Performance evaluation of volatility estimation methods for Exabel

Øyvind Grotmol¹*, Martin Jullum², Kjersti Aas², Michael Scheuerer², ...

Abstract
Quantifying both historic and future volatility is key in portfolio risk management. This note presents and compare estimation strategies for volatility estimation in an estimation universe consisting on 28 629 unique companies from February 2010 to April 2021, with 858 different portfolios. The estimation methods are compared in terms of how they rank the volatility of the different subsets of portfolios. The overall best performing approach estimates volatility from direct entity returns using a GARCH model for variance estimation.

¹Exabel, Oslo, Norway
²Department for Statistical modeling, Machine learning and Artificial intelligence, Norsk Regnesentral, Oslo, Norway
*Corresponding author: grotmol@exabel.com

1. Introduction
The aim of this note is to describe and summarize a quantitative study comparing strategies for forecasting investment risk. The study concentrates on the investment risk of a portfolio relative to other portfolios, and uses volatility as the risk measure. Denote by \( r_{k,t} \), the log-return of a company \( k \in K := \{1, \ldots, K\} \) on day \( t \) (hereafter for simplicity referred to as the return or company return). The (daily) volatility of company \( k \) during a time period \( T \) of \( |T| \) trading days is then defined as the sample standard deviation

\[
vol_k(T) = \sqrt{\frac{1}{|T|-1} \sum_{t \in T} \left( r_{k,t} - \frac{1}{|T|} \sum_{t \in T} r_{k,t} \right)^2},
\]

with an analogous definition for portfolios. While it is common to scale \( vol_k(T) \) to reflect the volatility over \( T \) or a full year, we omit that here as we deem daily volatility more interpretable.

As a short-term investment tool, estimates of future volatility is more useful than computations of historic volatility. In this study, we will therefore consider methods for estimating (daily) volatility over consecutive test periods \( T \) of duration 3 months. We will consider both direct approaches estimating the volatility using the historic observed returns of the companies, and indirect approaches using Exabel’s factor model. We will also consider two different weighting strategies for data back in time: A naive window based approach which puts equal weight on all observations \( q \) months back in time, and an approach which puts more weight on the most recent observations according to an estimated GARCH model. Exabel’s factor model is briefly described in Section 2, while all different volatility estimation methods are precisely defined in Section 3.

The estimation universe we will be working with here consists of a total of 28 629 companies from all over the world, with company information and returns ranging from February 2010 to April 2021. As some companies have only existed for a portion of the full time period, the trading days are not the same for all companies. Due to this, and other missing data complications, some data pre-processing is required. We describe the removals, imputations and other actions we take to circumvent missing data issues in Section 4.

The volatilities are estimated for a range of portfolios, consisting of different companies from the estimation universe. The portfolios are defined based on the factors described in Section 2 to reflect investment strategies following the different factors. In total we are estimating volatilities for 858 different portfolios over 97 different 3-month test periods \( T \). See Section 5 for details.

The estimation methods will be compared and measured according to their ability to rank portfolios based on their
risk. We will do this by computing Kendall’s $\tau$ [1] between the estimated volatilities and the target volatilities of a set of portfolios (after the latter have been observed). Details about the performance evaluation is provided in Section 6. Different sets of portfolios are considered, and their results are presented in Section 7. Some conclusions are provided in Section 8.

## 2. Factor model

With a factor model we would like to explain the returns of all companies in $\mathbb{K}$ as well as possible through a small number of interpretable factors, i.e. attributes of a company which appear to influence its return. The factor loading $X_{k,t}^s$ of company $k$ quantifies its exposure to a particular factor $s$ at day $t$. Three different types of factor loadings are considered in Exabel’s factor model:

- **Style-loadings** $X_{k,t}^s$, $s \in \mathbb{S}$ (set of style factors)
- **Country-loadings** $X_{k,t}^c$, $c \in \mathbb{C}$ (set of countries)
- **Industry-loadings** $X_{k,t}^i$, $i \in \mathbb{I}$ (set of industries)

Using a cross-sectional regression model, we then want to estimate (separately for each day) factor returns $f_t^s$, $s \in \mathbb{S}$, $f_t^c$, $c \in \mathbb{C}$, and $f_t^i$, $i \in \mathbb{I}$ together with the overall market return $f_t^m$:

$$r_{k,t} = f_t^m + \sum_{s \in \mathbb{S}} X_{k,t}^s f_t^s + \sum_{c \in \mathbb{C}} X_{k,t}^c f_t^c + \sum_{i \in \mathbb{I}} X_{k,t}^i f_t^i + \epsilon_{k,t}, \quad (2)$$

where $\epsilon_{k,t} \sim \mathcal{N}(0, \sigma^2_{k,t})$ is the residual term that accounts for the fraction of the return $r_{k,t}$ not explained by the company’s exposure to the different factors. In Exabel’s factor model, the factor loadings $X_{k,t}^s$ are updated on a monthly basis, i.e. they are constant within each month. All style factors $X_{k,t}^s$ are market cap centered, meaning that they satisfy $\sum_{k \in \mathbb{K}} m_{c_{k,t}} X_{k,t}^s$, where $m_{c_{k,t}}$ denotes the market capitalization of company $k$. Furthermore, the estimates $\hat{f}_t^s$ for the factor $f_t^s$ are obtained using a constrained weighted least squares approach where the companies are weighted by $m_{c_{k,t}}$. See [2] for further details about Exabel’s factor model.

In total there are $l = 77$ different factors. Using matrix notation, (2) can be written as

$$r_{k,t} = X_{k,t}^T \mathbf{f}_t + \epsilon_{k,t}, \quad (3)$$

where $\mathbf{f}_t$ is the $l$-dimensional vector of all factor returns, and $X_{k,t}$ is the $l$-dimensional vector of all factor loadings (we define $X_{k,t}^m = 1$).

## 3. Volatility estimation and computation

As mentioned, the purpose of this note is to compare different methods for estimating the future volatility of portfolios. In this section we extend the single company volatility definition in (1) to the volatility of a portfolio, and describe the different methods we will be using for estimating this ahead of time. All quantities in this section are specific to a portfolio $\mathbb{P} = \{k_1, \ldots, k_p\}$ of $p$ companies, with corresponding portfolio weights $w_{k,t}$ for company $k$ at day $t$. Every portfolio has the property that $\sum_{k=1}^p |w_{k,t}| = 1$ for all $t$. To simplify notation, we will, however, often omit $\mathbb{P}$. When it is clear from the context, we will also sometimes omit the test period $T$ from our notation.

### 3.1 Target portfolio volatility

We start by defining the target portfolio volatility, i.e. the future volatility that the different estimation methods will attempt to predict. As mentioned in the introduction, we will consider consecutive 3-months test periods $T$, and shall for simplicity assume that all portfolios have fixed portfolio weights $w_{k,t}$ within each of the test periods $T$, i.e. $w_{k,t} = w_k$ for all $t \in T, k \in \mathbb{P}$. With these simplifications in place, we define the target portfolio volatility for time period $T$ as

$$\text{vol}(T) = \sqrt{\mathbf{w}^T \Sigma(T) \mathbf{w}}, \quad (4)$$

where $\mathbf{w} = (w_1, \ldots, w_p)^T$ is the vector of portfolio weights in time period $T$. Furthermore, $\Sigma(T)$ is the $p \times p$-dimensional sample covariance matrix of the daily returns in time period $T$, for the companies in $\mathbb{P}$.

An alternative to the portfolio volatility definition in (4) would be to compute the daily returns of the portfolio first, and then compute the variance of these returns across $T$. This is, however, deemed more sensitive to missing company returns, as missing a single return for one day with nonzero weight would leave the portfolio return for that day undefined. See Section 4 for how we handle missing company returns when estimating the covariance matrix $\Sigma(T)$.

### 3.2 Estimating future portfolio volatility

In this section we discuss methods to estimate the target portfolio volatility in Section 3.1 using historical data. We will discuss two types of approaches. The first uses the actual returns directly, while the second is an indirect approach which uses the factor model discussed in Section 2. Both approaches uses the portfolio volatility formula in (4) to estimate the future portfolio volatility, except that the (future) sample covariance matrix $\Sigma(T)$ (which is unknown at the estimation time point) is replaced with an estimate. In the subsequent subsections, we first present general formulae for the two types of estimating approaches, before we provide precise estimation methods for the unknown quantities.

#### 3.2.1 The direct return approach

For the direct (company) return approach, the estimate of the future covariance matrix takes the form

$$\hat{\Sigma}_T = \hat{D}_T \hat{R}_T \hat{D}_T,$$

where $\hat{R}_T$ is an estimate of the $p \times p$ dimensional future correlation matrix of the returns, and $\hat{D}_T$ is a $p \times p$ diagonal matrix with estimates of the standard deviation of the company returns on the diagonal. The portfolio volatility estimate using
the direct return approach (r) thus takes the form
\[ \widehat{\text{vol}}_t(T) = \sqrt{w^\top \Sigma_t w}. \] (5)

### 3.2.2 The indirect factor based approach
To write up a general volatility expression for the indirect factor based approach (f), recall the matrix formulation of the estimated factor model in (3). With estimated parameters, this factor model takes the form \( \tilde{\mathbf{r}}_{k,t} = \mathbf{X}_{1,t}^\top \mathbf{f}_t \), where \( \mathbf{f}_t \) is the estimated analogue of \( \mathbf{f} \). Let us also denote by \( L \) the \( l \times p \) dimensional matrix obtained by columnwise concatenation of the \( l \)-dimensional vectors \( \mathbf{X}_{1,t}, \mathbf{X}_{2,t}, \ldots, \mathbf{X}_{p,t} \). For the indirect, factor based approach, the estimate of the future covariance may then be written as
\[ \widehat{\Sigma}_t = L^\top \widehat{\Sigma}_t^f L = L^\top \widehat{D}_t \widehat{R}_t L, \] (6)
where \( \widehat{\Sigma}_t^f \) and \( \widehat{R}_t \) are estimates of the \( l \times l \) dimensional future covariance and correlation matrices of the factors \( \mathbf{f}_t \). The \( \widehat{D}_t \) is an \( l \times l \) diagonal matrix with estimates of the (future) standard deviation of the factors at the diagonal.

When using the factor model to estimate the volatility, we also need to account for the idiosyncratic variance \( \sigma^2_{\epsilon} \) of the residual term \( \epsilon_{k,t} \). As the factor model assumes that the error terms are independent, the estimated portfolio variance then becomes \( \sum_{k \in \mathbb{P}} w_{k,t}^2 \widehat{\sigma}^2_{\epsilon} \), which in matrix form may be written \( w^\top \Sigma_t w \), where \( \sigma^2_{\epsilon} \) is the \( p \times p \)-dimensional diagonal matrix with estimates for the variances of the companies in the portfolio \( \mathbb{P} \) on the diagonal. Combining the insertion of \( \Sigma_t^f \) for \( \Sigma(T) \) in (4) with the additional residual variance term, we get the following general formula for the portfolio volatility estimate using the factor based approach (f):
\[ \widehat{\text{vol}}_t(T) = \sqrt{w^\top \Sigma_t w + w^\top \Sigma_t^f w}. \] (7)

Furthermore, by pre-computing a \( l \times p \)-dimensional portfolio factor loading matrix \( \mathbf{L}_p = w^\top L^\top \), we don’t need to compute the full \( p \times p \) dimensional \( \Sigma_t \), but only the \( l \times l \)-dimensional \( \Sigma_t^f \). Thus, (7) can also be written as
\[ \widehat{\text{vol}}_t(T) = \sqrt{\mathbf{L}_p^\top \Sigma_t^f \mathbf{L}_p + w^\top \sigma^2_{\epsilon} w}, \] (8)
which is computationally more efficient for portfolios with many companies.

### 3.2.3 Estimating the unknown quantities
The formulae for the estimated volatility in (5) and (8) requires specification of the estimated covariance matrices \( \Sigma = \widehat{D}_t \widehat{R}_t \widehat{D}_t \) and \( \Sigma_t^f = \widehat{D}_t \widehat{R}_t \widehat{D}_t \), respectively. We will provide two different methods for estimating each of them based on historic data, in addition to an estimation method for the residual variances in the factor model.

### Naive window estimation method
The simplest way to estimate a covariance matrix for future companies/factors is to use the sample covariance matrix for data in a window covering a certain number of days back in time. This is a naive estimation method which weights all company/factor returns within the window equally. If the window is too long one risks that the oldest company/factor returns are not representative, while a too short window risks giving unstable estimates as there is not enough data. Let us denote by \( \hat{\Sigma}_{f,\text{naive}}(q) \) and \( \hat{\Sigma}_{r,\text{naive}}(q) \) the company/factor return sample covariance matrices with a window covering \( q \) months prior to the test period \( T \).

#### GARCH model for variance estimation
The naive method above has the problem of requiring a rather long period (large \( q \)) in order to get stable and good estimates, while at the same time such a long period makes it react rather slow to changes in the volatility over time. One way to overcome this is to put more weight on the more recent observations than the older ones. Our approach in that direction is to modify the estimation of the covariance matrices by combining the naive window method for the correlation structure \( \widehat{R}_t \) and \( \widehat{R}_t \) with separate univariate GARCH models [3] for the variances of the companies \( (\widehat{D}_t \) and \( \widehat{D}_t \) \). There exists multivariate GARCH models that could estimate covariance structure directly, but in the high dimensional settings present here, that would be computationally infeasible. GARCH stands for generalized autoregressive conditional heteroskedasticity, and is essentially an autoregressive moving average (ARMA) model on the error variance of a time series model. Here we will just assume a constant mean model, and use both autoregressive (AR) and moving average (MA) terms of order 1 for the variance. This is the most commonly used GARCH form within financial modelling [3]. For company \( k \), this GARCH model takes the form
\[
\epsilon_{k,t} = \mu_k + \epsilon_{G,k,t},
\]
\[
\epsilon_{G,k,t} | \epsilon_{G,k,t-1}, \Omega^2_{G,k,t-1} \sim N(0, \Omega^2_{G,k,t-1}),
\]
\[
\Omega^2_{G,k,t} = \alpha_k + \alpha_k \epsilon^2_{G,k,t-1} + \beta_k \Omega^2_{G,k,t-1},
\]
where \( \mu_k, \alpha_k, \beta_k \) are parameters that have to be estimated. For the factor model, the formulation is completely analogous, except that company returns \( r_{k,t} \) are replaced by the fitted factor returns \( \tilde{f}_t \). When estimating the parameters in these model specifications, we will always use a history of 3 years prior to the test period \( T \). The diagonal matrices \( \mathbf{D}_t = \mathbf{D}_t \mathbf{GARCH} \) and \( \mathbf{D}_t = \mathbf{D}_t \mathbf{GARCH} \) are then filled with the modelled company/factor return variances for the latest observed day. Combining these diagonal variance matrices with naive window based estimators for the correlation structure (i.e. \( \widehat{R}_{c,\text{naive}}(q) \) and \( \widehat{R}_{f,\text{naive}}(q) \)), we get
\[
\hat{\Sigma}_{f,\text{GARCH}}(q) = \widehat{D}_t \mathbf{GARCH} \widehat{R}_{c,\text{naive}}(q) \widehat{D}_t \mathbf{GARCH},
\]
\[
\hat{\Sigma}_{r,\text{GARCH}}(q) = \widehat{D}_t \mathbf{GARCH} \widehat{R}_{f,\text{naive}}(q) \widehat{D}_t \mathbf{GARCH}.
\]
basic sample variance of the factor residuals $\varepsilon_{k,t}$ for $t$ in a window $q$ months prior to $T$ as an estimate of $\sigma^2_{k,t}$, $k \in \mathbb{P}$. We use the same window length $q$ as for $\hat{\Sigma}_{t, \text{GARCH}}^f(q)$ and $\hat{\Sigma}_{t, \text{GARCH}}(q)$, and denote the estimate of the residual variance diagonal matrix $\hat{\Sigma}_P^2$ by $\hat{\Sigma}_P^2 = \hat{\Sigma}_{P, \text{naive}}^2(q)$.

3.2.4 Final estimation schemes
To summarize, we have the following types of estimation schemes for the portfolio variance

$$\hat{\sigma}_{t, \text{GARCH}}(q) = \sqrt{\mathbf{w}^\top \hat{\Sigma}_{t, \text{GARCH}}(q) \mathbf{w}},$$

$$\hat{\sigma}_{t, \text{naive}}(q) = \sqrt{\mathbf{w}^\top \hat{\Sigma}_{t, \text{naive}}(q) \mathbf{w}},$$

$$\hat{\sigma}_{t, \text{GARCH}}(q) = \sqrt{L_P^\top \hat{\Sigma}_{t, \text{GARCH}}(q)L_P + \mathbf{w}^\top \hat{\Sigma}_{P, \text{naive}}^2(q) \mathbf{w}},$$

$$\hat{\sigma}_{t, \text{naive}}(q) = \sqrt{L_P^\top \hat{\Sigma}_{t, \text{naive}}(q)L_P + \mathbf{w}^\top \hat{\Sigma}_{P, \text{naive}}^2(q) \mathbf{w}},$$

which will be combined with different window lengths $q$.

4. Data pre-processing

In this section we briefly describe the data in our estimation universe, and list the data processing steps which filter or adjusts the original data.

Our data set consists of daily returns and monthly factor loadings for a total of 28,629 unique companies from February 2010 to April 2021. From these, we have estimated factor returns corresponding to Exabel’s factor model, described in Section 2, with further details in [2].

We have made the following adjustments and filtering of our data:

- Missing industry and country loadings $X_{k,t}^c$ and $X_{k,t}^l$ are filled with 0.
- Missing factor returns $f_{k,t}^s$ are filled with 0.
- Missing market caps are replaced by the square of the factor loading company weights.
- Companies with no registered factor loadings for a certain month are excluded from the study that entire month.
- Companies with no registered company return a certain month are excluded from the study that entire month.
- All sample covariance/correlation matrices are estimated pairwise using daily data, after removing all missing pairs. Non-estimable variances or covariances (due to completely missing data) are replaced by 0.
- We use the naive sample variance instead of the GARCH estimate for single companies/factors when there is either less than 100 observed company(factor) returns, the GARCH model does not converge, or it is not able to produce an estimate for other reasons.
- When estimating the covariance/correlation matrices pairwise, and adjusting non-estimable elements one may end up with covariance/correlation matrices which are not (computationally) positive definite. In such cases, we replace it by the nearest positive definite matrix using the algorithm of [4].

As a consequence of the above filtration and adjustments, companies that have missing observations in the entire estimation window ($q$ months) contribute with zero volatility to the portfolios they are involved in.

5. Portfolios

This section describes and summarizes the portfolios we use in our study. The portfolios are constructed based on the factors described in Section 2 and the market cap $mc$, to reflect different investment strategies. We construct a total of 286 different portfolios for the entire time span. These consist of both unrestricted, region restricted and subregion restricted factor based portfolios. We consider the following regions: America, Asia and Europe, and the following subregions: Latin America and the Caribbean, Northern America, Eastern Asia, South Eastern Asia, Southern Asia, Western Asia, Northern Europe, Southern Europe, and Western Europe. The regions are defined using the ISO 3166 country codes and region/subregion definitions. Each portfolio has a minimum of 40, and a maximum of 300 companies in each test period $T$. The portfolio weights $w_{k,t}$ for company $k$ and $t \in T$ are based on the factor loadings the last month before $T$. In the portfolios, long positions have positive weights, while short positions have negative weights. We will both consider portfolios with purely long positions, and market neutral portfolios where $\sum_{k=1}^P w_{k,t} = 0$. All portfolios are scaled such that $\sum_{k=1}^P |w_{k,t}| = 1$. We only allow companies which have a market cap of at least $200$ million. All long/short portfolios are required to have at least 20 companies with positive weight and 20 companies with negative weight.

When constructing the portfolios below, we will be filtering and weighting companies based on both factor loadings and the market cap $mc_{k,t}$. Although $mc$ has no factor loading, we will use that term also there to simplify the description. Thus, when constructing portfolios weighted by $mc$ below, the term ‘factor loading’ and the notation $X_{k,t}^f$ will refer to $mc_{k,t}$. All portfolio weights are constructed using factor loadings $X_{k,t}^s$ at the last time point $t$ prior to each test period $T$. Our portfolios are constructed as follows:

- **Long, unrestricted portfolios, all factors + $mc$:** For each test period $T$ and each of the 77 factors + $mc$, we filter out all companies without a positive factor loading $X_{k,t}^s$. Out of these, we keep the (maximum) 300 companies with the largest market cap. For the remaining companies, the portfolio weight $w_{k,t}$ is set proportional to $X_{k,t}^s$.
- **Long/short, unrestricted portfolios, all style factors:** For each test period $T$ and each of the $|S| = 11$ style
For the random portfolios, the number of companies per test period \( T \) varies between 40 and 300 as shown in the histogram in Figure 1. Table 1 shows the number of original portfolios of each type.

### Table 1. The number of original portfolios of each type.

| Long/short | Restriction | Type    | #    |
|------------|-------------|---------|------|
| Long       | Unrestricted| country | 27   |
| Long       | Unrestricted| industry| 12   |
| Long       | Unrestricted| style +mc| 12  |
| Long/short | Unrestricted| style | 11   |
| Long       | 3 regions   | style +mc| 36  |
| Long/short | 3 regions   | style | 33   |
| Long       | 9 subregions| style +mc| 79  |
| Long/short | 9 subregions| style | 76   |

The number of companies per test period \( T \) vary between 40 and 300 as shown in the histogram in Figure 1. Table 1 shows the number of portfolios of each type.

For each of the 286 portfolios above, we generate two additional ‘random’ portfolios. Each of these are constructed by sampling 50 companies with replacement from the original portfolio, using \( w_{k,t} \) as sampling weights. This gives 584 additional portfolios, such that we end up with a total of 876 unique portfolios. The purpose of including the additional ‘random’ portfolios is to control the sensitivity to strictly factor based portfolio definitions.

For the random portfolios, the number of companies per test period \( T \) vary between 18 and 50 as shown in the histogram in Figure 2.

### 6. Performance evaluation

In this section we describe how we evaluate the performance of the volatility estimation methods described in Section 3. When comparing different volatility estimation methods, we are mainly interested in their ability to correctly rank the portfolios according to their volatility (i.e. their investment risk). As a consequence, we will evaluate the performance of the estimation methods by computing Kendall’s \( \tau \) [1] between the estimated volatilities and the target volatilities for sets of portfolios within each test period \( T \). Kendall’s \( \tau \) measures the ordinal association between the estimated and target volatilities by comparing their sorting order, and thereby ignoring the actual numeric values. It takes the value 1 if the ordering is exactly the same, and 0 if there is no association between their orderings. We use Kendall’s \( \tau \)-b, which also accounts for ties in the ordering, see e.g. [5]. Since long and long/short portfolios are fundamentally different, we will compare these portfolio sets separately. We will evaluate the rankings of the full set of portfolios (still long and long/short separately), but also various subsets of portfolios. In particular we will perform separate computations for the original and random portfolios to check to what extent the results generalize. The performance across all test periods \( T \) will be summarized by taking the mean of the Kendall’s \( \tau \)’s over all test periods.

### 7. Results

In this section we summarize the results from the performance study. We compare the direct return and factor model ap-
Figure 3. Estimated volatility compared to target volatility across test periods for a portfolio based on long term momentum. The upper and lower panels shows, respectively, the volatilities corresponding to the $q = 1$ and $q = 6$ month estimation windows.

The factor model and the direct return estimation schemes are also quite similar for this portfolio, perhaps with a mild tendency towards less erratic behavior for the direct return approach, at least for $q = 6$. Note that the picture may look completely different for other portfolios.

| Data/Model     | Volatility modelling | $q$ | $\text{avg}(\tau)$ |
|----------------|----------------------|-----|---------------------|
| Direct return  | GARCH                | 12  | 0.576               |
| Direct return  | GARCH                | 6   | 0.570               |
| Direct return  | GARCH                | 3   | 0.558               |
| Direct return  | Naive window         | 6   | 0.554               |
| Direct return  | Naive window         | 3   | 0.539               |
| Direct return  | Naive window         | 12  | 0.539               |
| Direct return  | GARCH                | 1   | 0.519               |
| Direct return  | Naive window         | 1   | 0.492               |
| Factor model   | GARCH                | 6   | 0.470               |
| Factor model   | Naive window         | 6   | 0.470               |
| Factor model   | GARCH                | 12  | 0.464               |
| Factor model   | GARCH                | 3   | 0.463               |
| Factor model   | Naive window         | 12  | 0.458               |
| Factor model   | Naive window         | 3   | 0.456               |
| Factor model   | GARCH                | 1   | 0.434               |
| Factor model   | Naive window         | 1   | 0.422               |

Table 2. Average Kendall tau for portfolio group 'All long'. $q$ denotes the window length in months for the correlation structure (and also the variance structure for the naive window method)

Turning to the ordering ability of the different methods,
we first consider the ability to order all the 498 different long portfolios (both original and random portfolios, which are either unrestricted or restricted to certain regions/subregions). Figure 4 shows Kendall’s $\tau$ across all test periods. Only the best performing window length $q$ (measured by $\text{avg}(\tau)$) is plotted for each of the 4 approaches. We see that except for a few test periods, such as those around the turn of the year 2019/2020, the direct return approach performs better than the factor model. For both estimation schemes, the GARCH volatility model almost always performs better than the naive window approach. Table 2 shows the average of the Kendall’s $\tau$ across all the test periods $T$. It shows that overall, all direct approaches perform better than the factor models.

Similar results are also obtained for the subset of all the 360 long/short portfolios, see Figure 5 and Table 3.

While the factor model is not able to rank these large sets of portfolios as well as the direct return approach, it does perform comparably and sometimes better for other, smaller subsets of portfolios. Figure 6 and Table 4 show corresponding results for 12 style based portfolios restricted to North America. While Table 4 shows the factor model is not as good as the direct approach overall, the performance differences are not big. From Figure 6, we see that at least starting from 2016, the performance of the factor model is generally as good as that of the direct approaches. Further, in the first half of 2018, it is performing significantly better.

In addition to the three portfolio subsets we have looked at above, we have performed analysis for 12 additional portfolio subsets. The tables with average $\tau$ values for these are provided in Appendix A. The results for subsets containing only original portfolios and only random portfolios which otherwise are of the same type, are largely similar. This indicates that the results are not very sensitive to the companies included in the original portfolios and their specific weights.

Table 3. Average Kendall tau for portfolio group ‘All long/short’. $q$ denotes the window length in months for the correlation structure (and also the variance structure for the naive window method)

| Data/Model | Volatility modelling | $q$ | $\text{avg}(\tau)$ |
|------------|----------------------|-----|-------------------|
| Direct return | GARCH | 6 | 0.645 |
| Direct return | GARCH | 12 | 0.645 |
| Direct return | Naive window | 6 | 0.639 |
| Direct return | GARCH | 3 | 0.629 |
| Direct return | Naive window | 12 | 0.624 |
| Direct return | Naive window | 3 | 0.620 |
| Direct return | GARCH | 1 | 0.587 |
| Direct return | Naive window | 1 | 0.573 |
| Factor model | Naive window | 6 | 0.570 |
| Factor model | Naive window | 3 | 0.569 |
| Factor model | GARCH | 6 | 0.569 |
| Factor model | GARCH | 3 | 0.566 |
| Factor model | GARCH | 12 | 0.560 |
| Factor model | Naive window | 12 | 0.556 |
| Factor model | GARCH | 1 | 0.552 |
| Factor model | Naive window | 1 | 0.549 |

8. Conclusion

In this note we have described a set of volatility estimation methods, and studied their ability to rank portfolios according to their future volatility. This included both a direct return approach and using Exabel’s factor model, combined with either a GARCH or naive window based volatility modelling approach. In the estimation universe considered here, we have seen that the direct return approach is the overall best method. However, for some portfolios the factor model has similar or even slightly better performance over some time periods.
Figure 5. Kendall’s $\tau$ between target and estimated volatility across test periods for the all the 360 long/short portfolios. Only the best performing window lengths are plotted for each type of approach.

Figure 6. Kendall’s $\tau$ between target and estimated volatility across test periods for the 12 original North American style based portfolios. Only the best performing window lengths are plotted for each type of approach.
| Data/Model       | Volatility modelling | q | avg(τ) |
|------------------|----------------------|---|--------|
| Direct return GARCH | 6                   | 0.761   |
| Direct return GARCH | 3                   | 0.756   |
| Direct return Naive window | 6              | 0.753   |
| Direct return GARCH | 12                  | 0.751   |
| Direct return Naive window | 3             | 0.747   |
| Factor model GARCH | 6                   | 0.7545  |
| Factor model GARCH | 12                  | 0.730   |
| Factor model Naive window | 6          | 0.745   |
| Factor model GARCH | 1                   | 0.729   |
| Factor model Naive window | 12         | 0.728   |
| Direct return Naive window | 12            | 0.725   |
| Factor model Naive window | 12         | 0.722   |
| Factor model Naive window | 3            | 0.718   |
| Direct return Naive window | 1            | 0.712   |
| Factor model GARCH | 1                   | 0.694   |
| Factor model Naive window | 1            | 0.684   |

Table 4. Average Kendall tau for portfolio group ‘North America original long’. q denotes the window length in months for the correlation structure (and also the variance structure for the naive window method)

It is also clear from the analysis that a GARCH model is almost always preferable over a naive window based volatility modelling approach. The best performing window length q used for the correlation structure (and also the variance for the naive window based approach), seems to vary quite a bit between portfolio subsets. This is expected as the fluctuations in the volatility might be very different in different portfolio types. For most of the portfolio subsets considered here, window lengths of q = 6 or 12 months are the best performing ones.

A. Additional result tables for portfolio subsets

In Tables 5-15, we provide result with average τ values for 12 additional portfolio subsets. The subsets (with number of portfolios in parenthesis) are: "All original long" (166), "All original long/short" (120), "All random long" (332), "All random long/short" (240), "Unrestricted original long" (51), "Unrestricted original long/short" (11), "Unrestricted random long" (102), "Unrestricted random long/short" (22), "North America original long" (12), "North America original long/short" (11), "North America random long" (24), "North America random long/short" (22).

References

[1] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30(1/2):81–93, 1938.

[2] Ø. Grotmol, M. Scheuerer, K. Aas, and M. Jullum. Exabel’s factor model, 2021.

[3] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

[4] N. J. Higham. Computing the nearest correlation matrix—a problem from finance. *IMA journal of Numerical Analysis*, 22(3):329–343, 2002.

[5] A. Agresti. *Analysis of ordinal categorical data*, volume 656. John Wiley & Sons, 2010.
### Table 5. Average Kendall tau for portfolio group 'All original long'.

| Data/Model | Volatility modelling | q | avg(τ) |
|------------|----------------------|---|--------|
| Direct return | GARCH | 12 | 0.587 |
| Direct return | GARCH | 6  | 0.580 |
| Direct return | GARCH | 3  | 0.568 |
| Direct return | Naive window | 6 | 0.561 |
| Direct return | Naive window | 3 | 0.548 |
| Direct return | Naive window | 12 | 0.546 |
| Direct return | GARCH | 1  | 0.532 |
| Direct return | Naive window | 1 | 0.505 |
| Factor model | GARCH | 6  | 0.471 |
| Factor model | Naive window | 6 | 0.471 |
| Factor model | GARCH | 12 | 0.467 |
| Factor model | GARCH | 3  | 0.464 |
| Factor model | Naive window | 12 | 0.461 |
| Factor model | Naive window | 3 | 0.456 |
| Factor model | GARCH | 1  | 0.435 |
| Factor model | Naive window | 1 | 0.423 |

### Table 6. Average Kendall tau for portfolio group 'All original long/short'.

| Data/Model | Volatility modelling | q | avg(τ) |
|------------|----------------------|---|--------|
| Direct return | GARCH | 12 | 0.667 |
| Direct return | GARCH | 6  | 0.665 |
| Direct return | Naive window | 6 | 0.657 |
| Direct return | GARCH | 3  | 0.649 |
| Direct return | Naive window | 12 | 0.642 |
| Direct return | Naive window | 3 | 0.638 |
| Direct return | GARCH | 1  | 0.615 |
| Direct return | Naive window | 1 | 0.601 |
| Factor model | Naive window | 6 | 0.538 |
| Factor model | Naive window | 3 | 0.535 |
| Factor model | GARCH | 6  | 0.533 |
| Factor model | GARCH | 3  | 0.530 |
| Factor model | Naive window | 12 | 0.524 |
| Factor model | GARCH | 12 | 0.522 |
| Factor model | GARCH | 1  | 0.516 |
| Factor model | Naive window | 1 | 0.512 |

### Table 7. Average Kendall tau for portfolio group 'All random long'.

| Data/Model | Volatility modelling | q | avg(τ) |
|------------|----------------------|---|--------|
| Direct return | GARCH | 12 | 0.570 |
| Direct return | GARCH | 6  | 0.564 |
| Direct return | GARCH | 3  | 0.552 |
| Direct return | Naive window | 6 | 0.550 |
| Direct return | Naive window | 3 | 0.534 |
| Direct return | Naive window | 12 | 0.534 |
| Direct return | GARCH | 1  | 0.511 |
| Direct return | Naive window | 1 | 0.485 |
| Factor model | GARCH | 6  | 0.468 |
| Factor model | Naive window | 6 | 0.468 |
| Factor model | GARCH | 12 | 0.461 |
| Factor model | GARCH | 3  | 0.461 |
| Factor model | Naive window | 12 | 0.455 |
| Factor model | Naive window | 3 | 0.453 |
| Factor model | GARCH | 1  | 0.432 |
| Factor model | Naive window | 1 | 0.419 |

### Table 8. Average Kendall tau for portfolio group 'All random long/short'.

| Data/Model | Volatility modelling | q | avg(τ) |
|------------|----------------------|---|--------|
| Direct return | Naive window | 6 | 0.591 |
| Direct return | GARCH | 6  | 0.590 |
| Direct return | GARCH | 12 | 0.590 |
| Direct return | Naive window | 12 | 0.575 |
| Direct return | GARCH | 3  | 0.573 |
| Direct return | Naive window | 3 | 0.569 |
| Direct return | GARCH | 1  | 0.524 |
| Factor model | Naive window | 6 | 0.522 |
| Factor model | GARCH | 6  | 0.521 |
| Factor model | Naive window | 3 | 0.518 |
| Factor model | GARCH | 3  | 0.516 |
| Factor model | GARCH | 12 | 0.511 |
| Direct return | Naive window | 1 | 0.509 |
| Factor model | Naive window | 12 | 0.505 |
| Factor model | GARCH | 1  | 0.498 |
| Factor model | Naive window | 1 | 0.494 |
| Data/Model | Volatility modelling | q  | avg(τ) |
|-----------|----------------------|----|--------|
| Direct return | GARCH | 12  | 0.575  |
| Direct return | GARCH | 6   | 0.569  |
| Direct return | GARCH | 3   | 0.558  |
| Direct return | Naive window | 6   | 0.546  |
| Direct return | Naive window | 3   | 0.538  |
| Direct return | Naive window | 12  | 0.530  |
| Direct return | GARCH | 1   | 0.512  |
| Direct return | Naive window | 1   | 0.484  |
| Factor model | Naive window | 6   | 0.447  |
| Factor model | Naive window | 3   | 0.440  |
| Factor model | GARCH | 6   | 0.439  |
| Factor model | GARCH | 3   | 0.437  |
| Factor model | GARCH | 12  | 0.433  |
| Factor model | Naive window | 12  | 0.431  |
| Factor model | GARCH | 1   | 0.414  |
| Factor model | Naive window | 1   | 0.404  |

Table 9. Average Kendall tau for portfolio group 'Unrestricted original long'.
| Data/Model      | Volatility modelling | q | avg(τ) |
|----------------|----------------------|---|--------|
| Direct return  | GARCH                | 6 | 0.761  |
| Direct return  | GARCH                | 3 | 0.756  |
| Direct return  | Naive window         | 6 | 0.753  |
| Direct return  | GARCH                | 12| 0.751  |
| Direct return  | Naive window         | 3 | 0.747  |
| Factor model   | GARCH                | 6 | 0.745  |
| Factor model   | Naive window         | 6 | 0.745  |
| Direct return  | GARCH                | 1 | 0.730  |
| Factor model   | GARCH                | 12| 0.729  |
| Factor model   | GARCH                | 3 | 0.728  |
| Direct return  | Naive window         | 12| 0.725  |
| Factor model   | Naive window         | 12| 0.722  |
| Factor model   | Naive window         | 3 | 0.718  |
| Direct return  | Naive window         | 1 | 0.712  |
| Factor model   | GARCH                | 1 | 0.694  |
| Factor model   | Naive window         | 1 | 0.684  |

Table 13. Average Kendall tau for portfolio group 'North America original long'.

| Data/Model      | Volatility modelling | q | avg(τ) |
|----------------|----------------------|---|--------|
| Direct return  | Naive window         | 6 | 0.732  |
| Direct return  | GARCH                | 6 | 0.731  |
| Direct return  | GARCH                | 12| 0.730  |
| Direct return  | Naive window         | 3 | 0.728  |
| Direct return  | GARCH                | 3 | 0.726  |
| Direct return  | Naive window         | 12| 0.722  |
| Direct return  | Naive window         | 1 | 0.697  |
| Direct return  | GARCH                | 1 | 0.696  |
| Factor model   | Naive window         | 6 | 0.637  |
| Factor model   | Naive window         | 12| 0.633  |
| Factor model   | GARCH                | 6 | 0.623  |
| Factor model   | Naive window         | 3 | 0.622  |
| Factor model   | GARCH                | 3 | 0.609  |
| Factor model   | GARCH                | 12| 0.606  |
| Factor model   | GARCH                | 1 | 0.598  |
| Factor model   | Naive window         | 1 | 0.590  |

Table 14. Average Kendall tau for portfolio group 'North America original long/short'.

| Data/Model      | Volatility modelling | q | avg(τ) |
|----------------|----------------------|---|--------|
| Direct return  | Naive window         | 6 | 0.707  |
| Direct return  | Naive window         | 6 | 0.706  |
| Direct return  | GARCH                | 3 | 0.705  |
| Direct return  | GARCH                | 12| 0.704  |
| Factor model   | GARCH                | 6 | 0.702  |
| Factor model   | Naive window         | 6 | 0.698  |
| Direct return  | GARCH                | 12| 0.696  |
| Direct return  | Naive window         | 12| 0.692  |
| Factor model   | GARCH                | 12| 0.687  |
| Factor model   | Naive window         | 12| 0.686  |
| Factor model   | GARCH                | 3 | 0.680  |
| Factor model   | Naive window         | 3 | 0.680  |
| Direct return  | GARCH                | 1 | 0.667  |
| Factor model   | Naive window         | 1 | 0.656  |
| Factor model   | GARCH                | 1 | 0.655  |
| Direct return  | Naive window         | 1 | 0.644  |

Table 15. Average Kendall tau for portfolio group 'North America random long'.

| Data/Model      | Volatility modelling | q | avg(τ) |
|----------------|----------------------|---|--------|
| Direct return  | Naive window         | 6 | 0.653  |
| Direct return  | GARCH                | 6 | 0.645  |
| Direct return  | GARCH                | 12| 0.638  |
| Direct return  | Naive window         | 12| 0.638  |
| Direct return  | GARCH                | 3 | 0.628  |
| Direct return  | Naive window         | 3 | 0.626  |
| Direct return  | GARCH                | 1 | 0.586  |
| Direct return  | Naive window         | 1 | 0.570  |
| Factor model   | GARCH                | 6 | 0.564  |
| Factor model   | Naive window         | 6 | 0.564  |
| Factor model   | GARCH                | 12| 0.552  |
| Factor model   | Naive window         | 12| 0.551  |
| Factor model   | GARCH                | 3 | 0.546  |
| Factor model   | GARCH                | 3 | 0.545  |
| Factor model   | GARCH                | 1 | 0.515  |
| Factor model   | Naive window         | 1 | 0.510  |

Table 16. Average Kendall tau for portfolio group 'North America random long/short'.