.116994	3.437779	0.115940
per group	3.183233	0.121109	3.115830	0.119045	3.079522	0.117670	3.177321	0.116925
per stock L2	3.362924	0.129220	3.506979	0.127655	3.373131	0.124904	3.278439	0.131583
per stock L1	3.244234	0.117317	3.518983	0.115227	3.439094	0.115588	3.218300	0.115981
AA								
per group L2	5.936133	0.137936	6.029227	0.134038	6.183860	0.136151	6.026906	0.139193
per group L1	5.942635	0.125116	5.889891	0.123682	6.078153	0.123410	5.921649	0.122144
eq by eq	6.405642	0.132420	6.305772	0.126929	6.247689	0.124333	6.138289	0.123391
per group	5.906878	0.128295	6.003910	0.123983	6.144378	0.125168	6.128674	0.123952
per stock L2	6.238543	0.129006	6.236386	0.128310	6.013860	0.129057	5.993113	0.133550
per stock L1	6.176531	0.126151	6.155322	0.123965	6.215007	0.122814	5.938992	0.122001
AXP								
per group L2	6.961424	0.155507	6.848384	0.155683	6.712041	0.157744	6.786283	0.160717
per group L1	6.841520	0.139113	6.810177	0.138256	6.704493	0.137659	6.721202	0.137964
eq by eq	7.042937	0.151176	6.886359	0.146157	6.815204	0.143145	6.806886	0.140047
per group	6.962814	0.143713	6.891063	0.141471	6.682784	0.139900	6.797699	0.138527
per stock L2	6.857229	0.148451	6.857350	0.150675	6.675693	0.151736	6.741230	0.156429
per stock L1	6.845422	0.137879	6.807841	0.138585	6.730602	0.137589	6.726616	0.136941

	L = 250		L = 125		L = 94		L = 63	
Forecast model	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE
DD								
per group L2	2.639195	0.133755	2.645352	0.136062	2.542869	0.135068	2.578228	0.139961
per group L1	2.590795	0.116669	2.573450	0.116203	2.537447	0.116331	2.550720	0.116445
eq by eq	2.731225	0.126651	2.675548	0.121888	2.655639	0.119050	2.602761	0.116740
per group	2.657451	0.120807	2.620707	0.120031	2.539391	0.117281	2.545180	0.116641
per stock L2	2.642692	0.124139	2.631200	0.125059	2.564842	0.124156	2.582914	0.131076
per stock L1	2.602728	0.116318	2.588556	0.115308	2.586157	0.115146	2.546051	0.115487
GE								
per group L2	5.305041	0.157083	5.163905	0.155224	5.301305	0.160131	4.993005	0.161652
per group L1	5.179618	0.137003	5.071512	0.134032	5.142440	0.135131	4.966098	0.131540
eq by eq	5.551770	0.148580	5.412557	0.140560	5.383453	0.137492	5.166344	0.137102
per group	5.496047	0.144704	5.187457	0.135442	5.345590	0.135916	4.965012	0.133339
per stock L2	5.192273	0.146810	5.248224	0.148560	5.216201	0.146189	5.058001	0.155992
per stock L1	5.251327	0.136063	5.092149	0.135471	5.189138	0.134965	4.984866	0.134366
KO								
per group L2	1.647654	0.125644	1.628089	0.130256	1.608047	0.128773	1.631322	0.138839
per group L1	1.643980	0.109451	1.623183	0.107136	1.584581	0.105114	1.608276	0.106927
eq by eq	1.725270	0.117426	1.675702	0.110906	1.640995	0.108730	1.642240	0.109367
per group	1.671546	0.113105	1.685103	0.109728	1.635043	0.107717	1.623965	0.107771
per stock L2	1.678465	0.120917	1.633049	0.121774	1.584734	0.124376	1.614828	0.135902
per stock L1	1.641925	0.108163	1.598328	0.106765	1.580660	0.106923	1.606174	0.110008

The QLIKE and RMSE over different rolling windows $L = 63, 125, 94, 63$ for the different stocks.

Table 7: Diebold mariano BAC (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	0.569	0.19	-0.326	0.392	0.724
		(0.569)	(0.849)	(0.744)	(0.695)	(0.469)
per group L1	-0.569	-	-0.425	-0.576	-0.579	1.022
	(0.569)		(0.671)	(0.564)	(0.563)	(0.307)
eq by eq	-0.19	0.425	-	-0.639	0.171	0.56
	(0.849)	(0.671)		(0.523)	(0.865)	(0.576)
per group	0.326	0.576	0.639	-	0.43	0.71
	(0.744)	(0.564)	(0.523)		(0.668)	(0.478)
per stock L2	-0.392	0.579	-0.171	-0.43	-	0.822
	(0.695)	(0.563)	(0.865)	(0.668)		(0.411)
per stock L1	-0.724	-1.022	-0.56	-0.71	-0.822	-
	(0.469)	(0.307)	(0.576)	(0.478)	(0.411)	
L = 94						
per group L2	-	0.799	-0.255	0.016	0.806	0.591
		(0.425)	(0.799)	(0.987)	(0.42)	(0.555)
per group L1	-0.799	-	-0.786	-0.742	-0.614	-0.853
	(0.425)		(0.432)	(0.458)	(0.539)	(0.394)
eq by eq	0.255	0.786	-	0.332	0.83	0.602
	(0.799)	(0.432)		(0.74)	(0.407)	(0.547)
per group	-0.016	0.742	-0.332	-	0.725	0.554
	(0.987)	(0.458)	(0.74)		(0.468)	(0.58)
per stock L2	-0.806	0.614	-0.83	-0.725	-	0.259
	(0.42)	(0.539)	(0.407)	(0.468)		(0.796)
per stock L1	-0.591	0.853	-0.602	-0.554	-0.259	-
	(0.555)	(0.394)	(0.547)	(0.58)	(0.796)	
L = 125						
per group L2	-	0.644	0.324	0.133	1.35	0.741
		(0.519)	(0.746)	(0.894)	(0.177)	(0.459)
per group L1	-0.644	-	-0.6	-0.641	-0.183	0.543
	(0.519)		(0.548)	(0.521)	(0.855)	(0.587)
eq by eq	-0.324	0.6	-	-0.387	1.27	0.703
	(0.746)	(0.548)		(0.699)	(0.204)	(0.482)
per group	-0.133	0.641	0.387	-	1.363	0.739
	(0.894)	(0.521)	(0.699)		(0.173)	(0.46)
per stock L2	-1.35	0.183	-1.27	-1.363	-	0.319
	(0.177)	(0.855)	(0.204)	(0.173)		(0.749)
per stock L1	-0.741	-0.543	-0.703	-0.739	-0.319	-
	(0.459)	(0.587)	(0.482)	(0.46)	(0.749)	
L = 250						
per group L2	-	1.695*	0.02	0.671	-0.083	0.418
		(0.09)	(0.984)	(0.502)	(0.934)	(0.676)
per group L1	-1.695*	-	-1.545	-1.268	-1.691*	-0.892
	(0.09)		(0.122)	(0.205)	(0.091)	(0.373)
eq by eq	-0.02	1.545	-	1.225	-0.214	0.752
	(0.984)	(0.122)		(0.221)	(0.83)	(0.452)
per group	-0.671	1.268	-1.225	-	-1.41	0.115
	(0.502)	(0.205)	(0.221)		(0.159)	(0.908)
per stock L2	0.083	1.691*	0.214	1.41	-	0.685
	(0.934)	(0.091)	(0.83)	(0.159)		(0.494)
per stock L1	-0.418	0.892	-0.752	-0.115	-0.685	-
	(0.676)	(0.373)	(0.452)	(0.908)	(0.494)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 8: Diebold mariano BAC (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	6.281***	0.666	2.883***	0.45	5.484***
		(0.0)	(0.505)	(0.004)	(0.653)	(0.0)
per group L1	-6.281***	-	-3.814***	-1.878*	-6.628***	-1.623
	(0.0)		(0.0)	(0.061)	(0.0)	(0.105)
eq by eq	-0.666	3.814***	-	4.406***	-0.417	3.508***
	(0.505)	(0.0)		(0.0)	(0.677)	(0.0)
per group	-2.883***	1.878*	-4.406***	-	-2.716***	1.311
	(0.004)	(0.061)	(0.0)		(0.007)	(0.19)
per stock L2	-0.45	6.628***	0.417	2.716***	-	6.27***
	(0.653)	(0.0)	(0.677)	(0.007)		(0.0)
per stock L1	-5.484***	1.623	-3.508***	-1.311	-6.27***	-
	(0.0)	(0.105)	(0.0)	(0.19)	(0.0)	
L = 94						
per group L2	-	5.482***	3.567***	5.596***	0.997	5.319***
		(0.0)	(0.0)	(0.0)	(0.319)	(0.0)
per group L1	-5.482***	-	-3.076***	-2.092**	-5.954***	-0.881
	(0.0)		(0.002)	(0.037)	(0.0)	(0.378)
eq by eq	-3.567***	3.076***	-	2.832***	-2.582***	2.972***
	(0.0)	(0.002)		(0.005)	(0.01)	(0.003)
per group	-5.596***	2.092**	-2.832***	-	-4.557***	1.591
	(0.0)	(0.037)	(0.005)		(0.0)	(0.112)
per stock L2	-0.997	5.954***	2.582***	4.557***	-	6.417***
	(0.319)	(0.0)	(0.01)	(0.0)		(0.0)
per stock L1	-5.319***	0.881	-2.972***	-1.591	-6.417***	-
	(0.0)	(0.378)	(0.003)	(0.112)	(0.0)	
L = 125						
per group L2	-	6.743***	5.756***	6.668***	1.35	6.638***
		(0.0)	(0.0)	(0.0)	(0.177)	(0.0)
per group L1	-6.743***	-	-3.006***	-2.217**	-5.429***	0.426
	(0.0)		(0.003)	(0.027)	(0.0)	(0.67)
eq by eq	-5.756***	3.006***	-	2.199**	-4.059***	3.227***
	(0.0)	(0.003)		(0.028)	(0.0)	(0.001)
per group	-6.668***	2.217**	-2.199**	-	-4.838***	2.158**
	(0.0)	(0.027)	(0.028)		(0.0)	(0.031)
per stock L2	-1.35	5.429***	4.059***	4.838***	-	4.872***
	(0.177)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-6.638***	-0.426	-3.227***	-2.158**	-4.872***	-
	(0.0)	(0.67)	(0.001)	(0.031)	(0.0)	
L = 250						
per group L2	-	5.657***	6.75***	7.229***	0.642	6.671***
		(0.0)	(0.0)	(0.0)	(0.521)	(0.0)
per group L1	-5.657***	-	-1.236	-1.432	-5.404***	0.604
	(0.0)		(0.217)	(0.152)	(0.0)	(0.546)
eq by eq	-6.75***	1.236	-	0.296	-6.889***	2.141**
	(0.0)	(0.217)		(0.767)	(0.0)	(0.032)
per group	-7.229***	1.432	-0.296	-	-6.413***	2.548**
	(0.0)	(0.152)	(0.767)		(0.0)	(0.011)
per stock L2	-0.642	5.404***	6.889***	6.413***	-	6.274***
	(0.521)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-6.671***	-0.604	-2.141**	-2.548**	-6.274***	-
	(0.0)	(0.546)	(0.032)	(0.011)	(0.0)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 9: Diebold mariano JPM (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	-0.485	-0.818	1.269	-0.886	-1.036
		(0.628)	(0.413)	(0.204)	(0.376)	(0.3)
per group L1	0.485	-	-0.755	1.069	-0.842	-1.041
	(0.628)		(0.45)	(0.285)	(0.4)	(0.298)
eq by eq	0.818	0.755	-	0.934	-0.081	-0.958
	(0.413)	(0.45)		(0.35)	(0.935)	(0.338)
per group	-1.269	-1.069	-0.934	-	-1.006	-1.124
	(0.204)	(0.285)	(0.35)		(0.314)	(0.261)
per stock L2	0.886	0.842	0.081	1.006	-	-1.161
	(0.376)	(0.4)	(0.935)	(0.314)		(0.246)
per stock L1	1.036	1.041	0.958	1.124	1.161	-
	(0.3)	(0.298)	(0.338)	(0.261)	(0.246)	
L = 94						
per group L2	-	-0.344	-0.206	0.694	0.765	-0.107
		(0.731)	(0.837)	(0.488)	(0.445)	(0.915)
per group L1	0.344	-	-0.06	0.837	1.371	0.161
	(0.731)		(0.952)	(0.403)	(0.171)	(0.872)
eq by eq	0.206	0.06	-	0.413	0.908	0.491
	(0.837)	(0.952)		(0.679)	(0.364)	(0.623)
per group	-0.694	-0.837	-0.413	-	0.713	-0.348
	(0.488)	(0.403)	(0.679)		(0.476)	(0.728)
per stock L2	-0.765	-1.371	-0.908	-0.713	-	-0.952
	(0.445)	(0.171)	(0.364)	(0.476)		(0.341)
per stock L1	0.107	-0.161	-0.491	0.348	0.952	-
	(0.915)	(0.872)	(0.623)	(0.728)	(0.341)	
L = 125						
per group L2	-	-0.717	-0.656	-1.215	1.151	-0.556
		(0.473)	(0.512)	(0.225)	(0.25)	(0.578)
per group L1	0.717	-	-0.005	-1.36	1.061	0.264
	(0.473)		(0.996)	(0.174)	(0.289)	(0.791)
eq by eq	0.656	0.005	-	-1.318	1.033	0.301
	(0.512)	(0.996)		(0.188)	(0.302)	(0.764)
per group	1.215	1.36	1.318	-	1.441	1.353
	(0.225)	(0.174)	(0.188)		(0.15)	(0.176)
per stock L2	-1.151	-1.061	-1.033	-1.441	-	-1.064
	(0.25)	(0.289)	(0.302)	(0.15)		(0.287)
per stock L1	0.556	-0.264	-0.301	-1.353	1.064	-
	(0.578)	(0.791)	(0.764)	(0.176)	(0.287)	
L = 250						
per group L2	-	2.261**	-0.394	0.23	0.494	0.092
		(0.024)	(0.694)	(0.818)	(0.621)	(0.927)
per group L1	-2.261**	-	-1.169	-1.862*	-1.339	-0.739
	(0.024)		(0.243)	(0.063)	(0.181)	(0.46)
eq by eq	0.394	1.169	-	0.599	0.828	1.804*
	(0.694)	(0.243)		(0.549)	(0.408)	(0.071)
per group	-0.23	1.862*	-0.599	-	0.47	0.03
	(0.818)	(0.063)	(0.549)		(0.638)	(0.976)
per stock L2	-0.494	1.339	-0.828	-0.47	-	-0.153
	(0.621)	(0.181)	(0.408)	(0.638)		(0.878)
per stock L1	-0.092	0.739	-1.804*	-0.03	0.153	-
	(0.927)	(0.46)	(0.071)	(0.976)	(0.878)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 10: Diebold mariano JPM (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	7.689***	-0.616	4.308***	-0.108	-0.063
		(0.0)	(0.538)	(0.0)	(0.914)	(0.95)
per group L1	-7.689***	-	-2.232**	-2.072**	-2.731***	-1.033
	(0.0)		(0.026)	(0.038)	(0.006)	(0.302)
eq by eq	0.616	2.232**	-	1.909*	1.166	0.626
	(0.538)	(0.026)		(0.056)	(0.244)	(0.531)
per group	-4.308***	2.072**	-1.909*	-	-2.05**	-0.763
	(0.0)	(0.038)	(0.056)		(0.04)	(0.446)
per stock L2	0.108	2.731***	-1.166	2.05**	-	-0.035
	(0.914)	(0.006)	(0.244)	(0.04)		(0.972)
per stock L1	0.063	1.033	-0.626	0.763	0.035	-
	(0.95)	(0.302)	(0.531)	(0.446)	(0.972)	
L = 94						
per group L2	-	8.25***	2.716***	7.993***	2.651***	7.231***
		(0.0)	(0.007)	(0.0)	(0.008)	(0.0)
per group L1	-8.25***	-	-2.915***	-2.04**	-5.287***	-0.345
	(0.0)		(0.004)	(0.041)	(0.0)	(0.73)
eq by eq	-2.716***	2.915***	-	2.658***	-1.375	3.524***
	(0.007)	(0.004)		(0.008)	(0.169)	(0.0)
per group	-7.993***	2.04**	-2.658***	-	-5.154***	1.229
	(0.0)	(0.041)	(0.008)		(0.0)	(0.219)
per stock L2	-2.651***	5.287***	1.375	5.154***	-	6.707***
	(0.008)	(0.0)	(0.169)	(0.0)		(0.0)
per stock L1	-7.231***	0.345	-3.524***	-1.229	-6.707***	-
	(0.0)	(0.73)	(0.0)	(0.219)	(0.0)	
L = 125						
per group L2	-	8.269***	5.889***	7.461***	4.547***	8.079***
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
per group L1	-8.269***	-	-2.473**	-3.238***	-5.136***	0.24
	(0.0)		(0.013)	(0.001)	(0.0)	(0.81)
eq by eq	-5.889***	2.473**	-	0.747	-3.291***	2.93***
	(0.0)	(0.013)		(0.455)	(0.001)	(0.003)
per group	-7.461***	3.238***	-0.747	-	-4.024***	2.748***
	(0.0)	(0.001)	(0.455)		(0.0)	(0.006)
per stock L2	-4.547***	5.136***	3.291***	4.024***	-	5.963***
	(0.0)	(0.0)	(0.001)	(0.0)		(0.0)
per stock L1	-8.079***	-0.24	-2.93***	-2.748***	-5.963***	-
	(0.0)	(0.81)	(0.003)	(0.006)	(0.0)	
L = 250						
per group L2	-	8.118***	7.246***	8.498***	3.752***	7.87***
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
per group L1	-8.118***	-	-1.858*	-2.203**	-5.557***	-0.096
	(0.0)		(0.063)	(0.028)	(0.0)	(0.924)
eq by eq	-7.246***	1.858*	-	0.453	-5.658***	1.89*
	(0.0)	(0.063)		(0.651)	(0.0)	(0.059)
per group	-8.498***	2.203**	-0.453	-	-5.417***	1.741*
	(0.0)	(0.028)	(0.651)		(0.0)	(0.082)
per stock L2	-3.752***	5.557***	5.658***	5.417***	-	5.994***
	(0.0)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-7.87***	0.096	-1.89*	-1.741*	-5.994***	-
	(0.0)	(0.924)	(0.059)	(0.082)	(0.0)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 11: Diebold mariano IBM (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	0.378	-0.314	-0.929	0.154	0.455
		(0.705)	(0.753)	(0.353)	(0.877)	(0.649)
per group L1	-0.378	-	-0.529	-0.95	-0.134	0.421
	(0.705)		(0.597)	(0.342)	(0.893)	(0.674)
eq by eq	0.314	0.529	-	0.067	0.569	0.808
	(0.753)	(0.597)		(0.946)	(0.569)	(0.419)
per group	0.929	0.95	-0.067	-	0.536	0.89
	(0.353)	(0.342)	(0.946)		(0.592)	(0.374)
per stock L2	-0.154	0.134	-0.569	-0.536	-	0.791
	(0.877)	(0.893)	(0.569)	(0.592)		(0.429)
per stock L1	-0.455	-0.421	-0.808	-0.89	-0.791	-
	(0.649)	(0.674)	(0.419)	(0.374)	(0.429)	
L = 94						
per group L2	-	1.869*	0.168	1.264	0.357	1.335
		(0.062)	(0.866)	(0.207)	(0.721)	(0.182)
per group L1	-1.869*	-	-0.415	-0.877	-0.628	0.411
	(0.062)		(0.678)	(0.381)	(0.53)	(0.681)
eq by eq	-0.168	0.415	-	0.199	0.053	0.565
	(0.866)	(0.678)		(0.842)	(0.958)	(0.572)
per group	-1.264	0.877	-0.199	-	-0.178	0.741
	(0.207)	(0.381)	(0.842)		(0.858)	(0.459)
per stock L2	-0.357	0.628	-0.053	0.178	-	1.356
	(0.721)	(0.53)	(0.958)	(0.858)		(0.175)
per stock L1	-1.335	-0.411	-0.565	-0.741	-1.356	-
	(0.182)	(0.681)	(0.572)	(0.459)	(0.175)	
L = 125						
per group L2	-	-0.274	-0.915	-0.517	-0.174	-0.373
		(0.784)	(0.36)	(0.605)	(0.862)	(0.709)
per group L1	0.274	-	-0.484	0.015	0.066	-0.379
	(0.784)		(0.629)	(0.988)	(0.947)	(0.705)
eq by eq	0.915	0.484	-	0.788	0.831	0.388
	(0.36)	(0.629)		(0.431)	(0.406)	(0.698)
per group	0.517	-0.015	-0.788	-	0.044	-0.135
	(0.605)	(0.988)	(0.431)		(0.965)	(0.893)
per stock L2	0.174	-0.066	-0.831	-0.044	-	-0.264
	(0.862)	(0.947)	(0.406)	(0.965)		(0.791)
per stock L1	0.373	0.379	-0.388	0.135	0.264	-
	(0.709)	(0.705)	(0.698)	(0.893)	(0.791)	
L = 250						
per group L2	-	1.977**	1.419	-0.203	0.849	1.467
		(0.048)	(0.156)	(0.839)	(0.396)	(0.143)
per group L1	-1.977**	-	0.155	-0.992	-0.029	0.812
	(0.048)		(0.877)	(0.321)	(0.977)	(0.417)
eq by eq	-1.419	-0.155	-	-1.021	-0.219	0.433
	(0.156)	(0.877)		(0.307)	(0.827)	(0.665)
per group	0.203	0.992	1.021	-	0.611	0.951
	(0.839)	(0.321)	(0.307)		(0.541)	(0.342)
per stock L2	-0.849	0.029	0.219	-0.611	-	1.035
	(0.396)	(0.977)	(0.827)	(0.541)		(0.301)
per stock L1	-1.467	-0.812	-0.433	-0.951	-1.035	-
	(0.143)	(0.417)	(0.665)	(0.342)	(0.301)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 12: Diebold mariano IBM (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	6.044***	0.731	4.591***	1.141	5.746***
		(0.0)	(0.465)	(0.0)	(0.254)	(0.0)
per group L1	-6.044***	-	-3.49***	-0.789	-6.342***	-0.435
	(0.0)		(0.0)	(0.43)	(0.0)	(0.663)
eq by eq	-0.731	3.49***	-	3.512***	0.007	3.639***
	(0.465)	(0.0)		(0.0)	(0.994)	(0.0)
per group	-4.591***	0.789	-3.512***	-	-4.983***	0.54
	(0.0)	(0.43)	(0.0)		(0.0)	(0.589)
per stock L2	-1.141	6.342***	-0.007	4.983***	-	7.304***
	(0.254)	(0.0)	(0.994)	(0.0)		(0.0)
per stock L1	-5.746***	0.435	-3.639***	-0.54	-7.304***	-
	(0.0)	(0.663)	(0.0)	(0.589)	(0.0)	
L = 94						
per group L2	-	7.73***	3.218***	6.463***	0.371	6.447***
		(0.0)	(0.001)	(0.0)	(0.711)	(0.0)
per group L1	-7.73***	-	-4.841***	-3.264***	-7.321***	-3.145***
	(0.0)		(0.0)	(0.001)	(0.0)	(0.002)
eq by eq	-3.218***	4.841***	-	3.608***	-3.394***	3.584***
	(0.001)	(0.0)		(0.0)	(0.001)	(0.0)
per group	-6.463***	3.264***	-3.608***	-	-6.011***	0.195
	(0.0)	(0.001)	(0.0)		(0.0)	(0.845)
per stock L2	-0.371	7.321***	3.394***	6.011***	-	7.023***
	(0.711)	(0.0)	(0.001)	(0.0)		(0.0)
per stock L1	-6.447***	3.145***	-3.584***	-0.195	-7.023***	-
	(0.0)	(0.002)	(0.0)	(0.845)	(0.0)	
L = 125						
per group L2	-	8.162***	6.694***	7.705***	2.596***	8.066***
		(0.0)	(0.0)	(0.0)	(0.009)	(0.0)
per group L1	-8.162***	-	-3.455***	-2.678***	-6.783***	0.381
	(0.0)		(0.001)	(0.007)	(0.0)	(0.703)
eq by eq	-6.694***	3.455***	-	2.212**	-4.524***	3.71***
	(0.0)	(0.001)		(0.027)	(0.0)	(0.0)
per group	-7.705***	2.678***	-2.212**	-	-5.392***	2.63***
	(0.0)	(0.007)	(0.027)		(0.0)	(0.009)
per stock L2	-2.596***	6.783***	4.524***	5.392***	-	7.052***
	(0.009)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-8.066***	-0.381	-3.71***	-2.63***	-7.052***	-
	(0.0)	(0.703)	(0.0)	(0.009)	(0.0)	
L = 250						
per group L2	-	5.15***	5.176***	5.605***	1.475	5.009***
		(0.0)	(0.0)	(0.0)	(0.14)	(0.0)
per group L1	-5.15***	-	-2.346**	-0.332	-7.009***	-1.624
	(0.0)		(0.019)	(0.74)	(0.0)	(0.104)
eq by eq	-5.176***	2.346**	-	3.453***	-7.66***	1.82*
	(0.0)	(0.019)		(0.001)	(0.0)	(0.069)
per group	-5.605***	0.332	-3.453***	-	-7.968***	-0.284
	(0.0)	(0.74)	(0.001)		(0.0)	(0.776)
per stock L2	-1.475	7.009***	7.66***	7.968***	-	6.838***
	(0.14)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-5.009***	1.624	-1.82*	0.284	-6.838***	-
	(0.0)	(0.104)	(0.069)	(0.776)	(0.0)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 13: Diebold mariano MSFT (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	2.049**	0.433	0.414	1.071	1.434
		(0.041)	(0.665)	(0.679)	(0.285)	(0.152)
per group L1	-2.049**	-	-0.488	-1.185	0.083	0.545
	(0.041)		(0.626)	(0.236)	(0.934)	(0.586)
eq by eq	-0.433	0.488	-	-0.261	0.811	1.045
	(0.665)	(0.626)		(0.794)	(0.417)	(0.296)
per group	-0.414	1.185	0.261	-	0.831	1.129
	(0.679)	(0.236)	(0.794)		(0.406)	(0.259)
per stock L2	-1.071	-0.083	-0.811	-0.831	-	1.057
	(0.285)	(0.934)	(0.417)	(0.406)		(0.29)
per stock L1	-1.434	-0.545	-1.045	-1.129	-1.057	-
	(0.152)	(0.586)	(0.296)	(0.259)	(0.29)	
L = 94						
per group L2	-	2.122**	0.408	1.905*	0.918	2.211**
		(0.034)	(0.683)	(0.057)	(0.359)	(0.027)
per group L1	-2.122**	-	-0.555	-1.208	-0.558	0.78
	(0.034)		(0.579)	(0.227)	(0.577)	(0.435)
eq by eq	-0.408	0.555	-	0.242	0.276	0.906
	(0.683)	(0.579)		(0.809)	(0.782)	(0.365)
per group	-1.905*	1.208	-0.242	-	-0.062	1.579
	(0.057)	(0.227)	(0.809)		(0.95)	(0.114)
per stock L2	-0.918	0.558	-0.276	0.062	-	1.147
	(0.359)	(0.577)	(0.782)	(0.95)		(0.251)
per stock L1	-2.211**	-0.78	-0.906	-1.579	-1.147	-
	(0.027)	(0.435)	(0.365)	(0.114)	(0.251)	
L = 125						
per group L2	-	0.287	-0.741	-0.178	-0.433	0.271
		(0.774)	(0.459)	(0.859)	(0.665)	(0.787)
per group L1	-0.287	-	-0.708	-0.336	-0.753	0.074
	(0.774)		(0.479)	(0.737)	(0.451)	(0.941)
eq by eq	0.741	0.708	-	0.734	0.242	0.659
	(0.459)	(0.479)		(0.463)	(0.809)	(0.51)
per group	0.178	0.336	-0.734	-	-0.19	0.301
	(0.859)	(0.737)	(0.463)		(0.85)	(0.763)
per stock L2	0.433	0.753	-0.242	0.19	-	0.748
	(0.665)	(0.451)	(0.809)	(0.85)		(0.454)
per stock L1	-0.271	-0.074	-0.659	-0.301	-0.748	-
	(0.787)	(0.941)	(0.51)	(0.763)	(0.454)	
L = 250						
per group L2	-	2.004**	0.396	2.735***	-0.153	1.365
		(0.045)	(0.692)	(0.006)	(0.878)	(0.172)
per group L1	-2.004**	-	-0.737	0.287	-1.437	-0.136
	(0.045)		(0.461)	(0.774)	(0.151)	(0.891)
eq by eq	-0.396	0.737	-	0.979	-0.428	0.425
	(0.692)	(0.461)		(0.328)	(0.669)	(0.671)
per group	-2.735***	-0.287	-0.979	-	-1.828*	-0.268
	(0.006)	(0.774)	(0.328)		(0.068)	(0.789)
per stock L2	0.153	1.437	0.428	1.828*	-	1.6
	(0.878)	(0.151)	(0.669)	(0.068)		(0.11)
per stock L1	-1.365	0.136	-0.425	0.268	-1.6	-
	(0.172)	(0.891)	(0.671)	(0.789)	(0.11)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 14: Diebold mariano MSFT (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	8.348***	2.235**	6.966***	1.972**	7.571***
		(0.0)	(0.025)	(0.0)	(0.049)	(0.0)
per group L1	-8.348***	-	-4.705***	-2.442**	-6.545***	-1.607
	(0.0)		(0.0)	(0.015)	(0.0)	(0.108)
eq by eq	-2.235**	4.705***	-	3.994***	-0.778	4.83***
	(0.025)	(0.0)		(0.0)	(0.437)	(0.0)
per group	-6.966***	2.442**	-3.994***	-	-4.937***	1.058
	(0.0)	(0.015)	(0.0)		(0.0)	(0.29)
per stock L2	-1.972**	6.545***	0.778	4.937***	-	7.013***
	(0.049)	(0.0)	(0.437)	(0.0)		(0.0)
per stock L1	-7.571***	1.607	-4.83***	-1.058	-7.013***	-
	(0.0)	(0.108)	(0.0)	(0.29)	(0.0)	
L = 94						
per group L2	-	7.913***	5.728***	7.112***	3.85***	8.13***
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
per group L1	-7.913***	-	-3.165***	-2.33**	-5.793***	0.866
	(0.0)		(0.002)	(0.02)	(0.0)	(0.387)
eq by eq	-5.728***	3.165***	-	2.399**	-3.678***	3.925***
	(0.0)	(0.002)		(0.017)	(0.0)	(0.0)
per group	-7.112***	2.33**	-2.399**	-	-4.907***	2.717***
	(0.0)	(0.02)	(0.017)		(0.0)	(0.007)
per stock L2	-3.85***	5.793***	3.678***	4.907***	-	6.634***
	(0.0)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-8.13***	-0.866	-3.925***	-2.717***	-6.634***	-
	(0.0)	(0.387)	(0.0)	(0.007)	(0.0)	
L = 125						
per group L2	-	7.855***	5.67***	7.67***	1.624	7.407***
		(0.0)	(0.0)	(0.0)	(0.104)	(0.0)
per group L1	-7.855***	-	-2.905***	-1.137	-6.211***	-1.367
	(0.0)		(0.004)	(0.256)	(0.0)	(0.172)
eq by eq	-5.67***	2.905***	-	2.895***	-5.209***	2.875***
	(0.0)	(0.004)		(0.004)	(0.0)	(0.004)
per group	-7.67***	1.137	-2.895***	-	-6.127***	-0.124
	(0.0)	(0.256)	(0.004)		(0.0)	(0.902)
per stock L2	-1.624	6.211***	5.209***	6.127***	-	6.458***
	(0.104)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-7.407***	1.367	-2.875***	0.124	-6.458***	-
	(0.0)	(0.172)	(0.004)	(0.902)	(0.0)	
L = 250						
per group L2	-	7.922***	7.494***	8.247***	1.067	7.657***
		(0.0)	(0.0)	(0.0)	(0.286)	(0.0)
per group L1	-7.922***	-	-2.347**	-1.401	-7.974***	-0.934
	(0.0)		(0.019)	(0.161)	(0.0)	(0.35)
eq by eq	-7.494***	2.347**	-	1.902*	-7.871***	2.138**
	(0.0)	(0.019)		(0.057)	(0.0)	(0.033)
per group	-8.247***	1.401	-1.902*	-	-8.229***	0.545
	(0.0)	(0.161)	(0.057)		(0.0)	(0.586)
per stock L2	-1.067	7.974***	7.871***	8.229***	-	7.886***
	(0.286)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-7.657***	0.934	-2.138**	-0.545	-7.886***	-
	(0.0)	(0.35)	(0.033)	(0.586)	(0.0)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 15: Diebold mariano XOM (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	-0.241 (0.81)	-1.123 (0.262)	-0.472 (0.637)	-1.151 (0.25)	-0.711 (0.477)
per group L1	0.241 (0.81)	-	-1.118 (0.264)	-0.197 (0.844)	-1.191 (0.234)	-0.691 (0.49)
eq by eq	1.123 (0.262)	1.118 (0.264)	-	1.079 (0.281)	1.043 (0.297)	1.213 (0.225)
per group	0.472 (0.637)	0.197 (0.844)	-1.079 (0.281)	-	-1.007 (0.314)	-0.502 (0.616)
per stock L2	1.151 (0.25)	1.191 (0.234)	-1.043 (0.297)	1.007 (0.314)	-	1.529 (0.126)
per stock L1	0.711 (0.477)	0.691 (0.49)	-1.213 (0.225)	0.502 (0.616)	-1.529 (0.126)	-
L = 94						
per group L2	-	1.372 (0.17)	-1.037 (0.3)	0.387 (0.699)	-1.056 (0.291)	-0.93 (0.352)
per group L1	-1.372 (0.17)	-	-1.083 (0.279)	-0.856 (0.392)	-1.108 (0.268)	-0.977 (0.329)
eq by eq	1.037 (0.3)	1.083 (0.279)	-	1.089 (0.276)	0.162 (0.871)	0.013 (0.989)
per group	-0.387 (0.699)	0.856 (0.392)	-1.089 (0.276)	-	-1.108 (0.268)	-0.972 (0.331)
per stock L2	1.056 (0.291)	1.108 (0.268)	-0.162 (0.871)	1.108 (0.268)	-	-0.176 (0.86)
per stock L1	0.93 (0.352)	0.977 (0.329)	-0.013 (0.989)	0.972 (0.331)	0.176 (0.86)	-
L = 125						
per group L2	-	-0.223 (0.824)	-1.087 (0.277)	-0.206 (0.837)	-0.98 (0.327)	-1.103 (0.27)
per group L1	0.223 (0.824)	-	-1.112 (0.266)	0.1 (0.921)	-1.004 (0.316)	-1.129 (0.259)
eq by eq	1.087 (0.277)	1.112 (0.266)	-	1.063 (0.288)	1.35 (0.177)	0.4 (0.689)
per group	0.206 (0.837)	-0.1 (0.921)	-1.063 (0.288)	-	-0.951 (0.342)	-1.075 (0.283)
per stock L2	0.98 (0.327)	1.004 (0.316)	-1.35 (0.177)	0.951 (0.342)	-	-1.137 (0.256)
per stock L1	1.103 (0.27)	1.129 (0.259)	-0.4 (0.689)	1.075 (0.283)	1.137 (0.256)	-
L = 250						
per group L2	-	0.198 (0.843)	-0.895 (0.371)	-0.445 (0.656)	-0.988 (0.323)	-0.966 (0.334)
per group L1	-0.198 (0.843)	-	-0.843 (0.399)	-0.359 (0.719)	-0.874 (0.382)	-0.895 (0.371)
eq by eq	0.895 (0.371)	0.843 (0.399)	-	0.983 (0.326)	0.803 (0.422)	0.821 (0.412)
per group	0.445 (0.656)	0.359 (0.719)	-0.983 (0.326)	-	-1.237 (0.216)	-0.71 (0.478)
per stock L2	0.988 (0.323)	0.874 (0.382)	-0.803 (0.422)	1.237 (0.216)	-	0.801 (0.423)
per stock L1	0.966 (0.334)	0.895 (0.371)	-0.821 (0.412)	0.71 (0.478)	-0.801 (0.423)	-

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 16: Diebold mariano XOM (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	7.658*** (0.0)	1.97** (0.049)	6.138*** (0.0)	1.943* (0.052)	7.404*** (0.0)
per group L1	-7.658*** (0.0)	-	-2.936*** (0.003)	-2.615*** (0.009)	-5.949*** (0.0)	0.379 (0.704)
eq by eq	-1.97** (0.049)	2.936*** (0.003)	-	1.987** (0.047)	-0.915 (0.36)	3.623*** (0.0)
per group	-6.138*** (0.0)	2.615*** (0.009)	-1.987** (0.047)	-	-4.074*** (0.0)	2.642*** (0.008)
per stock L2	-1.943* (0.052)	5.949*** (0.0)	0.915 (0.36)	4.074*** (0.0)	-	7.216*** (0.0)
per stock L1	-7.404*** (0.0)	-0.379 (0.704)	-3.623*** (0.0)	-2.642*** (0.008)	-7.216*** (0.0)	-
L = 94						
per group L2	-	9.449*** (0.0)	6.19*** (0.0)	8.485*** (0.0)	3.623*** (0.0)	9.083*** (0.0)
per group L1	-9.449*** (0.0)	-	-2.151** (0.032)	-3.339*** (0.001)	-6.352*** (0.0)	0.336 (0.737)
eq by eq	-6.19*** (0.0)	2.151** (0.032)	-	0.26 (0.795)	-4.874*** (0.0)	3.18*** (0.001)
per group	-8.485*** (0.0)	3.339*** (0.001)	-0.26 (0.795)	-	-4.7*** (0.0)	3.04*** (0.002)
per stock L2	-3.623*** (0.0)	6.352*** (0.0)	4.874*** (0.0)	4.7*** (0.0)	-	7.527*** (0.0)
per stock L1	-9.083*** (0.0)	-0.336 (0.737)	-3.18*** (0.001)	-3.04*** (0.002)	-7.527*** (0.0)	-
L = 125						
per group L2	-	8.396*** (0.0)	7.317*** (0.0)	8.289*** (0.0)	5.043*** (0.0)	8.581*** (0.0)
per group L1	-8.396*** (0.0)	-	0.412 (0.68)	-0.055 (0.956)	-4.502*** (0.0)	2.801*** (0.005)
eq by eq	-7.317*** (0.0)	-0.412 (0.68)	-	-0.484 (0.628)	-4.692*** (0.0)	1.101 (0.271)
per group	-8.289*** (0.0)	0.055 (0.956)	0.484 (0.628)	-	-4.308*** (0.0)	2.032** (0.042)
per stock L2	-5.043*** (0.0)	4.502*** (0.0)	4.692*** (0.0)	4.308*** (0.0)	-	5.788*** (0.0)
per stock L1	-8.581*** (0.0)	-2.801*** (0.005)	-1.101 (0.271)	-2.032** (0.042)	-5.788*** (0.0)	-
L = 250						
per group L2	-	7.405*** (0.0)	8.169*** (0.0)	8.006*** (0.0)	3.108*** (0.002)	7.518*** (0.0)
per group L1	-7.405*** (0.0)	-	0.769 (0.442)	0.128 (0.898)	-5.99*** (0.0)	1.4 (0.162)
eq by eq	-8.169*** (0.0)	-0.769 (0.442)	-	-0.837 (0.403)	-8.184*** (0.0)	-0.034 (0.973)
per group	-8.006*** (0.0)	-0.128 (0.898)	0.837 (0.403)	-	-6.502*** (0.0)	0.795 (0.427)
per stock L2	-3.108*** (0.002)	5.99*** (0.0)	8.184*** (0.0)	6.502*** (0.0)	-	6.711*** (0.0)
per stock L1	-7.518*** (0.0)	-1.4 (0.162)	0.034 (0.973)	-0.795 (0.427)	-6.711*** (0.0)	-

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 17: Diebold mariano AA (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	-0.175 (0.861)	-1.217 (0.224)	0.741 (0.459)	-1.387 (0.166)	-1.112 (0.266)
per group L1	0.175 (0.861)	-	-1.185 (0.236)	0.826 (0.409)	-1.332 (0.183)	-1.092 (0.275)
eq by eq	1.217 (0.224)	1.185 (0.236)	-	1.298 (0.194)	0.94 (0.347)	1.085 (0.278)
per group	-0.741 (0.459)	-0.826 (0.409)	-1.298 (0.194)	-	-1.498 (0.134)	-1.242 (0.214)
per stock L2	1.387 (0.166)	1.332 (0.183)	-0.94 (0.347)	1.498 (0.134)	-	0.733 (0.464)
per stock L1	1.112 (0.266)	1.092 (0.275)	-1.085 (0.278)	1.242 (0.214)	-0.733 (0.464)	-
L = 94						
per group L2	-	2.368** (0.018)	-0.859 (0.39)	0.251 (0.802)	-0.751 (0.453)	-0.482 (0.63)
per group L1	-2.368** (0.018)	-	-1.228 (0.22)	-1.145 (0.252)	-1.195 (0.232)	-0.976 (0.329)
eq by eq	0.859 (0.39)	1.228 (0.22)	-	0.973 (0.331)	0.523 (0.601)	0.976 (0.329)
per group	-0.251 (0.802)	1.145 (0.252)	-0.973 (0.331)	-	-0.796 (0.426)	-0.528 (0.598)
per stock L2	0.751 (0.453)	1.195 (0.232)	-0.523 (0.601)	0.796 (0.426)	-	1.638 (0.101)
per stock L1	0.482 (0.63)	0.976 (0.329)	-0.976 (0.329)	0.528 (0.598)	-1.638 (0.101)	-
L = 125						
per group L2	-	1.311 (0.19)	-0.506 (0.613)	0.571 (0.568)	1.225 (0.221)	-0.295 (0.768)
per group L1	-1.311 (0.19)	-	-1.295 (0.195)	-0.735 (0.463)	0.523 (0.601)	-1.507 (0.132)
eq by eq	0.506 (0.613)	1.295 (0.195)	-	0.984 (0.325)	1.536 (0.125)	0.291 (0.771)
per group	-0.571 (0.568)	0.735 (0.463)	-0.984 (0.325)	-	0.83 (0.407)	-0.609 (0.543)
per stock L2	-1.225 (0.221)	-0.523 (0.601)	-1.536 (0.125)	-0.83 (0.407)	-	-1.479 (0.139)
per stock L1	0.295 (0.768)	1.507 (0.132)	-0.291 (0.771)	0.609 (0.543)	1.479 (0.139)	-
L = 250						
per group L2	-	1.658* (0.098)	-0.77 (0.441)	-0.716 (0.474)	0.559 (0.576)	1.336 (0.182)
per group L1	-1.658* (0.098)	-	-1.152 (0.25)	-1.1 (0.271)	-1.352 (0.176)	-0.306 (0.759)
eq by eq	0.77 (0.441)	1.152 (0.25)	-	0.417 (0.677)	0.833 (0.405)	1.304 (0.192)
per group	0.716 (0.474)	1.1 (0.271)	-0.417 (0.677)	-	0.777 (0.437)	1.219 (0.223)
per stock L2	-0.559 (0.576)	1.352 (0.176)	-0.833 (0.405)	-0.777 (0.437)	-	0.724 (0.469)
per stock L1	-1.336 (0.182)	0.306 (0.759)	-1.304 (0.192)	-1.219 (0.223)	-0.724 (0.469)	-

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 18: Diebold mariano AA (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	7.013***	1.928*	5.191***	4.303***	5.604***
		(0.0)	(0.054)	(0.0)	(0.0)	(0.0)
per group L1	-7.013***	-	-2.882***	-1.889*	-2.682***	-1.037
	(0.0)		(0.004)	(0.059)	(0.007)	(0.3)
eq by eq	-1.928*	2.882***	-	2.216**	1.606	2.521**
	(0.054)	(0.004)		(0.027)	(0.108)	(0.012)
per group	-5.191***	1.889*	-2.216**	-	-0.39	1.114
	(0.0)	(0.059)	(0.027)		(0.697)	(0.266)
per stock L2	-4.303***	2.682***	-1.606	0.39	-	2.665***
	(0.0)	(0.007)	(0.108)	(0.697)		(0.008)
per stock L1	-5.604***	1.037	-2.521**	-1.114	-2.665***	-
	(0.0)	(0.3)	(0.012)	(0.266)	(0.008)	
L = 94						
per group L2	-	7.722***	3.385***	6.701***	3.893***	6.553***
		(0.0)	(0.001)	(0.0)	(0.0)	(0.0)
per group L1	-7.722***	-	-1.785*	-0.252	-3.795***	-0.342
	(0.0)		(0.074)	(0.801)	(0.0)	(0.732)
eq by eq	-3.385***	1.785*	-	2.259**	-0.916	1.802*
	(0.001)	(0.074)		(0.024)	(0.36)	(0.072)
per group	-6.701***	0.252	-2.259**	-	-3.098***	0.014
	(0.0)	(0.801)	(0.024)		(0.002)	(0.989)
per stock L2	-3.893***	3.795***	0.916	3.098***	-	4.357***
	(0.0)	(0.0)	(0.36)	(0.002)		(0.0)
per stock L1	-6.553***	0.342	-1.802*	-0.014	-4.357***	-
	(0.0)	(0.732)	(0.072)	(0.989)	(0.0)	
L = 125						
per group L2	-	7.77***	5.567***	6.54***	4.282***	7.437***
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
per group L1	-7.77***	-	-0.568	-1.489	-4.202***	0.734
	(0.0)		(0.57)	(0.137)	(0.0)	(0.463)
eq by eq	-5.567***	0.568	-	-0.795	-3.205***	1.074
	(0.0)	(0.57)		(0.427)	(0.001)	(0.283)
per group	-6.54***	1.489	0.795	-	-2.697***	1.852*
	(0.0)	(0.137)	(0.427)		(0.007)	(0.064)
per stock L2	-4.282***	4.202***	3.205***	2.697***	-	5.274***
	(0.0)	(0.0)	(0.001)	(0.007)		(0.0)
per stock L1	-7.437***	-0.734	-1.074	-1.852*	-5.274***	-
	(0.0)	(0.463)	(0.283)	(0.064)	(0.0)	
L = 250						
per group L2	-	8.375***	7.523***	7.768***	3.263***	8.301***
		(0.0)	(0.0)	(0.0)	(0.001)	(0.0)
per group L1	-8.375***	-	-1.177	-2.171**	-5.918***	0.311
	(0.0)		(0.239)	(0.03)	(0.0)	(0.756)
eq by eq	-7.523***	1.177	-	-0.823	-5.599***	1.384
	(0.0)	(0.239)		(0.41)	(0.0)	(0.167)
per group	-7.768***	2.171**	0.823	-	-5.123***	2.053**
	(0.0)	(0.03)	(0.41)		(0.0)	(0.04)
per stock L2	-3.263***	5.918***	5.599***	5.123***	-	6.182***
	(0.001)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-8.301***	-0.311	-1.384	-2.053**	-6.182***	-
	(0.0)	(0.756)	(0.167)	(0.04)	(0.0)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 19: Diebold mariano AXP (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	1.109	-1.465	-0.041	1.026	0.964
		(0.267)	(0.143)	(0.967)	(0.305)	(0.335)
per group L1	-1.109	-	-1.698*	-1.077	-0.391	-0.174
	(0.267)		(0.09)	(0.282)	(0.696)	(0.862)
eq by eq	1.465	1.698*	-	1.938*	1.757*	1.592
	(0.143)	(0.09)		(0.053)	(0.079)	(0.112)
per group	0.041	1.077	-1.938*	-	0.977	0.949
	(0.967)	(0.282)	(0.053)		(0.329)	(0.343)
per stock L2	-1.026	0.391	-1.757*	-0.977	-	0.304
	(0.305)	(0.696)	(0.079)	(0.329)		(0.761)
per stock L1	-0.964	0.174	-1.592	-0.949	-0.304	-
	(0.335)	(0.862)	(0.112)	(0.343)	(0.761)	
L = 94						
per group L2	-	0.623	-0.491	-0.571	-0.131	0.637
		(0.533)	(0.624)	(0.568)	(0.896)	(0.524)
per group L1	-0.623	-	-0.984	-1.073	-0.637	0.168
	(0.533)		(0.325)	(0.283)	(0.525)	(0.867)
eq by eq	0.491	0.984	-	-0.281	1.183	1.067
	(0.624)	(0.325)		(0.779)	(0.237)	(0.286)
per group	0.571	1.073	0.281	-	1.404	1.157
	(0.568)	(0.283)	(0.779)		(0.161)	(0.248)
per stock L2	0.131	0.637	-1.183	-1.404	-	0.687
	(0.896)	(0.525)	(0.237)	(0.161)		(0.492)
per stock L1	-0.637	-0.168	-1.067	-1.157	-0.687	-
	(0.524)	(0.867)	(0.286)	(0.248)	(0.492)	
L = 125						
per group L2	-	0.129	-1.121	0.756	0.711	-0.247
		(0.897)	(0.262)	(0.45)	(0.477)	(0.805)
per group L1	-0.129	-	-1.689*	0.511	0.557	-0.589
	(0.897)		(0.091)	(0.609)	(0.578)	(0.556)
eq by eq	1.121	1.689*	-	1.698*	1.823*	1.735*
	(0.262)	(0.091)		(0.09)	(0.068)	(0.083)
per group	-0.756	-0.511	-1.698*	-	0.192	-0.927
	(0.45)	(0.609)	(0.09)		(0.848)	(0.354)
per stock L2	-0.711	-0.557	-1.823*	-0.192	-	-1.031
	(0.477)	(0.578)	(0.068)	(0.848)		(0.303)
per stock L1	0.247	0.589	-1.735*	0.927	1.031	-
	(0.805)	(0.556)	(0.083)	(0.354)	(0.303)	
L = 250						
per group L2	-	0.827	-0.366	-0.219	1.139	0.849
		(0.408)	(0.714)	(0.827)	(0.255)	(0.396)
per group L1	-0.827	-	-1.345	-1.298	-0.208	-0.169
	(0.408)		(0.179)	(0.195)	(0.835)	(0.866)
eq by eq	0.366	1.345	-	0.532	0.85	1.719*
	(0.714)	(0.179)		(0.595)	(0.396)	(0.086)
per group	0.219	1.298	-0.532	-	0.714	1.52
	(0.827)	(0.195)	(0.595)		(0.475)	(0.129)
per stock L2	-1.139	0.208	-0.85	-0.714	-	0.172
	(0.255)	(0.835)	(0.396)	(0.475)		(0.863)
per stock L1	-0.849	0.169	-1.719*	-1.52	-0.172	-
	(0.396)	(0.866)	(0.086)	(0.129)	(0.863)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 20: Diebold mariano AXP (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	6.879*** (0.0)	1.456 (0.146)	5.299*** (0.0)	2.812*** (0.005)	6.857*** (0.0)
per group L1	-6.879*** (0.0)	-	-5.172*** (0.0)	-3.572*** (0.0)	-4.302*** (0.0)	1.18 (0.238)
eq by eq	-1.456 (0.146)	5.172*** (0.0)	-	3.84*** (0.0)	1.302 (0.193)	6.308*** (0.0)
per group	-5.299*** (0.0)	3.572*** (0.0)	-3.84*** (0.0)	-	-2.17** (0.03)	3.624*** (0.0)
per stock L2	-2.812*** (0.005)	4.302*** (0.0)	-1.302 (0.193)	2.17** (0.03)	-	6.007*** (0.0)
per stock L1	-6.857*** (0.0)	-1.18 (0.238)	-6.308*** (0.0)	-3.624*** (0.0)	-6.007*** (0.0)	-
L = 94						
per group L2	-	6.623*** (0.0)	3.795*** (0.0)	5.982*** (0.0)	2.274** (0.023)	6.763*** (0.0)
per group L1	-6.623*** (0.0)	-	-4.281*** (0.0)	-3.084*** (0.002)	-5.108*** (0.0)	-0.315 (0.753)
eq by eq	-3.795*** (0.0)	4.281*** (0.0)	-	3.403*** (0.001)	-2.478** (0.013)	4.576*** (0.0)
per group	-5.982*** (0.0)	3.084*** (0.002)	-3.403*** (0.001)	-	-3.987*** (0.0)	2.035** (0.042)
per stock L2	-2.274** (0.023)	5.108*** (0.0)	2.478** (0.013)	3.987*** (0.0)	-	6.21*** (0.0)
per stock L1	-6.763*** (0.0)	0.315 (0.753)	-4.576*** (0.0)	-2.035** (0.042)	-6.21*** (0.0)	-
L = 125						
per group L2	-	6.695*** (0.0)	5.4*** (0.0)	6.611*** (0.0)	2.183** (0.029)	6.907*** (0.0)
per group L1	-6.695*** (0.0)	-	-3.615*** (0.0)	-2.245** (0.025)	-4.95*** (0.0)	0.11 (0.912)
eq by eq	-5.4*** (0.0)	3.615*** (0.0)	-	2.771*** (0.006)	-4.13*** (0.0)	4.278*** (0.0)
per group	-6.611*** (0.0)	2.245** (0.025)	-2.771*** (0.006)	-	-4.523*** (0.0)	2.181** (0.029)
per stock L2	-2.183** (0.029)	4.95*** (0.0)	4.13*** (0.0)	4.523*** (0.0)	-	5.352*** (0.0)
per stock L1	-6.907*** (0.0)	-0.11 (0.912)	-4.278*** (0.0)	-2.181** (0.029)	-5.352*** (0.0)	-
L = 250						
per group L2	-	6.713*** (0.0)	7.272*** (0.0)	8.158*** (0.0)	1.699* (0.089)	7.162*** (0.0)
per group L1	-6.713*** (0.0)	-	-1.336 (0.182)	-0.448 (0.654)	-5.389*** (0.0)	1.473 (0.141)
eq by eq	-7.272*** (0.0)	1.336 (0.182)	-	1.846* (0.065)	-5.666*** (0.0)	2.263** (0.024)
per group	-8.158*** (0.0)	0.448 (0.654)	-1.846* (0.065)	-	-5.949*** (0.0)	1.257 (0.209)
per stock L2	-1.699* (0.089)	5.389*** (0.0)	5.666*** (0.0)	5.949*** (0.0)	-	5.928*** (0.0)
per stock L1	-7.162*** (0.0)	-1.473 (0.141)	-2.263** (0.024)	-1.257 (0.209)	-5.928*** (0.0)	-

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 21: Diebold mariano DD (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	1.268 (0.205)	-0.978 (0.328)	-0.655 (0.512)	-0.078 (0.938)	0.737 (0.461)
per group L1	-1.268 (0.205)	-	-1.41 (0.159)	-1.764* (0.078)	-1.062 (0.288)	-0.377 (0.706)
eq by eq	0.978 (0.328)	1.41 (0.159)	-	0.877 (0.38)	1.287 (0.198)	1.604 (0.109)
per group	0.655 (0.512)	1.764* (0.078)	-0.877 (0.38)	-	0.326 (0.745)	1.14 (0.254)
per stock L2	0.078 (0.938)	1.062 (0.288)	-1.287 (0.198)	-0.326 (0.745)	-	1.246 (0.213)
per stock L1	-0.737 (0.461)	0.377 (0.706)	-1.604 (0.109)	-1.14 (0.254)	-1.246 (0.213)	-
L = 94						
per group L2	-	2.449** (0.014)	-0.334 (0.739)	0.993 (0.321)	0.235 (0.815)	1.341 (0.18)
per group L1	-2.449** (0.014)	-	-1.108 (0.268)	-2.039** (0.042)	-0.912 (0.362)	-0.449 (0.654)
eq by eq	0.334 (0.739)	1.108 (0.268)	-	0.629 (0.529)	0.71 (0.478)	1.22 (0.222)
per group	-0.993 (0.321)	2.039** (0.042)	-0.629 (0.529)	-	-0.164 (0.87)	0.811 (0.417)
per stock L2	-0.235 (0.815)	0.912 (0.362)	-0.71 (0.478)	0.164 (0.87)	-	1.004 (0.315)
per stock L1	-1.341 (0.18)	0.449 (0.654)	-1.22 (0.222)	-0.811 (0.417)	-1.004 (0.315)	-
L = 125						
per group L2	-	0.099 (0.921)	-1.347 (0.178)	0.096 (0.924)	-0.543 (0.587)	-0.606 (0.545)
per group L1	-0.099 (0.921)	-	-1.146 (0.252)	-0.031 (0.976)	-0.558 (0.577)	-1.152 (0.25)
eq by eq	1.347 (0.178)	1.146 (0.252)	-	1.559 (0.119)	1.251 (0.211)	0.833 (0.405)
per group	-0.096 (0.924)	0.031 (0.976)	-1.559 (0.119)	-	-0.567 (0.571)	-0.59 (0.555)
per stock L2	0.543 (0.587)	0.558 (0.577)	-1.251 (0.211)	0.567 (0.571)	-	-0.406 (0.684)
per stock L1	0.606 (0.545)	1.152 (0.25)	-0.833 (0.405)	0.59 (0.555)	0.406 (0.684)	-
L = 250						
per group L2	-	1.329 (0.184)	-0.671 (0.502)	1.82* (0.069)	-0.294 (0.769)	1.372 (0.17)
per group L1	-1.329 (0.184)	-	-1.52 (0.129)	0.368 (0.713)	-1.423 (0.155)	0.167 (0.867)
eq by eq	0.671 (0.502)	1.52 (0.129)	-	1.856* (0.064)	0.606 (0.545)	1.143 (0.253)
per group	-1.82* (0.069)	-0.368 (0.713)	-1.856* (0.064)	-	-2.156** (0.031)	-0.031 (0.975)
per stock L2	0.294 (0.769)	1.423 (0.155)	-0.606 (0.545)	2.156** (0.031)	-	1.537 (0.125)
per stock L1	-1.372 (0.17)	-0.167 (0.867)	-1.143 (0.253)	0.031 (0.975)	-1.537 (0.125)	-

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 22: Diebold mariano DD (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	9.107***	2.066**	6.945***	4.37***	8.274***
		(0.0)	(0.039)	(0.0)	(0.0)	(0.0)
per group L1	-9.107***	-	-3.386***	-3.902***	-4.82***	0.411
	(0.0)		(0.001)	(0.0)	(0.0)	(0.681)
eq by eq	-2.066**	3.386***	-	2.088**	1.075	3.986***
	(0.039)	(0.001)		(0.037)	(0.282)	(0.0)
per group	-6.945***	3.902***	-2.088**	-	-1.886*	3.29***
	(0.0)	(0.0)	(0.037)		(0.059)	(0.001)
per stock L2	-4.37***	4.82***	-1.075	1.886*	-	6.008***
	(0.0)	(0.0)	(0.282)	(0.059)		(0.0)
per stock L1	-8.274***	-0.411	-3.986***	-3.29***	-6.008***	-
	(0.0)	(0.681)	(0.0)	(0.001)	(0.0)	
L = 94						
per group L2	-	9.196***	5.138***	8.01***	5.297***	8.847***
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
per group L1	-9.196***	-	-3.083***	-3.418***	-5.782***	1.333
	(0.0)		(0.002)	(0.001)	(0.0)	(0.183)
eq by eq	-5.138***	3.083***	-	1.345	-1.836*	4.027***
	(0.0)	(0.002)		(0.179)	(0.067)	(0.0)
per group	-8.01***	3.418***	-1.345	-	-3.593***	3.881***
	(0.0)	(0.001)	(0.179)		(0.0)	(0.0)
per stock L2	-5.297***	5.782***	1.836*	3.593***	-	6.731***
	(0.0)	(0.0)	(0.067)	(0.0)		(0.0)
per stock L1	-8.847***	-1.333	-4.027***	-3.881***	-6.731***	-
	(0.0)	(0.183)	(0.0)	(0.0)	(0.0)	
L = 125						
per group L2	-	8.622***	6.218***	8.253***	4.972***	8.475***
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
per group L1	-8.622***	-	-1.836*	-1.103	-4.922***	1.699*
	(0.0)		(0.067)	(0.27)	(0.0)	(0.089)
eq by eq	-6.218***	1.836*	-	1.423	-3.306***	3.232***
	(0.0)	(0.067)		(0.155)	(0.001)	(0.001)
per group	-8.253***	1.103	-1.423	-	-4.69***	2.171**
	(0.0)	(0.27)	(0.155)		(0.0)	(0.03)
per stock L2	-4.972***	4.922***	3.306***	4.69***	-	5.991***
	(0.0)	(0.0)	(0.001)	(0.0)		(0.0)
per stock L1	-8.475***	-1.699*	-3.232***	-2.171**	-5.991***	-
	(0.0)	(0.089)	(0.001)	(0.03)	(0.0)	
L = 250						
per group L2	-	8.569***	8.169***	9.036***	4.227***	8.406***
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
per group L1	-8.569***	-	-0.259	-0.271	-6.814***	1.541
	(0.0)		(0.796)	(0.786)	(0.0)	(0.124)
eq by eq	-8.169***	0.259	-	0.12	-7.153***	1.21
	(0.0)	(0.796)		(0.904)	(0.0)	(0.226)
per group	-9.036***	0.271	-0.12	-	-7.292***	1.383
	(0.0)	(0.786)	(0.904)		(0.0)	(0.167)
per stock L2	-4.227***	6.814***	7.153***	7.292***	-	6.943***
	(0.0)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-8.406***	-1.541	-1.21	-1.383	-6.943***	-
	(0.0)	(0.124)	(0.226)	(0.167)	(0.0)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 23: Diebold mariano GE (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	1.724*	-1.819*	-1.532	1.006	0.934
		(0.085)	(0.069)	(0.126)	(0.315)	(0.35)
per group L1	-1.724*	-	-1.994**	-1.783*	-0.144	-1.334
	(0.085)		(0.046)	(0.075)	(0.885)	(0.182)
eq by eq	1.819*	1.994**	-	1.23	1.597	2.022**
	(0.069)	(0.046)		(0.219)	(0.11)	(0.043)
per group	1.532	1.783*	-1.23	-	1.35	1.689*
	(0.126)	(0.075)	(0.219)		(0.177)	(0.091)
per stock L2	-1.006	0.144	-1.597	-1.35	-	-0.578
	(0.315)	(0.885)	(0.11)	(0.177)		(0.564)
per stock L1	-0.934	1.334	-2.022**	-1.689*	0.578	-
	(0.35)	(0.182)	(0.043)	(0.091)	(0.564)	
L = 94						
per group L2	-	1.888*	-1.146	-0.489	-1.084	1.191
		(0.059)	(0.252)	(0.625)	(0.279)	(0.234)
per group L1	-1.888*	-	-1.418	-1.831*	-1.762*	-0.621
	(0.059)		(0.156)	(0.067)	(0.078)	(0.534)
eq by eq	1.146	1.418	-	1.066	1.09	1.389
	(0.252)	(0.156)		(0.286)	(0.276)	(0.165)
per group	0.489	1.831*	-1.066	-	-0.685	1.183
	(0.625)	(0.067)	(0.286)		(0.493)	(0.237)
per stock L2	1.084	1.762*	-1.09	0.685	-	1.679*
	(0.279)	(0.078)	(0.276)	(0.493)		(0.093)
per stock L1	-1.191	0.621	-1.389	-1.183	-1.679*	-
	(0.234)	(0.534)	(0.165)	(0.237)	(0.093)	
L = 125						
per group L2	-	1.721*	-1.169	-0.647	1.458	1.879*
		(0.085)	(0.242)	(0.517)	(0.145)	(0.06)
per group L1	-1.721*	-	-1.72*	-1.503	-1.193	-1.097
	(0.085)		(0.086)	(0.133)	(0.233)	(0.273)
eq by eq	1.169	1.72*	-	1.159	1.52	1.808*
	(0.242)	(0.086)		(0.247)	(0.129)	(0.071)
per group	0.647	1.503	-1.159	-	1.179	1.55
	(0.517)	(0.133)	(0.247)		(0.238)	(0.121)
per stock L2	-1.458	1.193	-1.52	-1.179	-	0.624
	(0.145)	(0.233)	(0.129)	(0.238)		(0.533)
per stock L1	-1.879*	1.097	-1.808*	-1.55	-0.624	-
	(0.06)	(0.273)	(0.071)	(0.121)	(0.533)	
L = 250						
per group L2	-	1.047	-1.51	0.737	-1.274	0.181
		(0.295)	(0.131)	(0.461)	(0.203)	(0.856)
per group L1	-1.047	-	-1.878*	0.04	-1.612	-0.505
	(0.295)		(0.061)	(0.968)	(0.107)	(0.614)
eq by eq	1.51	1.878*	-	2.122**	1.217	1.41
	(0.131)	(0.061)		(0.034)	(0.224)	(0.159)
per group	-0.737	-0.04	-2.122**	-	-1.658*	-0.353
	(0.461)	(0.968)	(0.034)		(0.098)	(0.724)
per stock L2	1.274	1.612	-1.217	1.658*	-	1.026
	(0.203)	(0.107)	(0.224)	(0.098)		(0.305)
per stock L1	-0.181	0.505	-1.41	0.353	-1.026	-
	(0.856)	(0.614)	(0.159)	(0.724)	(0.305)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 24: Diebold mariano GE (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	7.312***	2.606***	4.652***	3.497***	7.144***
		(0.0)	(0.009)	(0.0)	(0.0)	(0.0)
per group L1	-7.312***	-	-3.96***	-3.321***	-4.63***	0.943
	(0.0)		(0.0)	(0.001)	(0.0)	(0.346)
eq by eq	-2.606***	3.96***	-	2.147**	0.621	4.617***
	(0.009)	(0.0)		(0.032)	(0.534)	(0.0)
per group	-4.652***	3.321***	-2.147**	-	-0.789	3.716***
	(0.0)	(0.001)	(0.032)		(0.43)	(0.0)
per stock L2	-3.497***	4.63***	-0.621	0.789	-	6.026***
	(0.0)	(0.0)	(0.534)	(0.43)		(0.0)
per stock L1	-7.144***	-0.943	-4.617***	-3.716***	-6.026***	-
	(0.0)	(0.346)	(0.0)	(0.0)	(0.0)	
L = 94						
per group L2	-	8.617***	4.956***	8.466***	2.679***	7.761***
		(0.0)	(0.0)	(0.0)	(0.007)	(0.0)
per group L1	-8.617***	-	-3.058***	-1.703*	-5.884***	-1.197
	(0.0)		(0.002)	(0.089)	(0.0)	(0.231)
eq by eq	-4.956***	3.058***	-	2.662***	-3.63***	3.133***
	(0.0)	(0.002)		(0.008)	(0.0)	(0.002)
per group	-8.466***	1.703*	-2.662***	-	-5.48***	-0.021
	(0.0)	(0.089)	(0.008)		(0.0)	(0.983)
per stock L2	-2.679***	5.884***	3.63***	5.48***	-	6.305***
	(0.007)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-7.761***	1.197	-3.133***	0.021	-6.305***	-
	(0.0)	(0.231)	(0.002)	(0.983)	(0.0)	
L = 125						
per group L2	-	9.316***	7.647***	9.464***	5.383***	8.854***
		(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
per group L1	-9.316***	-	-1.488	-0.878	-5.55***	0.211
	(0.0)		(0.137)	(0.38)	(0.0)	(0.833)
eq by eq	-7.647***	1.488	-	1.2	-5.02***	1.907*
	(0.0)	(0.137)		(0.23)	(0.0)	(0.057)
per group	-9.464***	0.878	-1.2	-	-5.43***	0.853
	(0.0)	(0.38)	(0.23)		(0.0)	(0.394)
per stock L2	-5.383***	5.55***	5.02***	5.43***	-	6.192***
	(0.0)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-8.854***	-0.211	-1.907*	-0.853	-6.192***	-
	(0.0)	(0.833)	(0.057)	(0.394)	(0.0)	
L = 250						
per group L2	-	7.348***	6.051***	7.923***	1.71*	6.62***
		(0.0)	(0.0)	(0.0)	(0.087)	(0.0)
per group L1	-7.348***	-	-3.727***	-1.853*	-7.541***	-2.117**
	(0.0)		(0.0)	(0.064)	(0.0)	(0.034)
eq by eq	-6.051***	3.727***	-	2.605***	-6.926***	2.053**
	(0.0)	(0.0)		(0.009)	(0.0)	(0.04)
per group	-7.923***	1.853*	-2.605***	-	-7.665***	-0.6
	(0.0)	(0.064)	(0.009)		(0.0)	(0.549)
per stock L2	-1.71*	7.541***	6.926***	7.665***	-	7.562***
	(0.087)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-6.62***	2.117**	-2.053**	0.6	-7.562***	-
	(0.0)	(0.034)	(0.04)	(0.549)	(0.0)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 25: Diebold mariano KO (MSE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	0.093	-1.316	-1.021	-0.607	0.122
		(0.926)	(0.188)	(0.307)	(0.544)	(0.903)
per group L1	-0.093	-	-1.084	-0.532	-1.253	0.121
	(0.926)		(0.278)	(0.594)	(0.21)	(0.904)
eq by eq	1.316	1.084	-	1.13	0.609	1.073
	(0.188)	(0.278)		(0.259)	(0.543)	(0.283)
per group	1.021	0.532	-1.13	-	-0.113	0.509
	(0.307)	(0.594)	(0.259)		(0.91)	(0.611)
per stock L2	0.607	1.253	-0.609	0.113	-	2.024**
	(0.544)	(0.21)	(0.543)	(0.91)		(0.043)
per stock L1	-0.122	-0.121	-1.073	-0.509	-2.024**	-
	(0.903)	(0.904)	(0.283)	(0.611)	(0.043)	
L = 94						
per group L2	-	0.141	-1.213	-1.505	-0.177	0.871
		(0.888)	(0.225)	(0.133)	(0.859)	(0.384)
per group L1	-0.141	-	-0.97	-1.181	-0.352	2.055**
	(0.888)		(0.332)	(0.238)	(0.725)	(0.04)
eq by eq	1.213	0.97	-	-0.762	0.908	1.519
	(0.225)	(0.332)		(0.446)	(0.364)	(0.129)
per group	1.505	1.181	0.762	-	1.115	1.716*
	(0.133)	(0.238)	(0.446)		(0.265)	(0.086)
per stock L2	0.177	0.352	-0.908	-1.115	-	1.183
	(0.859)	(0.725)	(0.364)	(0.265)		(0.237)
per stock L1	-0.871	-2.055**	-1.519	-1.716*	-1.183	-
	(0.384)	(0.04)	(0.129)	(0.086)	(0.237)	
L = 125						
per group L2	-	1.14	-0.908	-1.001	0.815	1.27
		(0.254)	(0.364)	(0.317)	(0.415)	(0.204)
per group L1	-1.14	-	-1.268	-1.536	-0.004	0.281
	(0.254)		(0.205)	(0.125)	(0.996)	(0.779)
eq by eq	0.908	1.268	-	0.26	1.979**	1.551
	(0.364)	(0.205)		(0.795)	(0.048)	(0.121)
per group	1.001	1.536	-0.26	-	1.575	1.643
	(0.317)	(0.125)	(0.795)		(0.115)	(0.101)
per stock L2	-0.815	0.004	-1.979**	-1.575	-	0.15
	(0.415)	(0.996)	(0.048)	(0.115)		(0.881)
per stock L1	-1.27	-0.281	-1.551	-1.643	-0.15	-
	(0.204)	(0.779)	(0.121)	(0.101)	(0.881)	
L = 250						
per group L2	-	1.286	-0.479	0.462	0.52	0.65
		(0.199)	(0.632)	(0.644)	(0.603)	(0.516)
per group L1	-1.286	-	-1.449	-0.935	-0.164	0.049
	(0.199)		(0.147)	(0.35)	(0.869)	(0.961)
eq by eq	0.479	1.449	-	1.13	0.664	0.789
	(0.632)	(0.147)		(0.259)	(0.506)	(0.43)
per group	-0.462	0.935	-1.13	-	0.238	0.41
	(0.644)	(0.35)	(0.259)		(0.812)	(0.682)
per stock L2	-0.52	0.164	-0.664	-0.238	-	0.601
	(0.603)	(0.869)	(0.506)	(0.812)		(0.548)
per stock L1	-0.65	-0.049	-0.789	-0.41	-0.601	-
	(0.516)	(0.961)	(0.43)	(0.682)	(0.548)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

Table 26: Diebold mariano KO (QLIKE)

	per group L2	per group L1	eq by eq	per group	per stock L2	per stock L1
L = 63						
per group L2	-	5.635***	2.845***	4.707***	1.598	5.872***
		(0.0)	(0.004)	(0.0)	(0.11)	(0.0)
per group L1	-5.635***	-	-3.872***	-2.611***	-5.375***	1.109
	(0.0)		(0.0)	(0.009)	(0.0)	(0.267)
eq by eq	-2.845***	3.872***	-	2.786***	-1.705*	4.755***
	(0.004)	(0.0)		(0.005)	(0.088)	(0.0)
per group	-4.707***	2.611***	-2.786***	-	-3.48***	2.87***
	(0.0)	(0.009)	(0.005)		(0.001)	(0.004)
per stock L2	-1.598	5.375***	1.705*	3.48***	-	7.318***
	(0.11)	(0.0)	(0.088)	(0.001)		(0.0)
per stock L1	-5.872***	-1.109	-4.755***	-2.87***	-7.318***	-
	(0.0)	(0.267)	(0.0)	(0.004)	(0.0)	
L = 94						
per group L2	-	8.455***	7.346***	7.749***	2.934***	9.053***
		(0.0)	(0.0)	(0.0)	(0.003)	(0.0)
per group L1	-8.455***	-	-2.436**	-2.152**	-5.508***	0.381
	(0.0)		(0.015)	(0.032)	(0.0)	(0.703)
eq by eq	-7.346***	2.436**	-	0.985	-4.259***	2.987***
	(0.0)	(0.015)		(0.325)	(0.0)	(0.003)
per group	-7.749***	2.152**	-0.985	-	-4.271***	2.159**
	(0.0)	(0.032)	(0.325)		(0.0)	(0.031)
per stock L2	-2.934***	5.508***	4.259***	4.271***	-	6.036***
	(0.003)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-9.053***	-0.381	-2.987***	-2.159**	-6.036***	-
	(0.0)	(0.703)	(0.003)	(0.031)	(0.0)	
L = 125						
per group L2	-	8.567***	7.453***	7.578***	1.621	7.897***
		(0.0)	(0.0)	(0.0)	(0.105)	(0.0)
per group L1	-8.567***	-	-2.558**	-2.288**	-6.873***	-1.698*
	(0.0)		(0.011)	(0.022)	(0.0)	(0.09)
eq by eq	-7.453***	2.558**	-	0.9	-6.083***	1.416
	(0.0)	(0.011)		(0.368)	(0.0)	(0.157)
per group	-7.578***	2.288**	-0.9	-	-5.728***	0.525
	(0.0)	(0.022)	(0.368)		(0.0)	(0.599)
per stock L2	-1.621	6.873***	6.083***	5.728***	-	6.764***
	(0.105)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-7.897***	1.698*	-1.416	-0.525	-6.764***	-
	(0.0)	(0.09)	(0.157)	(0.599)	(0.0)	
L = 250						
per group L2	-	7.985***	8.609***	8.428***	0.929	7.212***
		(0.0)	(0.0)	(0.0)	(0.353)	(0.0)
per group L1	-7.985***	-	-1.811*	-0.793	-7.652***	-2.354**
	(0.0)		(0.07)	(0.428)	(0.0)	(0.019)
eq by eq	-8.609***	1.811*	-	1.583	-8.637***	-0.44
	(0.0)	(0.07)		(0.113)	(0.0)	(0.66)
per group	-8.428***	0.793	-1.583	-	-7.881***	-1.449
	(0.0)	(0.428)	(0.113)		(0.0)	(0.148)
per stock L2	-0.929	7.652***	8.637***	7.881***	-	7.213***
	(0.353)	(0.0)	(0.0)	(0.0)		(0.0)
per stock L1	-7.212***	2.354**	0.44	1.449	-7.213***	-
	(0.0)	(0.019)	(0.66)	(0.148)	(0.0)	

DM statistics with a p-value smaller than 0.1, 0.05 and 0.01 are indicated by *, ** and *** respectively.

A.5 Loss Ratios (DD)

Figure 33: $\frac{QLIKE_j}{QLIKE_{eq\ by\ eq}}$ over time. Comparing model j to HAR equation by equation for rolling window of 63 (DD).

A.6 LSTM Diebold Mariano

Table 27: Diebold Mariano statistics for the per group ℓ_1 HAR and equation by equation HAR, rolling window of 250

	per group ℓ_1	eq by eq	LSTM
per group ℓ_1	-	-1.133 (0.258)	-0.534 (0.594)
eq by eq	1.133 (0.258)	-	-0.242 (0.809)
LSTM	0.534 (0.594)	0.242 (0.809)	-

Other tables and figures

Available upon request.