

Large Dynamic Covariance Matrices: Enhancements Based on Intraday Data

Gianluca De Nard
Department of Banking and Finance
University of Zurich
CH-8032 Zurich, Switzerland
gianluca.denard@bf.uzh.ch

Robert F. Engle
Department of Finance
New York University
New York, NY 10012, USA
rengle@stern.nyu.edu

Olivier Ledoit*
Department of Economics
University of Zurich
CH-8032 Zurich, Switzerland
olivier.ledoit@econ.uzh.ch

Michael Wolf†
Department of Economics
University of Zurich
CH-8032 Zurich, Switzerland
michael.wolf@econ.uzh.ch

July 2020

Abstract

Modeling and forecasting dynamic (or time-varying) covariance matrices has many important applications in finance, such as Markowitz portfolio selection. A popular tool to this end are multivariate GARCH models. Historically, such models did not perform well in large dimensions due to the so-called curse of dimensionality. The recent DCC-NL model of [Engle et al. \(2019\)](#) is able to overcome this curse via nonlinear shrinkage estimation of the unconditional correlation matrix. In this paper, we show how performance can be increased further by using open/high/low/close (OHLC) price data instead of simply using daily returns. A key innovation, for the improved modeling of not only dynamic variances but also of dynamic covariances, is the concept of a *regularized return*, obtained from a volatility proxy in conjunction with a smoothed sign (function) of the observed return.

KEY WORDS: Dynamic conditional correlations; intraday data; Markowitz portfolio selection; multivariate GARCH; nonlinear shrinkage.

JEL CLASSIFICATION NOS: C13, C58, G11.

*Second affiliation: AlphaCrest Capital Management, New York, NY 10036, USA

†Corresponding author. Postal address: Zürichbergstrasse 14, CH-8032 Zurich, Switzerland.

1 Introduction

Modeling and forecasting dynamic (or time-varying) covariance matrices of a vector of asset returns has many important applications in finance, such as Markowitz portfolio selection. A popular tool to this end are multivariate GARCH models. Historically, such models did not perform well in large dimensions due to the so-called curse of dimensionality. The recent DCC-NL model of [Engle et al. \(2019\)](#) is able to overcome this curse via *nonlinear shrinkage* (NL) estimation of the unconditional correlation matrix.

Just as the original *dynamic conditional correlations* (DCC) model of [Engle \(2002\)](#), also the DCC-NL model is (primarily) based on daily returns, for the modeling of both univariate dynamic variances and the dynamic (pseudo-)correlation matrix; those quantities are then combined for the modeling of the dynamic covariance matrix. It has been known for quite some time that the use of intraday return data can lead to improved modeling of univariate dynamic variances. Even if one ‘only’ uses the four pieces of information given by open/high/low/close (OHLC) prices, the improvements can be noticeable; for example, see [Garman and Klass \(1980\)](#) and [Molnár \(2016\)](#). Further improvements can be obtained using 5- or 10-minute returns; for example, see [Hansen et al. \(2012\)](#). The intuition is that using an improved volatility proxy, such as the high-low range (even though very simple) or realized volatility (at the other end of the sophistication spectrum) as an innovation in a GARCH(1,1) model works better than using the simple-minded volatility proxy of the squared daily return, which is the original way of doing it.

Therefore, it is natural to use such an approach in the first step (out of three) of a DCC(-NL) model: the modeling of univariate dynamic variances; and, indeed, doing so yields better forecasts of dynamic covariance matrices, as we find empirically. But why stop there? The second step of a DCC(-NL) model consists of modeling the dynamic (pseudo-)correlation matrix where the innovation is the outer product of the vector of daily (devolitized) returns. So a further idea is to find a better innovation, in this second step as well, that is based on volatility proxies instead of returns. The counterpart of a volatility proxy itself is the squared return. Hence, it is natural to take the square root of the volatility proxy, together with sign of the return, as the counterpart of the return itself. This already works well, but a certain ‘smoothed’ sign of the return, together with the square root of the volatility proxy works even better, which is motivated by both theoretical reasoning and empirical findings. We call this construct *regularized return* and it constitutes a key innovation of our DCC(-NL) models based on intraday data.

Importantly, our models remain computationally feasible also for large dimensions of $N \geq 1000$ assets. In contrast, the HEAVY-DCC model, which is also based on intraday data, can be applied only to small investment universes; for example, see the empirical analyses with $N \leq 10$ of [Xu \(2019\)](#) and [Noureldin et al. \(2012\)](#).

The remainder of the paper is organized as follows. Section 2 briefly reviews DCC(-NL) models. Section 3 details our new models based on intraday data. Section 4 gives a brief

description of existing variance estimators we deem the most useful for our purpose. Section 5 describes the empirical methodology and presents the results of an out-of-sample backtest exercise based on real-life stock return data. Section 6 concludes. An appendix contains all figures and tables.

2 Large Dynamic Covariance Matrices

2.1 Notation

In what follows, the subscript i indexes the assets and covers the range of integers from 1 to N , where N denotes the dimension of the investment universe; the subscript t indexes the dates and covers the range of integers from 1 to T , where T denotes the sample size. The notation $\text{Cor}(\cdot)$ represents the correlation matrix of a random vector, the notation $\text{Cov}(\cdot)$ represents the covariance matrix of a random vector, and the notation $\text{Diag}(\cdot)$ represents the function that sets to zero all the off-diagonal elements of a matrix. Furthermore, we use the following notations:

- $o_{i,t}$: observed opening price (“open”) for asset i at date t , stacked into $o_t := (o_{1,t}, \dots, o_{N,t})'$
- $h_{i,t}$: observed highest price (“high”) transacted for asset i at date t , stacked into $h_t := (h_{1,t}, \dots, h_{N,t})'$
- $l_{i,t}$: observed lowest price (“low”) transacted for asset i at date t , stacked into $l_t := (l_{1,t}, \dots, l_{N,t})'$
- $c_{i,t}$: observed closing price (“close”) for asset i at date t , stacked into $c_t := (c_{1,t}, \dots, c_{N,t})'$
- $r_{i,t}$: observed return for asset i at date t , stacked into $r_t := (r_{1,t}, \dots, r_{N,t})'$
- $\tilde{r}_{i,t}$: regularized return for asset i at date t , stacked into $\tilde{r}_t := (\tilde{r}_{1,t}, \dots, \tilde{r}_{N,t})'$
- $x_{i,t}$: underlying time-series for covariance matrix estimation; thus $x_{i,t} \in \{r_{i,t}, \tilde{r}_{i,t}\}$
- $d_{i,t}^2 := \text{Var}(x_{i,t} | \mathcal{F}_{t-1})$: conditional variance of the i th asset at t
- $s_{i,t} := x_{i,t}/d_{i,t}$: devolatilized series, stacked into $s_t := (s_{1,t}, \dots, s_{N,t})'$
- D_t : the N -dimensional diagonal matrix whose i th diagonal element is $d_{i,t}$
- $R_t := \text{Cor}(x_t | \mathcal{F}_{t-1}) = \text{Cov}(s_t | \mathcal{F}_{t-1})$: conditional correlation matrix at date t
- $\Sigma_t := \text{Cov}(x_t | \mathcal{F}_{t-1})$: conditional covariance matrix at date t ; thus $\text{Diag}(\Sigma_t) = D_t^2$
- $C := \mathbb{E}(R_t) = \text{Cor}(x_t) = \text{Cov}(s_t)$: unconditional correlation matrix

Here, the symbol $:=$ is a definition sign where the left-hand side is defined to be equal to the right-hand side, whereas the symbol $=:$ (to be used below) is a definition sign where the right-hand side is defined to be equal to the left-hand side.

Remark 2.1 (Terminology). Note that in this paper, the terms “dynamic” and “conditional” are interchangeable. As an example a dynamic covariance matrix means the same thing as a conditional covariance matrix, such as the covariance matrix Σ_t defined above; analogously for a

correlation matrix and, necessarily then, also for any entries of such matrices, such as a variance, a covariance, or a correlation. ■

2.2 Averaged Forecasting of Dynamic Covariance Matrices

In our empirical analysis, as is common in the literature, we use (intra-)daily data to forecast dynamic covariance matrices but then hold the portfolio for an entire ‘month’ (that is, for a period of 21 subsequent trading days) before updating it again. This creates a certain ‘mismatch’ for dynamic models that assume that the (conditional) covariance matrix changes at the forecast frequency, that is, at the daily level: Why use a covariance matrix forecasted only for the next day to construct a portfolio that will then be held for an entire month?

To address this mismatch, we use an ‘averaged-forecasting’ approach for all dynamic models; this approach was first suggested by [De Nard et al. \(2020\)](#). At portfolio construction date k , forecast the covariance matrix for all days of the upcoming month, that is, for $t = k, k + 1, \dots, k + 20$; then average those 21 forecasts and use this ‘averaged forecasts’ to construct the portfolio at date k .

To model conditional variances, we use a GARCH(1,1) process:

$$d_{i,t}^2 = \omega_i + \delta_{1,i}x_{i,t-1}^2 + \delta_{2,i}d_{i,t-1}^2, \quad (2.1)$$

where $(\omega_i, \delta_{1,i}, \delta_{2,i})$ are the variable-specific GARCH(1,1) parameters. We assume that the evolution of the conditional correlation matrix over time is governed as in the DCC-NL model of [Engle et al. \(2019\)](#):

$$Q_t = (1 - \delta_1 - \delta_2)C + \delta_1 s_{t-1} s'_{t-1} + \delta_2 Q_{t-1}, \quad (2.2)$$

where (δ_1, δ_2) are the DCC-NL parameters analogous to $(\delta_{1,i}, \delta_{2,i})$. The matrix Q_t can be interpreted as a conditional pseudo-correlation matrix, or a conditional covariance matrix of devolatilized residuals. It cannot be used directly because its diagonal elements, although close to one, are not exactly equal to one. From this representation, we obtain the conditional correlation matrix and the conditional covariance matrix as

$$R_t := \text{Diag}(Q_t)^{-1/2} Q_t \text{Diag}(Q_t)^{-1/2} \quad (2.3)$$

$$\Sigma_t := D_t R_t D_t, \quad (2.4)$$

and the data-generating process is driven by the multivariate normal law

$$x_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, \Sigma_t). \quad (2.5)$$

Hence, to determine the average of the L forecasts of the conditional covariance matrices $\Sigma_{k+l} = D_{k+l} R_{k+l} D_{k+l}$, for $l = 0, 1, \dots, L - 1$, we suggest a three-step approach where D_{k+l} and R_{k+l} can be forecasted separately.

2.2.1 Step One: Forecasting Conditional Univariate Volatilities

According to [Baillie and Bollerslev \(1992\)](#), the multi-step ahead forecasts of the $i = 1, \dots, N$ GARCH(1,1) volatilities can be written as

$$\mathbb{E}[d_{i,k+l}^2 | \mathcal{F}_{k-1}] = \sum_{j=0}^{l-1} \omega_i (\delta_{1,i} + \delta_{2,i})^j + (\delta_{1,i} + \delta_{2,i})^l \mathbb{E}[d_{i,k}^2 | \mathcal{F}_{k-1}] , \quad (2.6)$$

where $\mathbb{E}[d_{i,k}^2 | \mathcal{F}_{k-1}] = \omega_i + \delta_{1,i} x_{i,k-1}^2 + \delta_{2,i} d_{i,k-1}^2$. Therefore, we compute the forecasts of the N -dimensional diagonal matrix D_{k+l} as

$$\mathbb{E}[D_{k+l} | \mathcal{F}_{k-1}] = \text{Diag} \left(\sqrt{\mathbb{E}[d_{1,k+l}^2 | \mathcal{F}_{k-1}]}, \dots, \sqrt{\mathbb{E}[d_{N,k+l}^2 | \mathcal{F}_{k-1}]} \right) . \quad (2.7)$$

2.2.2 Step Two: Forecasting Conditional Correlation Matrices

For the multivariate case we consider the approach of [Engle and Sheppard \(2001\)](#) where the multi-step ahead forecasts of the conditional correlation matrices are computed as

$$\mathbb{E}[R_{k+l} | \mathcal{F}_{k-1}] = \sum_{j=0}^{l-1} (1 - \delta_1 - \delta_2) C (\delta_1 + \delta_2)^j + (\delta_1 + \delta_2)^l \mathbb{E}[R_k | \mathcal{F}_{k-1}] , \quad (2.8)$$

using the approximation $\mathbb{E}[R_k | \mathcal{F}_{k-1}] \approx \mathbb{E}[Q_k | \mathcal{F}_{k-1}]$. In practice, the diagonal elements of the matrix C tend to deviate from one slightly, in spite of the fact that devolatilized returns are used as inputs. Therefore, every column and every row has to be divided by the square root of the corresponding diagonal entry, so as to produce a proper correlation matrix.

2.2.3 Step Three: Averaging Forecasted Conditional Covariance Matrices

By using the notation $\hat{\Sigma}_{k+l} := \mathbb{E}[\Sigma_{k+l} | \mathcal{F}_{k-1}]$, $\hat{R}_{k+l} := \mathbb{E}[R_{k+l} | \mathcal{F}_{k-1}]$ and $\hat{D}_{k+l} := \mathbb{E}[D_{k+l} | \mathcal{F}_{k-1}]$ we finally calculate $\hat{\Sigma}_{k+l} := \hat{D}_{k+l} \hat{R}_{k+l} \hat{D}_{k+l}$, for $l = 0, 1, \dots, L-1$. Therefore, to get the estimated covariance matrix on portfolio construction day k we average over the L forecasts:

$$\hat{\Sigma}_k := \frac{1}{L} \sum_{l=0}^{L-1} \hat{\Sigma}_{k+l} . \quad (2.9)$$

2.3 Estimation of Parameters

Note that in practice, the GARCH parameters in step one and the DCC(-NL) parameters in step two need to be estimated first. In doing so, we mainly follow the suggestions of [Engle et al. \(2019, Section 3\)](#).

In step one, the GARCH parameters of Equation (2.1) are estimated using (pseudo) maximum likelihood assuming normality. This results in estimators $(\hat{w}_i, \hat{\delta}_{1,i}, \hat{\delta}_{2,i})$ that are used for devolatilizing returns and are also used for forecasting conditional variances via Equation (2.6).

In step two, the correlation-targeting matrix C of Equation (2.2) is estimated in one of two ways. For DCC, we use the sample covariance matrix of the devolatilized returns $\{s_t\}$; for

DCC-NL we use nonlinear shrinkage applied to the $\{s_t\}$, with post-processing to enforce a proper correlation matrix; to speed up the computations, we use the analytical nonlinear shrinkage method of [Ledoit and Wolf \(2020\)](#).¹ Having an estimator \hat{C} , in one of these two ways, we then estimate the DCC parameters (δ_1, δ_2) of Equation (2.2) using the (pseudo) composite likelihood method of [Pakel et al. \(2020\)](#) assuming normality.² In this way, $(\hat{w}_i, \hat{\delta}_{1,i}, \hat{\delta}_{2,i}, \hat{\delta}_1, \hat{\delta}_2)$ are used for forecasting conditional correlation matrices via Equation (2.8).

Combining forecasts of conditional variances with forecasts of conditional correlation matrices yields forecasts of conditional covariance matrices in the usual fashion.

3 Models Based on OHLC Data

One might ask why not use monthly data instead of daily data for the estimation of the various models given that the investment horizon is one month? The justification is that at the monthly frequency we do not have enough data to estimate a multivariate GARCH model. Another justification is that using daily data for the estimation tends to lead to better results even if the investment period is one month; for examples compare Tables 1 and 10 of [Ledoit and Wolf \(2017\)](#).³ Therefore, we propose to use even higher frequencies (that is, intraday data) to obtain (a) a less noisy volatility proxy and (b) a regularized return time series that can be used to estimate DCC(-NL) models.

In GARCH-type models, squared (possibly demeaned) returns are commonly used as innovations in models for conditional variances. Rewriting the GARCH(1,1) model in terms of observed variables (returns) shows that the GARCH(1,1) model in fact expresses volatility (to be interpreted as current conditional variance in this context) as a weighted moving average of past squared returns. Or, looking at it from a slightly different angle: If the squared return is taken to be as a proxy for the volatility of the corresponding day, the GARCH(1,1) model in fact expresses volatility as a weighted moving average of past volatilities.

This intuition has several interesting implications. Most importantly, replacing the squared returns by less noisy volatility proxies should lead to improved GARCH models, in terms of both in-sample fit and out-of-sample forecasting performance.

¹In contrast, [Engle et al. \(2019, Section 3\)](#) used the numerical method of [Ledoit and Wolf \(2015\)](#).

²As [Engle et al. \(2019, Section 3\)](#) do, we using neighboring pairs of assets to build up a (pseudo) composite likelihood.

³Note that the two out-of-sample investment periods are not the same. Nevertheless, for dimension $N = 1000$, using daily data for the estimation reduces the out-of-sample standard deviation of the estimated global minimum variance portfolio by 49% when upgrading from the equal-weighted portfolio ($1/N$) to nonlinear shrinkage; on the other hand, the corresponding improvement is only 36% when using monthly data for the estimation instead.

3.1 Volatility Proxy

Following the basic premise of Molnár (2016), we aim to replace the squared returns $r_{i,t-1}^2$ of Equation (2.1) with a less noisy volatility proxy $\hat{v}_{i,t-1}$:

$$\hat{d}_{i,t}^2 = \omega_i + \delta_{1,i}\hat{v}_{i,t-1} + \delta_{2,i}d_{i,t-1}^2 . \quad (3.1)$$

The goal is to use intraday information to hopefully generate a less noisy volatility proxy (compared to simply using the squared daily return), to be used for devolating returns (only). Note that the returns that get thus devolating are still the observed daily returns. Hence, we only modify the first step of DCC(-NL) models, modeling conditional variances. We call the resulting models ID-DCC(-NL), where ID stands for volatility proxy based on IntraDay data.

Even though the first step is vitally important to obtain a reliable conditional covariance matrix estimator via DCC-type models, it is not the end of the story. Especially in large dimensions, there are many more conditional covariances/correlations $((N^2 - N)/2)$ to estimate than conditional variances (N) . Consequently, we also consider using intraday data not only to improve upon the diagonal of a covariance matrix estimator but also in its off-diagonal. It stands to reason that intraday data reveal additional information that enables us to generate also a less noisy correlation proxy, and not only a less noisy volatility proxy. To this end, we introduce the new concept of a *regularized return*.

3.2 Regularized Returns

In this section, we propose to go beyond ID-DCC(-NL) models and to use the intraday information also for the estimation of (conditional) covariances in the second step of DCC(-NL) models. As mentioned before, in models for conditional variances, it is beneficial to use a ‘better’ innovation in GARCH-type models than the squared daily return: this is just the point of using an improved volatility proxy. So why not generalize this approach to the estimation of conditional covariances respectively correlations? We are working within the framework of DCC(-NL) models where the innovation in the estimation of the conditional (pseudo-)correlation matrix is the outer product of a certain vector, namely the vector of devolating daily returns. The natural idea is then to regularize returns quantities derived from improved volatility proxies (before devolating). If the improved volatility proxy is used instead of the *squared* return, then what should be used instead of the return itself? As a return can be both positive and negative, the first thought would be to use the *signed* root of the volatility proxy:

$$\tilde{r}_{i,t-1}^{\text{naïve}} := \text{sign}(r_{i,t-1})\sqrt{\hat{v}_{i,t-1}} . \quad (3.2)$$

However, it turns out that the following refinement works even better in practice:

$$\tilde{r}_{i,t-1} := \text{stanh}(r_{i,t-1}, \kappa)\sqrt{\hat{v}_{i,t-1}} , \quad (3.3)$$

where $\text{stanh}(\cdot, \cdot)$ denotes the ‘scaled’ hyperbolic tangent function:

$$\text{stanh}(r, \kappa) := \frac{e^{\kappa r} - 1}{e^{\kappa r} + 1}, \quad (3.4)$$

which is graphically displayed in Figure 1.

The scaling factor κ denotes the steepness of the hyperbolic tangent function. The larger κ , the faster the function converges to ± 1 as r moves from 0 to $\pm\infty$. Note that for $\kappa = 2$, one obtains the ‘standard’ hyperbolic tangent function

$$\text{stanh}(r, 2) = \frac{e^{2r} - 1}{e^{2r} + 1} = \frac{e^r - e^{-r}}{e^r + e^{-r}} = \frac{\sinh(r)}{\cosh(r)} =: \tanh(r), \quad (3.5)$$

and for $\kappa \rightarrow \infty$, one obtains the ordinary sign function in the limit:

$$\text{stanh}(r, \infty) := \lim_{\kappa \rightarrow \infty} \text{stanh}(r, \kappa) = \begin{cases} 1, & \text{for } r > 0 \\ 0, & \text{for } r = 0 \\ -1, & \text{for } r < 0 \end{cases} =: \text{sign}(r). \quad (3.6)$$

The reason why we use the more general scaled hyperbolic tangent function in our final definition (3.3) instead of the (ordinary) sign function is that the sign function is intuitively not ‘quite right’ when the observed return $r_{i,t-1}$ is close to zero. For example, assume the observed return is only 1 bp, then the sign function would set the regularized return equal to the root of the volatility proxy which is undesirable, as, by construction, the volatility proxy, and thus also its root, is bounded away from zero in practice; for an illustration, see Figure 2. Therefore, for observed returns close to zero, the difference between the observed return and the root of the volatility proxy can be large in some scenarios. Consequently, we propose to shrink the root of the volatility proxy towards zero for such cases, where the shrinkage intensity is governed by the steepness of the scaled hyperbolic tangent function: the smaller κ , the more pronounced is the truncation; again, see Figure 1.

In practice, the question is: what is a suitable scaling factor κ ? We argue, and empirically find, that the value of κ should be relatively high, as the observed returns and the signed roots of the volatility proxies are highly correlated (a typical number being 90%) and one should only apply shrinkage if the observed return is very close to zero. As a consequence, the sign function is expected to be suboptimal but actually not far from optimal. At the purely theoretical level, we would want ‘noticeable’ shrinkage to occur only when the absolute value of the observed return is less than the typical magnitude of the bid-ask bounce, which we can take to be approximately 5 bps. As shown in Figure 1, setting the parameter κ equal to 100 achieves this objective. Furthermore, empirical analyses in Section 5 indicate that this parameter choice also works ‘best’ in practice.

Remark 3.1 (Returns in percent vs. raw returns). The discussion above applies to returns in percent. If raw returns are used instead, the value of κ needs to be multiplied by 100 to achieve the same amount of shrinkage; see Equation (3.4) and Figure 1. For example, the choice $\kappa = 100$ for returns in percent corresponds to the choice $\kappa = 10,000$ for raw returns. ■

A representative example of the proposed regularized returns (time series) is given in Figure 3. One can see that due to the scaled hyperbolic tangent function and the mean-reverting property of conditional variances (in a stationary world), the regularized returns fluctuate around their representative positive and negative return level. Hence, regularized returns place more weight on the sign of the observed return relative to the magnitude of the observed return when they are used (in place of observed returns) in the estimation of the DCC dynamics in the second step. Arguably, this is a desirable feature, as the magnitude of daily stock returns is generally regarded as unpredictable, and overly ‘noisy’, whereas their sign is regarded as predictable to some extent; for example, see [Welch and Goyal \(2008\)](#); [Henriksson and Merton \(1981\)](#), [Pesaran and Timmermann \(1995\)](#), and [Christoffersen and Diebold \(2006\)](#).

Note that for internal consistency, when we use regularized returns in the second step to model conditional (pseudo-)correlation matrices, we also use (squared) regularized returns in the first step to model conditional variances:

$$d_{i,t}^2 = \omega_i + \delta_{1,i} \tilde{r}_{i,t-1}^2 + \delta_{2,i} d_{i,t-1}^2 . \quad (3.7)$$

We call the resulting models IDR-DCC(-NL), where R stands for “regularized returns”. Appendix B provides a detailed description of these models.

Remark 3.2 (Intraday data vs. high-frequency data). Based on what we have promoted so far, it might be tempting to go even further and use high-frequency data such as 5-minute returns, or even tick-by-tick data instead of ‘only’ intraday data in the form of OHLC prices. However, this would give rise to a number of difficulties. First, high-frequency data is not easily available to everybody, and when it is, it tends to be expensive, especially if one wants a large universe and a long time series. Second, even if such data is there, it requires expert cleaning to be put in usable form, which is tedious and time-consuming; for example, see [Barndorff-Nielsen et al. \(2009\)](#). Third, using high-frequency data stretches computing resources because the data takes a lot of space, and running multiple simulations or backtest exercises would be very slow. For all these reasons, we stick to intraday data in the form OHLC prices in this paper. Nevertheless, if someone wants to go down the high-frequency route, our methodology, including regularized returns, can be adapted easily. ■

4 Volatility Proxies/Estimators of Conditional Variance

In this section, we review the existing volatility proxies (or conditional-variance estimators) that we deem the most useful for our purpose. Note that returns are not demeaned, as is common practice in the literature when working with daily returns.⁴

⁴In essence, the expected (unconditional) return is generally indistinguishable from zero at the daily level.

4.1 Close/Close

From these data, we deduce a synthetic previous-day closing price:

$$\tilde{c}_{i,t-2} := \frac{c_{i,t-1}}{1 + r_{i,t-1}}. \quad (4.1)$$

Most of the time, $\tilde{c}_{i,t-2}$ is equal to $c_{i,t-2}$, except when there is an overnight dividend, stock split, or other corporate action, in which case $\tilde{c}_{i,t-2}$ is suitably rescaled to be economically compatible for use in a formula alongside $c_{i,t-1}$ and any other price recorded on day t .

The first, and most obvious, estimator of the conditional variance on day $t - 1$ is

$$r_{i,t-1}^2 = \left(\frac{c_{i,t-1}}{\tilde{c}_{i,t-2}} - 1 \right)^2. \quad (4.2)$$

$r_{i,t-1}^2$ is the usual ‘ingredient’ in the standard ARCH/GARCH models of [Engle \(1982\)](#) and [Bollerslev \(1986\)](#).

As will soon become apparent, most of the relevant literature works with continuously compounded rather than simple returns; therefore, for compatibility reasons, we define

$$\hat{v}_{i,t-1}^{\text{CC}} := [\log(1 + r_{i,t-1})]^2, \quad (4.3)$$

where $\log(\cdot)$ denotes the natural logarithm.

Taking logarithms at the daily frequency makes very little difference in numerical terms. However, we aim to reduce the noise of the proxy $\hat{v}_{i,t}^{\text{CC}}$ by incorporating new intraday data, such as open, high, and low, while preserving the dynamic features of the ARCH/GARCH framework. Such improvements have traditionally been couched in terms of continuously compounded returns due to mathematical grounding in the random-walk model favored by the [Black and Scholes \(1973\)](#) option-pricing formula.

4.2 Introducing the Open

One of the first contributions of [Garman and Klass \(1980\)](#) is to realize that decomposing the close-to-close log-return $\log(r_{i,t-1})$ into

$$\log(1 + r_{i,t-1}) = \log\left(\frac{o_{i,t-1}}{\tilde{c}_{i,t-2}}\right) + \log\left(\frac{c_{i,t-1}}{o_{i,t-1}}\right) \quad (4.4)$$

opens the door to a family of improved estimators. An issue is scaling: both overnight and open-market variances are on a different scale than daily, so they need to be adjusted appropriately. To this end, [Garman and Klass \(1980, Section III\)](#) introduce the factor $f \in (0, 1)$, which represents the proportion of variance realized when the market is closed. From this analysis, they derive an improved estimator:

$$\hat{v}_{i,t-1}^{\text{OC}} := \frac{1}{2f} \left[\log\left(\frac{o_{i,t-1}}{\tilde{c}_{i,t-2}}\right) \right]^2 + \frac{1}{2(1-f)} \left[\log\left(\frac{c_{i,t-1}}{o_{i,t-1}}\right) \right]^2. \quad (4.5)$$

[Yang and Zhang \(2000, p. 485\)](#), based on an empirical study of US equity data at the daily frequency, recommend setting $f = 0.25$.

4.3 High-Low Range

Parkinson (1980) proposes an estimator for the conditional variance during market hours on day $t - 1$ based on the range, determined by high and low:

$$\hat{v}_{i,t-1}^{\text{HL}} := \frac{1}{4 \log 2} \left[\log \left(\frac{h_{i,t-1}}{l_{i,t-1}} \right) \right]^2, \quad (4.6)$$

where the normalizing coefficient $4 \log 2$ is derived from random-walk mathematics. In reality, however, overnight jumps matter. To this end, Garman and Klass (1980, Section IV) propose to amend the range-based estimator as follows:

$$\hat{v}_{i,t-1}^{\text{OHLC}} := \frac{a_3}{f} \left[\log \left(\frac{o_{i,t-1}}{\tilde{c}_{i,t-2}} \right) \right]^2 + \frac{1 - a_3}{1 - f} \frac{1}{4 \log 2} \left[\log \left(\frac{h_{i,t-1}}{l_{i,t-1}} \right) \right]^2, \quad (4.7)$$

and claim that for the optimal choice of parameter $a_3 = 0.17$, $\hat{v}_{i,t-1}^{\text{OHLC}}$ is around 6.2 times more efficient than the naïve close-to-close estimator $\hat{v}_{i,t-1}^{\text{CC}}$.

4.4 Additional Refinements

Most of the analytical work in Garman and Klass (1980) is then devoted to (potentially) further improving upon the estimator $\hat{v}_{i,t-1}^{\text{OHLC}}$. Whereas some mild efficiency gains are attained on paper, they are heavily dependent upon the assumption that log-prices follow a Brownian motion with constant volatility, which is unlikely to hold in practice. Their final recommendation, which they denote by $\hat{\sigma}_5^2$ and which is often referred to as “the” Garman-Klass estimator⁵, overloads on the Parkinson (1980) estimator and compensates for it by *negatively* loading on the more naïve open-to-close estimator. Garman and Klass (1980, p. 71) motivate this somewhat surprising proposal as follows:

High and low prices during the trading interval require continuous monitoring to establish their values. The opening and closing prices, on the other hand, are merely “snapshots” of the process. Intuition would then tell us that high/low prices contain more information regarding volatility than do open/close prices.

However, analyzing real-life data, which may not conform to Garman and Klass’s theoretical assumptions, Molnár (2016, p. 4979) finds no improvement relative to the range-based estimator. For this reason we stick to the simpler, and arguably more robust, estimator $\hat{v}_{i,t-1}^{\text{OHLC}}$. This choice, however, still allows us to retain a comparative advantage over Molnár (2016): unlike he, we incorporate overnight jumps that account for a quarter of total variance.

⁵For example, see Equation (9) of Molnár (2016).

5 Empirical Analysis

5.1 Data and General Portfolio-Construction Rules

We download daily stock return data from the Center for Research in Security Prices (CRSP) starting on 01/01/1994 and ending on 12/31/2018. We restrict attention to stocks from the NYSE, AMEX, and NASDAQ stock exchanges. We also download daily OHLC price data (in dollars per share).

For simplicity, and in line with literature, we adopt the common convention that 21 consecutive trading days constitute one (trading)‘month’. The out-of-sample period ranges from 12/18/1998 through 12/31/2018, resulting in a total of 240 months (or 5,040 days). All portfolios are updated monthly.⁶ We denote the investment dates by $k = 1, \dots, 240$. At any investment date k , a covariance matrix is estimated based on the most recent 1260 daily returns, which roughly corresponds to using five years of past data.

We consider the following portfolio sizes $N \in \{100, 500, 1000\}$. For a given combination (k, N) , the investment universe is obtained as follows. We find the set of stocks that have an almost complete return history over the most recent $T = 1260$ days as well as a complete return ‘future’ over the next 21 days.⁷

Additionally, we require every stock in the universe to have all the data listed in Section 2.1 available at least 90% of the time and either high/low or open available at least 95% of the time.

We then look for possible pairs of highly correlated stocks, that is, pairs of stocks that have returns with a sample correlation exceeding 0.95 over the past 1260 days. In such pairs, if they should exist, we remove the stock with the lower market capitalization of the two on investment date k .⁸ Of the remaining set of stocks, we then pick the largest N stocks (as measured by their market capitalization on investment date k) as our investment universe. In this way, the investment universe changes relatively slowly from one investment date to the next.

There is a great advantage in having a well-defined rule that does not involve drawing stocks at random, as such a scheme would have to be replicated many times and averaged over to give stable results. As far as rules go, the one we have chosen seems the most reasonable because it avoids so-called “penny stocks” whose behavior is often erratic; also, high-market-cap stocks tend to have the lowest bid-ask spreads and the highest depth in the order book, which allows large investment funds to invest in them without breaching standard safety guidelines.

⁶Monthly updating is common practice to avoid an unreasonable amount of turnover and thus transaction costs. During a month, from one day to the next, we hold number of shares fixed rather than portfolio weights; in this way, there are no transactions at all during a month.

⁷The first restriction allows for up to 2.5% of missing returns over the most recent 1260 days, and replaces missing values by zero. The latter, ‘forward-looking’ restriction is not a feasible one in real life but is commonly applied in the related finance literature on the out-of-sample evaluation of portfolios.

⁸The reason is that we do not want to include highly similar stocks. In the early years, there are no such pairs; in the most recent years, there are never more than three such pairs.

In the best-case scenario where all the price data are available for a given stock on day $t - 1$, we use $\hat{v}_{i,t-1}^{\text{OHL C}}$ as the volatility proxy $\hat{v}_{i,t-1}$. If high/low are available but open is missing, we use $\hat{v}_{i,t-1}^{\text{HL}}$. If high/low are missing but open is available, we use $\hat{v}_{i,t-1}^{\text{OC}}$. If both high/low and open are missing, we go back to the traditional setting and use $\hat{v}_{i,t-1}^{\text{CC}}$.

5.2 Competing Models

The following models are included in our empirical analysis:

- **DCC**: the multivariate GARCH model of Engle (2002).
- **ID-DCC**: a model based on DCC with intraday-based volatility proxy in the first step; see formula (3.1).
- **IDR-DCC**: as ID-DCC but, additionally, with regularized returns as underlying time series for estimating DCC dynamics in the second step; see formula (3.3) and Appendix B.
- **DCC-NL**: the multivariate GARCH model of Engle et al. (2019) where the unconditional correlation matrix C is estimated via nonlinear shrinkage.
- **ID-DCC-NL**: a model based on DCC-NL with intraday-based volatility proxy in the first step; see formula (3.1).
- **IDR-DCC-NL**: as ID-DCC-NL but, additionally, with regularized returns as underlying time series for estimating DCC dynamics in the second step; see formula (3.3) and Appendix B.

Note that in the acronyms of the new models proposed, “ID” stands for “volatility proxy based on IntraDay data” and “R” stands for “regularized returns”.

5.3 Global Minimum Variance Portfolio

We consider the problem of estimating the global minimum variance (GMV) portfolio in the absence of short-sales constraints. The problem is formulated as

$$\min_w w' \Sigma_t w \quad (5.1)$$

$$\text{subject to } w' \mathbb{1} = 1, \quad (5.2)$$

where $\mathbb{1}$ denotes a vector of ones of dimension $N \times 1$. It has the analytical solution

$$w = \frac{\Sigma_t^{-1} \mathbb{1}}{\mathbb{1}' \Sigma_t^{-1} \mathbb{1}}. \quad (5.3)$$

The natural strategy in practice is to replace the unknown Σ_t by an estimator $\hat{\Sigma}_t$ in formula (5.3), yielding a feasible portfolio

$$\hat{w} := \frac{\hat{\Sigma}_t^{-1} \mathbb{1}}{\mathbb{1}' \hat{\Sigma}_t^{-1} \mathbb{1}}. \quad (5.4)$$

Estimating the GMV portfolio is a ‘clean’ problem in terms of evaluating the quality of a covariance matrix estimator, as it abstracts from having to estimate the vector of expected returns at the same time. In addition, researchers have established that estimated GMV portfolios have desirable out-of-sample properties not only in terms of risk but also in terms of reward-to-risk, that is, in terms of the information ratio; for example, see [Haugen and Baker \(1991\)](#), [Jagannathan and Ma \(2003\)](#), and [Nielsen and Aylursubramanian \(2008\)](#). As a result, such portfolios have become an addition to the large array of products sold by the mutual-fund industry.

In addition to Markowitz portfolios based on formula (5.4), we also include as a simple-minded benchmark the equal-weighted portfolio promoted by [DeMiguel et al. \(2009\)](#), among others, as it has been claimed to be difficult to outperform. We denote the equal-weighted portfolio by $\mathbf{1}/N$.

We report the following three out-of-sample performance measures for each scenario. (All of them are annualized and in percent for ease of interpretation.)

- **AV:** We compute the average of the 5,040 out-of-sample returns and then multiply by 252 to annualize.
- **SD:** We compute the standard deviation of the 5,040 out-of-sample returns and then multiply by $\sqrt{252}$ to annualize.
- **IR:** We compute the (annualized) information ratio as the ratio AV/SD.

Our stance is that in the context of the GMV portfolio, the most important performance measure is the out-of-sample standard deviation, SD. The true (but unfeasible) GMV portfolio is given by (5.3). It is designed to minimize the variance (and thus the standard deviation) rather than to maximize the expected return or the information ratio. Therefore, any portfolio that implements the GMV portfolio should be primarily evaluated by how successfully it achieves this goal. A high out-of-sample average return, AV, and a high out-of-sample information ratio, IR, are naturally also desirable, but should be considered of secondary importance from the point of view of evaluating the quality of a covariance matrix estimator.

We also consider the question of whether one estimation model delivers a lower out-of-sample standard deviation than another estimation model. As we compare 7 models, there are 21 pairwise comparisons. To avoid a (large) multiple testing problem and as a major goal of this paper is to show that using higher-frequency (intraday) data improves the estimation of large-dimensional covariances matrices, we restrict attention to two comparisons: (i) DCC with IDR-DCC and (ii) DCC-NL with IDR-DCC-NL. For a given universe size, a two-sided p -value for the null hypothesis of equal standard deviations is obtained by the prewhitened HAC_{PW} method described in [Ledoit and Wolf \(2011, Section 3.1\)](#).⁹

The results are presented in Table 1 as well as in Figure 4 and can be summarized as follows; unless stated otherwise, the findings are with respect to SD as performance measure.

⁹As the out-of-sample size is very large at 5,040, there is no need to use the computationally more involved bootstrap method described in [Ledoit and Wolf \(2011, Section 3.2\)](#), which is preferred for small sample sizes.

- All models consistently outperform $1/N$ by a wide margin.
- Each DCC-NL model consistently outperforms its DCC version.
- Each intraday model, IDR-DCC(-NL) and ID-DCC(-NL), consistently outperforms its traditional base model, DCC(-NL). Additionally, IDR-DCC(-NL) consistently outperforms its ID-DCC(-NL) version.
- Moreover, the outperformance of IDR-DCC-NL over DCC-NL, respectively IDR-DCC over DCC, is always statistically significant and also economically meaningful.
- There is a consistent ranking across all portfolio sizes N (from best to worst): IDR-DCC-NL, ID-DCC-NL, IDR-DCC, ID-DCC, DCC-NL, DCC, $1/N$.¹⁰

To sum up, models using intraday data such as IDR-DCC(-NL) and ID-DCC(-NL) consistently outperform the traditional DCC(-NL) models using daily data only. Furthermore IDR models, which use regularized returns in the second step, outperform their ID counterparts.

Remark 5.1 (Subperiod analysis). Table 1 presents ‘single’ results over the entire out-of-sample period 12/18/1998–12/31/2018. It might be natural to ask whether the relative performance of the various models is stable during that period or whether it changes during certain subperiods, such as periods of ‘boom’ vs. periods of ‘bust’. To address this question via a robustness check, we carry out a rolling-window analysis based on shorter out-of-sample periods: one month (21 days), one year (252 days), and five years (1260 days). The results are displayed in Figure 5, where any given number represents the out-of-sample standard deviation (SD) over the corresponding subperiod ending on that day; the universe size is $N = 1000$ always. It can be seen that the relative performance is remarkably stable over time and that, in particular, IDR-DCC-NL generally performs best during all subperiods. ■

Remark 5.2 (Choice of κ). In Tables 3 and 4, we examine how robust IDR-DCC(-NL) models are in terms of κ and that the choice $\kappa = 100$, which was previously motivated by theoretical reasons, actually works well in practice. In terms of the performance measure SD, the choice $\kappa = 100$ indeed works generally best (and always best for large universe sizes $N > 100$). In terms of the performance measure IR, the choice $\kappa = 100$ works generally best as well. Furthermore, in terms of both measures, performance is not very robust to the choice of κ and it seems more harmful to pick a too small value compared to picking a too large value. Overall, there is strong empirical support for the choice $\kappa = 100$. ■

Remark 5.3 (Further comparisons). Another way to improve upon the DCC-NL model, while sticking to daily observed returns, is to combine it with an approximate factor model. In

¹⁰With the single exception of $N = 1000$ where DCC-NL outperforms ID-DCC. Thus, for ‘medium’-sized investment universes, $N = 100, 500$, even the proposed DCC models using intraday data outperform the traditional DCC-NL model. However, for larger dimensions, $N = 1000$, the benefit of nonlinear shrinkage is too important to neglect; see Figure 4.

particular, [De Nard et al. \(2020\)](#) suggest the AFM1-DCC-NL model which is based on an approximate single-factor model, where the single factor is the market factor (that is, the first factor of any Fama-French factor model). Based on results not reported in detail here, we found that the performance of this model (in terms of SD) is generally somewhere between the performances of DCC-NL and ID-DCC, and always below the performance of IDR-DCC-NL.

There are more complicated multivariate volatility models based on intradaily data. For example, one can use all 5- or 10-minute returns during the day to compute realized covariance matrices and then use a sophisticated methodology, involving factor structure and regularization, to make corresponding forecasts. Using such an approach, [Brito et al. \(2018\)](#) obtain a reduction of 22.1 percentage points in SD (from 10.65 to a typical number of 8.3) compared to DCC-NL for a (constant) universe of $N = 430$ stocks; see their Table 5. Note that with our (much) simpler IDR-DCC-NL model, we get almost the same reduction in SD for a (time-varying) universe of $N = 500$ stocks, namely a reduction of 17.4 percentage points (from 9.01 to 7.44).¹¹ But unlike the IDR-DCC-NL method, it is doubtful whether the method of [Brito et al. \(2018\)](#) is computationally feasible for universe sizes of $N \geq 1000$, even leaving aside concerns about data availability (because, for a given stock, they need all 5-minute returns during a day whereas we, at most, need only the four values OHLC). ■

[DeMiguel et al. \(2009\)](#) claim that it is difficult to outperform $1/N$ in terms of the out-of-sample Sharpe ratio with sophisticated portfolios (that is, with Markowitz portfolios that estimate input parameters). It can be seen that all models consistently outperform $1/N$ in terms of the out-of-sample information ratio, which translates into outperformance in terms of the out-of-sample Sharpe ratio. For $N = 100$, IDR-DCC is best overall, whereas for $N = 500, 1000$, IDR-DCC-NL is best overall.

Additionally, we report results on average turnover and leverage, defined as follows.

- **TO:** We compute average (monthly) turnover as $\frac{1}{239} \sum_{k=1}^{239} \|\hat{w}_{k+1} - \hat{w}_k^{\text{hold}}\|_1$, where $\|\cdot\|_1$ denotes the L^1 norm and \hat{w}_k^{hold} denotes the vector of the ‘hold’ portfolio weights at the end of month k .¹²
- **GL:** We compute average (monthly) gross leverage as $\frac{1}{240} \sum_{k=1}^{240} \|\hat{w}_k\|_1$.
- **PL:** We compute average (monthly) proportion of leverage as $\frac{1}{240 \times N} \sum_{k=1}^{240} \sum_{i=1}^N \mathbb{1}_{\{\hat{w}_{i,k} < 0\}}$, where $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function.

The results are presented in Table 2 and can be summarized as follows; unless stated otherwise, the findings are with respect to the average monthly turnover as performance measure. Note that we do not constrain for the amounts of leverage and turnover in our optimization.

¹¹It makes sense to compare reduction in percentage points rather than nominal reduction, as the actual SD numbers of [Brito et al. \(2018\)](#) are not one-to-one comparable to ours because of different universes and different out-of-sample periods.

¹²The vector \hat{w}_k^{hold} is determined by the initial vector of portfolio weights, \hat{w}_k , together with the evolution of the various prices of the N stocks in the portfolio during month k .

- IDR-DCC-NL consistently and markedly outperforms all other models in terms of turnover and gross leverage.
- IDR-DCC consistently outperforms all other models in terms of proportion of leverage, although the differences are not large.
- Using regularized returns instead of observed returns in the estimation of the DCC(-NL) models consistently reduces turnover and gross leverage, but has no noticeable effect on percentage of leverage.

5.4 Markowitz Portfolio with Momentum Signal

We now turn attention to a ‘full’ Markowitz portfolio with a signal.

By now a large number of variables have been documented that can be used to construct a signal in practice. For simplicity and reproducibility, we use the well-known momentum factor (or simply momentum for short) of [Jegadeesh and Titman \(1993\)](#). For a given investment period k and a given stock, the momentum is the geometric average of the previous 252 returns on the stock but excluding the most recent 21 returns; in other words, one uses the geometric average over the previous ‘year’ but excluding the previous ‘month’. Collecting the individual momentums of all the N stocks contained in the portfolio universe yields the return-predictive signal, denoted by m .

In the absence of short-sales constraints, the investment problem is formulated as

$$\min_w w' \Sigma_t w \quad (5.5)$$

$$\text{subject to } w' m_t = b, \text{ and} \quad (5.6)$$

$$w' \mathbb{1} = 1, \quad (5.7)$$

where b is a selected target expected return. The problem has the analytical solution

$$w = c_1 \Sigma_t^{-1} \mathbb{1} + c_2 \Sigma_t^{-1} m_t, \quad (5.8)$$

$$\text{where } c_1 := \frac{C - bB}{AC - B^2} \quad \text{and} \quad c_2 := \frac{bA - B}{AC - B^2}, \quad (5.9)$$

$$\text{with } A := \mathbb{1}' \Sigma_t^{-1} \mathbb{1}, \quad B := \mathbb{1}' \Sigma_t^{-1} m_t, \quad \text{and} \quad C := m_t' \Sigma_t^{-1} m_t. \quad (5.10)$$

The natural strategy in practice is to replace the unknown Σ_t by an estimator $\widehat{\Sigma}_t$ in formulas (5.8)–(5.10), yielding a feasible portfolio

$$\widehat{w} := c_1 \widehat{\Sigma}_t^{-1} \mathbb{1} + c_2 \widehat{\Sigma}_t^{-1} m_t, \quad (5.11)$$

$$\text{where } c_1 := \frac{C - bB}{AC - B^2} \quad \text{and} \quad c_2 := \frac{bA - B}{AC - B^2}, \quad (5.12)$$

$$\text{with } A := \mathbb{1}' \widehat{\Sigma}_t^{-1} \mathbb{1}, \quad B := \mathbb{1}' \widehat{\Sigma}_t^{-1} m_t, \quad \text{and} \quad C := m_t' \widehat{\Sigma}_t^{-1} m_t. \quad (5.13)$$

In addition to Markowitz portfolios based on formulas (5.11)–(5.13), we also include as a simple-minded benchmark the equal-weighted portfolio among the top-quintile stocks (according

to momentum). This portfolio is obtained by sorting the stocks, from lowest to highest, according to their momentum and then putting equal weight on all the stocks in the top 20%, that is, in the top quintile. We call this portfolio **EW-TQ**.

Our stance is that in the context of a ‘full’ Markowitz portfolio, the most important performance measure is the out-of-sample information ratio, IR. In the ‘ideal’ investment problem (5.8)–(5.10), minimizing the variance (for a fixed target expected return b) is equivalent to maximizing the information ratio (for a fixed target expected return b). In practice, because of estimation error in the signal, the various strategies do not necessarily have the same expected return and, thus, focusing on the out-of-sample standard deviation is inappropriate.

We also consider the question whether IDR-DCC(-NL) delivers a higher out-of-sample information ratio than DCC(-NL) at a level that is statistically significant with the same reason as discussed in Section 5.3. For a given universe size, a two-sided p -value for the null hypothesis of equal information ratios is obtained by the prewhitened HAC_{PW} method described in Ledoit and Wolf (2008, Section 3.1).¹³

The results are presented in Table 5 and can be summarized as follows; unless stated otherwise, the findings are with respect to IR as performance measure.

- All models consistently outperform EWTQ by a wide margin.
- Each DCC-NL model consistently outperforms its DCC counterpart.
- IDR-DCC-NL consistently outperforms all other models.
- Having said that, we do not find statistical significance (for the two comparisons considered).
- Moreover, in terms of SD as performance measure, we find statistical significance (for the two comparisons considered) for $N = 500, 1000$, with IDR-DCC-NL again being the best model. the outperformance of IDR-DCC-NL over DCC-NL, respectively IDR-DCC over DCC, in terms of lower SD is always statistically significant and also economically meaningful.

DeMiguel et al. (2009) claim that it is difficult to outperform $1/N$ in terms of the out-of-sample Sharpe ratio with sophisticated portfolios (that is, with Markowitz portfolios that estimate input parameters). Comparing with Table 1, it can be seen that all models based on the (simple-minded) momentum signal consistently outperform $1/N$ in terms of the out-of-sample information ratio, which translates into outperformance in terms of the out-of-sample Sharpe ratio. Even though momentum is not a very powerful return-predictive signal, the differences compared to $1/N$ can be enormous. For example, for $N = 1000$, the information ratio of $1/N$ is only 0.54 whereas the information ratio of IDR-DCC-NL is 1.00, almost twice as large.

Engle and Colacito (2006) argue for the use of the out-of-sample standard deviation, SD, as a performance measure also in the context of a full Markowitz portfolio. Also for this alternative

¹³As the out-of-sample size is very large at 5,040, there is no need to use the computationally more expensive bootstrap method described in Ledoit and Wolf (2008, Section 3.2), which is preferred for small sample sizes.

performance measure, the high-frequency estimators IDR-DCC(-NL) and ID-DCC(-NL) perform the best.

Additionally, we report results on average turnover and leverage. The results are presented in Table 6 and can be summarized as follows; unless stated otherwise, the findings are with respect to the average monthly turnover as performance measure. Note that we do not constrain for the amounts of leverage and turnover in our optimization.

- IDR-DCC-NL consistently and markedly outperforms all other models in terms of turnover and gross leverage.
- IDR-DCC consistently outperforms all other models in terms of proportion of leverage, although the differences are not large.
- Using regularized returns instead of observed returns in the estimation of the DCC(-NL) models consistently reduces turnover and gross leverage, but has no noticeable effect on percentage of leverage.

Remark 5.4 (Transaction costs). We do not provide tables with performance measures net of transaction costs for two reasons. First, we do not impose constraints on turnover, leverage, or transactions costs in any of our portfolio optimization. Of course, such constraints would be used, to varying degrees, by real-life portfolio managers; but the main point of our paper is to study the accuracy of various estimators of the covariance matrix, a problem that does not depend on transaction costs. Second, there is always disagreement which transaction cost to use. Many finance papers, at least in the past, have used a transaction cost of 50 bps. But in this day and age, the average transaction cost is usually south of 5 bps for managers trading the 1000 most liquid US stocks.

At any rate, given the results presented in Tables 1–6, the reader can get a rough idea of the various performance measures net of transaction costs, for any choice of transaction cost, according to the rule of thumb that the return loss (per month) due to turnover is twice the amount of turnover times the chosen transaction cost. For example, assuming a transaction cost of 5 bps, a turnover of one would result in a return loss of 10 bps (per month) according to this rule. ■

6 Conclusion

In this paper we have shown that using intraday data, in the form of open/high/low/close (OHLC) prices, leads to improved forecasting of dynamic covariance matrices via multivariate GARCH models, where our focus has been on the original DCC model of Engle (2002) and its recent extension, the DCC-NL model of Engle et al. (2019).

The first step of a DCC(-NL) model consists of modeling dynamic univariate variances via a GARCH(1,1) model, where daily squared returns are used as innovations. Hence, the first idea

is to use a less noisy volatility proxy based on OHLC data, instead of squared returns, in this step. We call the resulting multivariate GARCH models ID-DCC(-NL), where ID stands for “volatility proxy based on IntraDay data”.

The second step of a DCC(-NL) model consists of modeling the dynamic (pseudo-)correlation matrix where the innovation is the outer product of the vector of daily (devolitized) returns. Hence, the second idea is to find a better innovation that is based on volatility proxies instead of returns. The counterpart of a volatility proxy itself is the squared return. Thus, it is natural to take the square root of the volatility proxy, together with sign of the return, as the counterpart of the return itself. This already works well, but a certain ‘smoothed’ sign of return, together with the square root of the volatility proxy works even better, and we call the resulting quantity *regularized return*. Using these regularized returns in the second step, and of course also using a volatility proxy based on intraday data in the first step, gives rise to multivariate GARCH models that we call IDR-DCC(-NL) models, where R stands for “regularized returns”.

An important feature of our newly proposed models is that they remain computationally feasible for universes of $N \geq 1000$ assets, unlike many other multivariate GARCH-type models based on intraday data.

Empirical backtest exercises using real-life stock data demonstrate that ID-DCC(-NL) models already deliver a substantial improvement over DCC(-NL) models, but that IDR-DCC(-NL) models deliver another improvement of roughly the same magnitude. In particular, the IDR-DCC-NL model is the clear winner and its performance for large investment universes is rather impressive; for example, the out-of-sample standard deviation of an estimated global-minimum-variance (GMV) portfolio of $N = 1000$ stocks improves from 7.88% to 5.90% (on an annual basis) when one upgrades from DCC-NL to IDR-DCC-NL. A further advantage of the IDR-DCC-NL model is that it leads to both reduced turnover and reduced gross leverage compared to the DCC-NL model.

References

- Baillie, R. T. and Bollerslev, T. (1992). Prediction in dynamic models with time-dependent conditional variances. *Journal of Econometrics*, 52:91–113.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *Econometrics Journal*, 12:C1–C32.
- Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327.
- Brito, D., Medeiros, M. C., and Ribeiro, R. (2018). Forecasting large realized covariance matrices: The benefits of factor models and shrinkage. Technical report, Department of Economics, PUC-Rio. Available at <https://ssrn.com/abstract=3163668>.
- Christoffersen, P. and Diebold, F. (2006). Financial asset returns, direction-of-change forecasting, and volatility dynamics. *Management Science*, 52(8):1273–1287.
- De Nard, G., Ledoit, O., and Wolf, M. (2020). Factor models for portfolio selection in large dimensions: The good, the better and the ugly. *Journal of Financial Econometrics*. Forthcoming.
- DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the $1/N$ portfolio strategy? *Review of Financial Studies*, 22:1915–1953.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007.
- Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350.
- Engle, R. F. and Colacito, R. (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business & Economic Statistics*, 24(2):238–253.
- Engle, R. F., Ledoit, O., and Wolf, M. (2019). Large dynamic covariance matrices. *Journal of Business & Economic Statistics*, 37(2):363–375.
- Engle, R. F. and Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Working Paper 8554, National Bureau of Economic Research.

- Garman, M. B. and Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, 53(1):67–78.
- Hansen, P. R., Huang, Z., and Shek, H. H. (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906.
- Haugen, R. A. and Baker, N. L. (1991). The efficient market inefficiency of capitalizationweighted stock portfolios. *Journal of Portfolio Management*, 17(3):35–40.
- Henriksson, R. and Merton, R. (1981). On market timing and investment performance. *Journal of Business*, 54(4):513–533.
- Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, 54(4):1651–1684.
- Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1):65–91.
- Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15:850–859.
- Ledoit, O. and Wolf, M. (2011). Robust performance hypothesis testing with the variance. *Wilmott Magazine*, September:86–89.
- Ledoit, O. and Wolf, M. (2015). Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions. *Journal of Multivariate Analysis*, 139(2):360–384.
- Ledoit, O. and Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. *Review of Financial Studies*, 30(12):4349–4388.
- Ledoit, O. and Wolf, M. (2020). Analytical nonlinear shrinkage of large-dimensional covariance matrices. *Annals of Statistics*. Forthcoming.
- Molnár, P. (2016). High-low range in GARCH models of stock return volatility. *Applied Economics*, 48(51):4977–4991.
- Nielsen, F. and Aylursubramanian, R. (2008). Far from the madding crowd — Volatility efficient indices. Research insights, MSCI Barra.
- Noureldin, D., Shephard, N., and Sheppard, K. (2012). Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 27:907–933.
- Pakel, C., Shephard, N., Sheppard, K., and Engle, R. F. (2020). Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics*. Forthcoming.

- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65.
- Pesaran, H. and Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *Journal of Finance*, 50(4):1201–1228.
- Welch, I. and Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4):1455–1508.
- Xu, Y. (2019). DCC-HEAVY: A multivariate GARCH model with realized measures of variance and correlation. Cardiff Economics Working Paper E2019/5, Cardiff Business School.
- Yang, D. and Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business*, 73(3):477–492.

A Figures and Tables

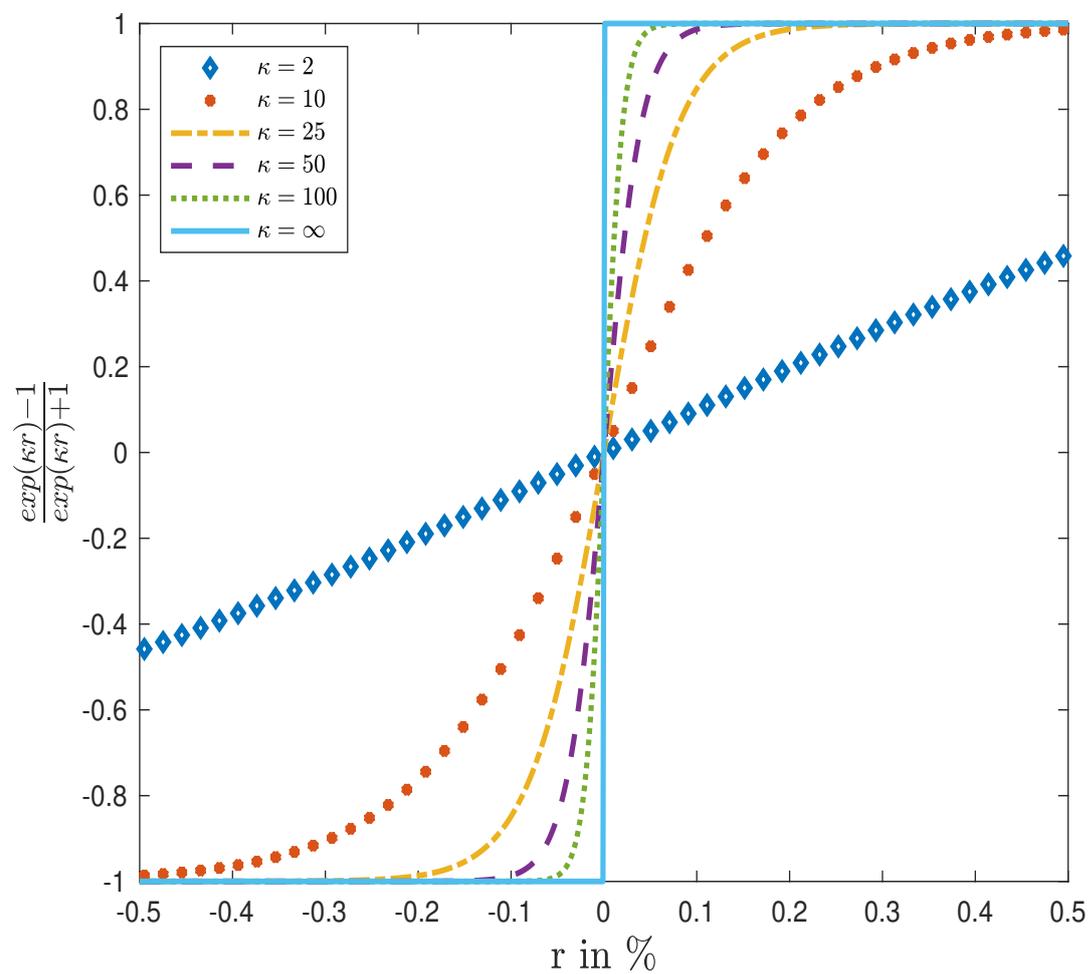


Figure 1: The scaled hyperbolic tangent function: some examples for various values of κ .

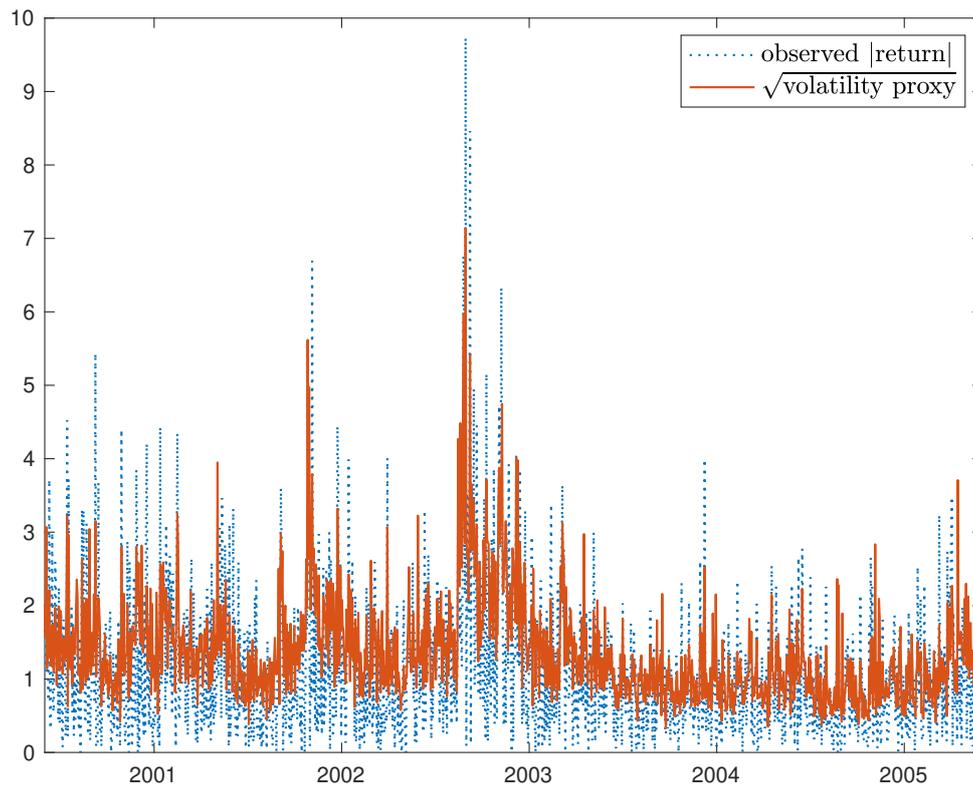


Figure 2: Observed absolute returns and the square root of the volatility proxies (in percent) for a representative stock (Exxon Mobil Corporation) and period (4/19/2000–4/26/2005) used to model the conditional univariate variances for DCC(-NL) respectively ID-DCC(-NL); see Equation (2.1) respectively Equation (3.1).

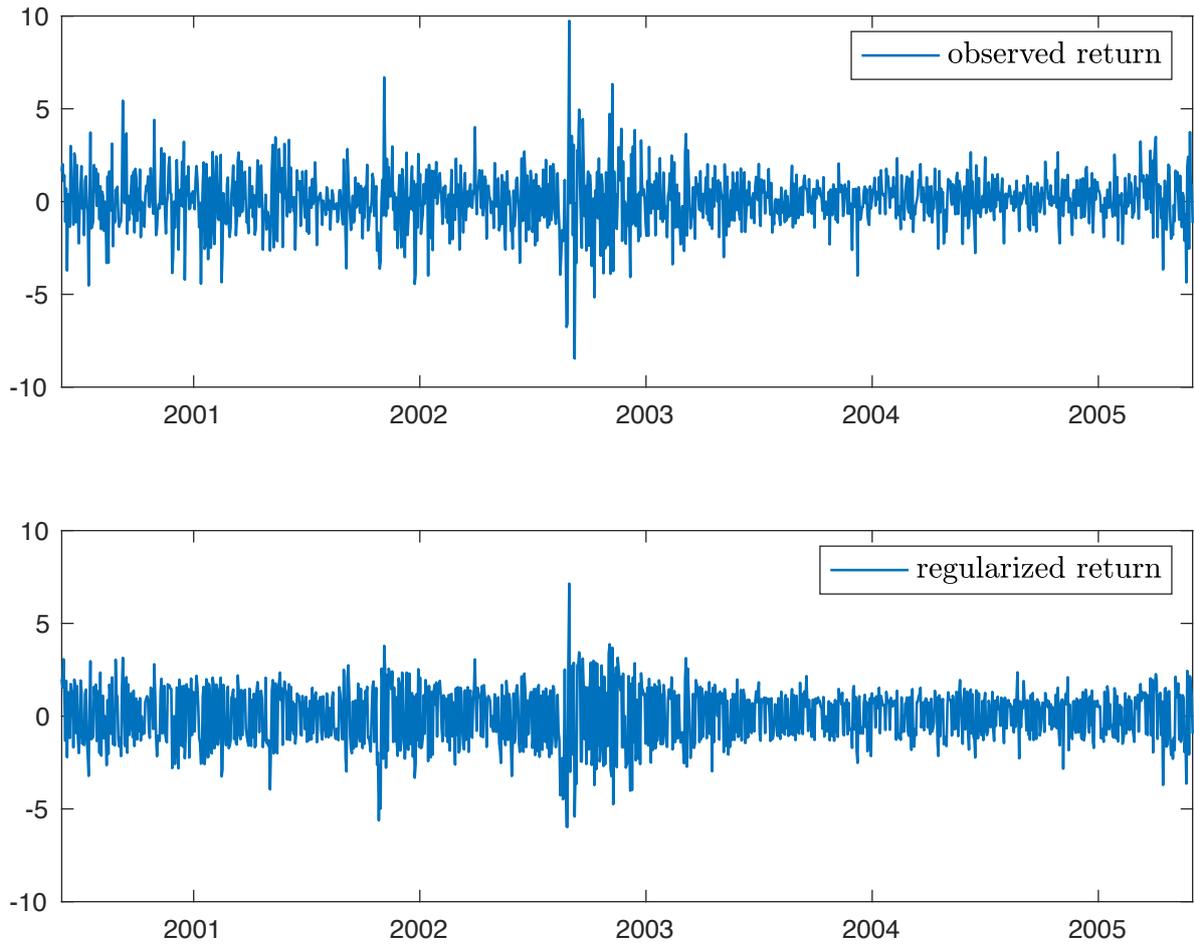


Figure 3: Observed returns and regularized returns (in percent) for a representative stock (Exxon Mobil Corporation) and period (4/19/2000–4/26/2005).

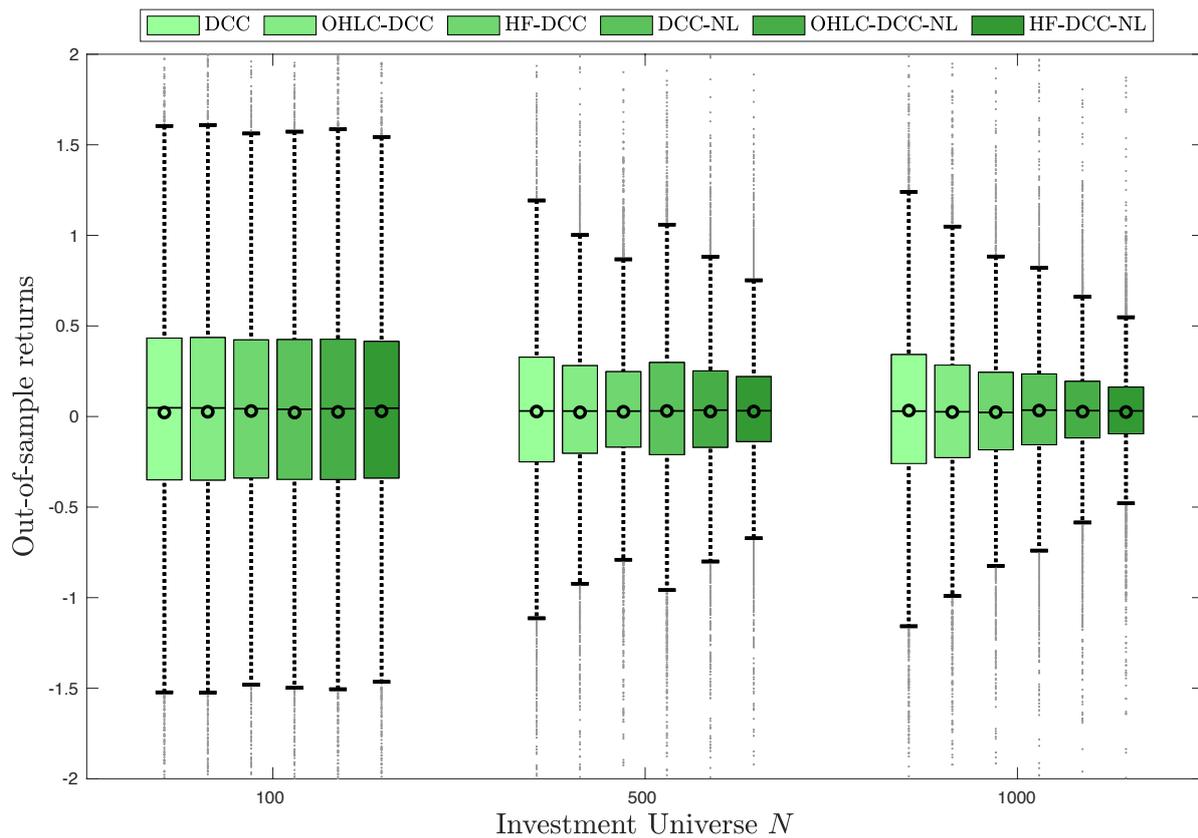


Figure 4: Boxplots of the 5,040 daily out-of-sample returns (in percent) for various estimators of the GMV portfolio; the period is 12/18/1998–12/31/2018. The relative benefit of using intraday data (ID respectively IDR) and nonlinear shrinkage (NL) gets more pronounced for larger dimensions.

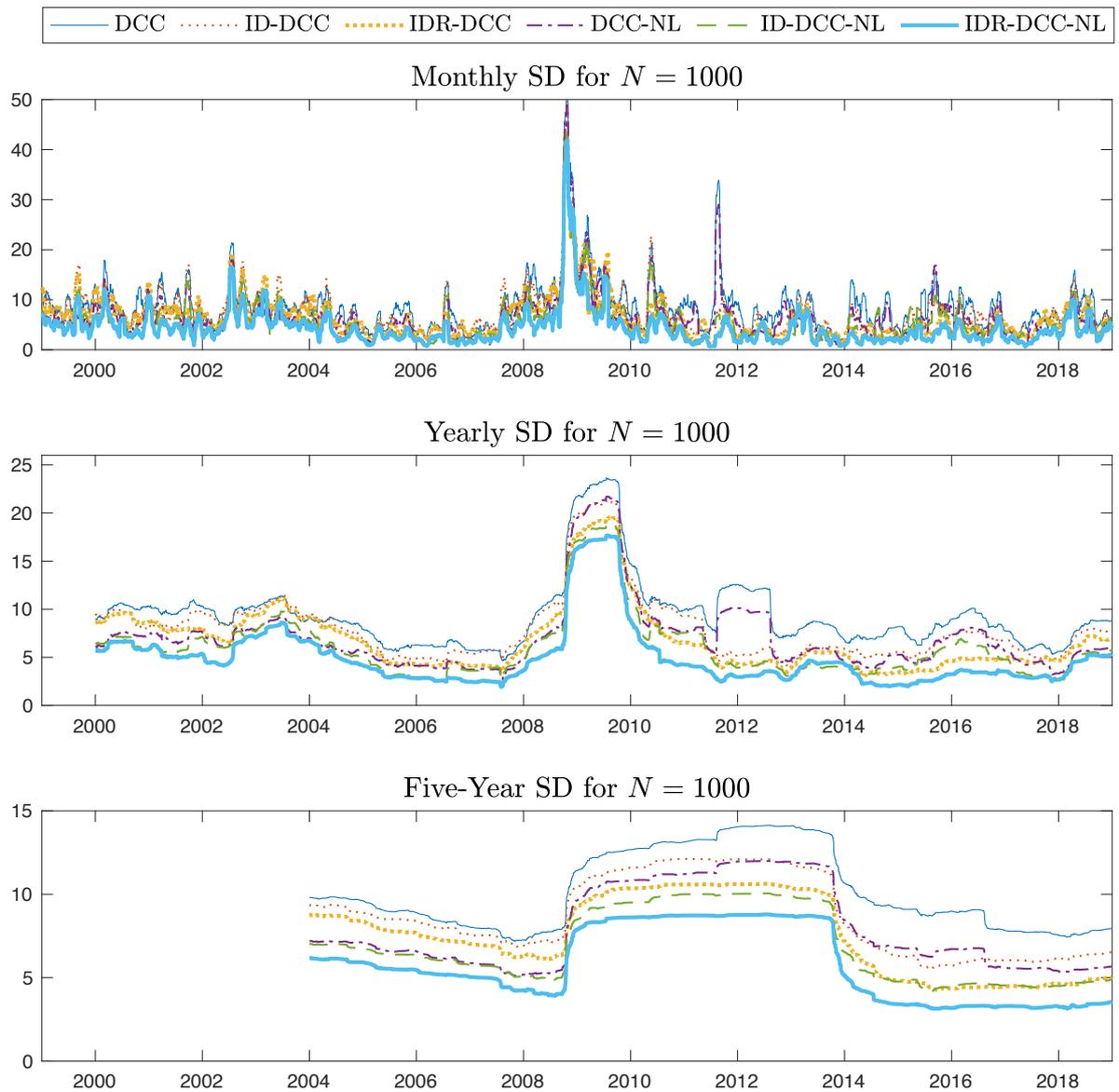


Figure 5: Rolling-window out-of-sample standard deviations (SD) for various models and $N = 1000$. The lengths of the out-of-sample period are one month, one year, and five years, respectively. Any given number represents the out-of-sample standard deviation (SD) over the corresponding subperiod ending on that day.

Period: 12/18/1998–12/31/2018							
	1/N	DCC	ID-DCC	IDR-DCC	DCC-NL	ID-DCC-NL	IDR-DCC-NL
$N = 100$							
AV	7.54	5.57	6.63	7.66	5.52	6.43	7.29
SD	18.94	12.45	12.24	12.18**	12.33	12.14	12.07**
IR	0.40	0.45	0.54	0.63	0.45	0.53	0.60
$N = 500$							
AV	10.05	6.94	5.88	6.62	7.75	6.91	7.04
SD	19.50	9.67	8.81	7.94***	9.01	8.19	7.44***
IR	0.52	0.72	0.67	0.83	0.86	0.84	0.95
$N = 1000$							
AV	10.91	8.33	6.20	6.08	8.32	6.76	6.29
SD	20.26	10.05	8.76	7.72***	7.88	6.85	5.90***
IR	0.54	0.83	0.71	0.79	1.06	0.99	1.07

Table 1: Annualized performance measures (in percent) for various estimators of the GMV portfolio. AV stands for average; SD stands for standard deviation; and IR stands for information ratio. All measures are based on 5,040 daily out-of-sample returns. In the rows labeled SD, the lowest number appears in **bold face**. In the columns labeled IDR-DCC respectively IDR-DCC-NL, significant outperformance over DCC respectively DCC-NL in terms of SD is denoted by asterisks: *** denotes significance at the 0.01 level; ** denotes significance at the 0.05 level; and * denotes significance at the 0.1 level.

Period: 12/18/1998–12/31/2018						
	DCC	ID-DCC	IDR-DCC	DCC-NL	ID-DCC-NL	IDR-DCC-NL
$N = 100$						
TO	2.84	3.06	2.44	2.64	2.87	2.26
GL	3.20	3.19	2.54	3.00	3.00	2.37
PL	0.47	0.46	0.45	0.46	0.46	0.45
$N = 500$						
TO	4.56	4.47	3.83	3.05	3.01	2.57
GL	4.99	4.41	3.83	3.46	3.12	2.62
PL	0.49	0.49	0.49	0.51	0.50	0.51
$N = 1000$						
TO	7.72	7.20	6.20	2.71	2.87	2.56
GL	7.90	6.80	6.07	3.07	2.83	2.54
PL	0.50	0.49	0.49	0.51	0.51	0.51

Table 2: Average monthly turnover and leverage for various estimators of the GMV portfolio. TO stands for average turnover; GL stands for average gross leverage; and PL stands for average proportion of leverage. All measures are based on 240 monthly weight vectors. In each row, the lowest (and thus best) number appears in **bold face**.

Period: 12/18/1998–12/31/2018					
IDR-DCC with					
	$\kappa = 2$	$\kappa = 50$	$\kappa = 100$	$\kappa = 200$	$\kappa = \infty$
$N = 100$					
AV	5.46	7.06	7.66	7.00	6.94
SD	12.59	12.18	12.18	12.24	12.17
IR	0.43	0.58	0.63	0.57	0.57
$N = 500$					
AV	6.35	5.99	6.62	5.88	6.84
SD	9.45	8.13	7.94***	8.08	8.60
IR	0.67	0.74	0.83	0.73	0.79
$N = 1000$					
AV	5.98	5.51	6.08	6.26	7.23
SD	8.80	7.88	7.72***	8.12	8.73
IR	0.68	0.70	0.79	0.77	0.82

Table 3: Annualized performance measures (in percent) for various IDR-DCC estimators of the GMV portfolio. AV stands for average; SD stands for standard deviation; and IR stands for information ratio. All measures are based on 5,040 daily out-of-sample returns. In the rows labeled SD, the lowest number appears in **bold face**. In the column labeled $\kappa = 100$, significant outperformance over $\kappa = \infty$ in terms of SD is denoted by asterisks: *** denotes significance at the 0.01 level; ** denotes significance at the 0.05 level; and * denotes significance at the 0.1 level.

Period: 12/18/1998–12/31/2018					
IDR-DCC-NL with					
	$\kappa = 2$	$\kappa = 50$	$\kappa = 100$	$\kappa = 200$	$\kappa = \infty$
$N = 100$					
AV	5.35	6.78	7.29	6.69	6.64
SD	12.47	12.07	12.07	12.13	12.07
IR	0.43	0.56	0.60	0.55	0.55
$N = 500$					
AV	6.80	6.75	7.04	6.65	7.62
SD	9.03	7.69	7.44***	7.50	8.14
IR	0.75	0.88	0.95	0.89	0.94
$N = 1000$					
AV	6.25	6.11	6.29	6.65	7.15
SD	7.72	6.13	5.90***	6.18	6.84
IR	0.81	1.00	1.07	1.08	1.04

Table 4: Annualized performance measures (in percent) for various IDR-DCC-NL estimators of the GMV portfolio. AV stands for average; SD stands for standard deviation; and IR stands for information ratio. All measures are based on 5,040 daily out-of-sample returns. In the rows labeled SD, the lowest number appears in **bold face**. In the column labeled $\kappa = 100$, significant outperformance over $\kappa = \infty$ in terms of SD is denoted by asterisks: *** denotes significance at the 0.01 level; ** denotes significance at the 0.05 level; and * denotes significance at the 0.1 level.

Period: 12/18/1998–12/31/2018							
	EW-TQ	DCC	ID-DCC	IDR-DCC	DCC-NL	ID-DCC-NL	IDR-DCC-NL
$N = 100$							
AV	9.75	9.41	10.23	10.76	9.63	10.38	10.69
SD	23.49	14.89	14.71	14.72	14.77	14.61	14.65
IR	0.42	0.63	0.70	0.73	0.65	0.71	0.73
$N = 500$							
AV	11.82	7.77	6.28	7.03	8.80	7.56	7.80
SD	22.97	10.81	9.96	9.12***	10.13	9.29	8.58***
IR	0.51	0.72	0.63	0.77	0.87	0.81	0.91
$N = 1000$							
AV	12.02	8.36	6.34	6.03	8.51	6.74	6.54
SD	23.66	11.32	10.02	8.73***	8.64	7.52	6.54***
IR	0.51	0.74	0.63	0.69	0.98	0.90	1.00

Table 5: Annualized performance measures (in percent) for various estimators of the Markowitz portfolio with momentum signal. AV stands for average; SD stands for standard deviation; and IR stands for information ratio. All measures are based on 5,040 daily out-of-sample returns. In the rows labeled IR, the largest number appears in **bold face**. In the columns labeled IDR-DCC respectively IDR-DCC-NL, significant outperformance over DCC respectively DCC-NL in terms of SD and IR (separately) is denoted by asterisks: *** denotes significance at the 0.01 level; ** denotes significance at the 0.05 level; and * denotes significance at the 0.1 level.

Period: 12/18/1998–12/31/2018						
	DCC	ID-DCC	IDR-DCC	DCC-NL	ID-DCC-NL	IDR-DCC-NL
$N = 100$						
TO	3.40	3.54	2.86	3.17	3.34	2.66
GL	3.90	3.87	3.14	3.67	3.65	2.95
PL	0.47	0.47	0.46	0.47	0.47	0.46
$N = 500$						
TO	5.04	4.89	4.19	3.41	3.43	2.84
GL	5.68	5.04	4.35	3.99	3.59	3.02
PL	0.50	0.50	0.50	0.50	0.50	0.51
$N = 1000$						
TO	8.99	8.33	6.95	3.05	3.24	2.83
GL	9.58	8.20	6.98	3.59	3.33	2.93
PL	0.50	0.50	0.50	0.51	0.51	0.51

Table 6: Average monthly turnover and leverage for various estimators of the Markowitz portfolio with momentum signal. TO stands for average turnover; GL stands for average gross leverage; and PL stands for average proportion of leverage. All measures are based on 240 monthly weight vectors. In each row, the lowest (and thus best) number appears in **bold face**.

B Detailed Description of IDR-DCC(-NL) Models

First, compute the regularized returns for a suitable scaling factor κ :¹⁴

$$\tilde{r}_{i,t-1} := \text{stanh}(r_{i,t-1}, \kappa) \sqrt{\hat{v}_{i,t-1}} , \quad (\text{B.1})$$

where $\text{stanh}(\cdot, \cdot)$ is the ‘scaled’ hyperbolic tangent function defined in Equation (3.4) and $\hat{v}_{i,t-1}$ denotes a generic volatility proxy based on OHLC price data; see Section 4 for specific proposals.

Second, for modeling conditional covariances, use a GARCH(1,1) model with squared regularized-returns, instead of squared returns, as innovations:

$$d_{i,t}^2 = \omega_i + \delta_{1,i} \tilde{r}_{i,t-1}^2 + \delta_{2,i} d_{i,t-1}^2 , \quad (\text{B.2})$$

where $(\omega_i, \delta_{1,i}, \delta_{2,i})$ are the asset-specific GARCH(1,1) parameters. Now use the conditional variances to devolatilize the regularized returns: $s_{i,t} := \tilde{r}_{i,t} / d_{i,t}$.

Third, for modeling the conditional (pseudo) correlation matrix use the DCC model with correlation targeting:

$$Q_t = (1 - \delta_1 - \delta_2)C + \delta_1 s_{t-1} s'_{t-1} + \delta_2 Q_{t-1} , \quad (\text{B.3})$$

where (δ_1, δ_2) are the DCC parameters and $C := \text{Cor}(\tilde{r}_t) = \text{Cov}(s_t)$ denotes the unconditional correlation matrix of the regularized returns. Note that Q_t cannot be used directly because its diagonal elements, although close to one, are not exactly equal to one. From this representation, we obtain the conditional correlation matrix and the conditional covariance matrix as

$$R_t := \text{Diag}(Q_t)^{-1/2} Q_t \text{Diag}(Q_t)^{-1/2} \quad (\text{B.4})$$

$$\Sigma_t := D_t R_t D_t . \quad (\text{B.5})$$

Finally, if the portfolio is held for more than one day use the ‘averaged-forecasting’ approach of De Nard et al. (2020). At portfolio construction date k , forecast the covariance matrix for all L days of the portfolio holding period, that is, for $t = k, k+1, \dots, k+(L-1)$; then average those L forecasts and use this ‘averaged forecast’ to construct the portfolio at date k ; see Section 2.2:

$$\hat{\Sigma}_k := \frac{1}{L} \sum_{l=0}^{L-1} \mathbb{E}[D_{k+l} | \mathcal{F}_{k-1}] \mathbb{E}[R_{k+l} | \mathcal{F}_{k-1}] \mathbb{E}[D_{k+l} | \mathcal{F}_{k-1}] . \quad (\text{B.6})$$

In practice, the parameters in Equations (B.2)–(B.3) need to be estimated. To this end, we follow the same prescriptions as in Section 2.3, just with regularized returns \tilde{r}_t in place of returns r_t . In particular, if the correlation-targeting matrix C of Equation (B.2) is estimated using the sample covariance matrix of the devolatilized regularized-returns $\{s_t\}$, the IDR-DCC model results; if instead C is estimated by applying nonlinear shrinkage to the $\{s_t\}$, with post-processing to obtain a proper correlation matrix, the IDR-DCC-NL model results.

¹⁴We suggest the choice $\kappa = 100$ for returns in percent, which corresponds to the choice $\kappa = 10,000$ for raw returns.