Lidan Grossmass* and Ser-Huang Poon

Estimating dynamic copula dependence using intraday data

Abstract: We estimate the dynamic daily dependence between assets by applying the Semiparametric Copula-Based Multivariate Dynamic (SCOMDY) model on intraday data. Using tick data of three stock returns of the period before and during the credit crisis, we find that our dependence estimator better captures the steep increase in dependence during the onset of the crisis as compared to other commonly used time-varying copula methods. Like other high-frequency estimators, we find that the dependence estimator exhibits long memory and forecast it using a HAR model. We show that for out-of-sample forecasts, our dependence estimator performs better than the constant estimator and other commonly used time-varying copula dependence estimators.

Keywords: copula; high frequency data; intraday dependence; time-varying dependence; value-at-risk.

JEL classification: C14; C18; C58; G17.

DOI 10.1515/snde-2013-0123

1 Introduction

This paper harnesses the rich information set in intraday data to obtain a useful estimate for the daily time-varying copula dependence. We wish to obtain from intraday data the daily nonparametric conditional dependence that is measured ex-post and term this the “intraday (copula) dependence.” Our intraday dependence measure is estimated using the Semiparametric Copula-Based Multivariate Dynamic (SCOMDY) model (Chen and Fan 2006) applied to intraday data. In SCOMDY models, the individual series’ conditional mean and conditional variance is parametrically specified while their joint distribution is specified by a semi-parametric copula (parametric copula with empirical marginals). Chen and Fan (2006) noted that in such models, the limiting distribution of the copula dependence parameters are unaffected by the first-step estimation of the dynamic parameters specifying the conditional mean and variance. It is however affected by the marginal distribution of the standardized residuals. Here, we adopt Chen and Fan (2006) SCOMDY approach but apply it to intraday data. We show empirically that the derived dependence from intraday data is a superior estimator of daily dependence as compared to commonly used time-varying copulas.

Multivariate volatility models have been a long-standing challenge in financial econometrics, especially with the estimation and modeling of correlations between time series. Recent financial crises have taught us that models based on normal distribution are inadequate, and correlation changes through time and increases sharply during crises. This paper brings together two main areas of econometric literature, namely that of time-varying copulas and of econometric estimators that harness high frequency data.

Sklar (1959) introduced a joint distribution function to be flexibly decomposed into its individual marginals and a dependence function, termed the copula function, linking the marginals. By specifying different

*Corresponding author: Lidan Grossmass, Department of Economics, University of Konstanz, Box 124, 78457 Konstanz, Germany, e-mail: rachellidan@gmail.com

Ser-Huang Poon: Manchester Business School, Crawford House, University of Manchester, Manchester, UK

copula functions, different dependence structures beyond linear correlations can be formed and estimated. This makes copulas popular tools in finance when dealing with multivariate distributions. A main shortcoming is that they are static, i.e., they assume i.i.d data, and hence only consider the spatial dependence. In reality, financial time series often exhibit time-varying dependences and structural breaks [see, e.g., Patton (2001)].

It is thus important to use time-varying copula models. One of the most traditional methods of obtaining time-varying dependence is by using rolling windows of observations for estimation. However, this method is arbitrary on the choice of window size, is slow to react to market changes, and still suffers from the effects of structural breaks. Developments in this area include the regime-switching models of Rodriguez (2007) and Chollete, Heinen, and Valdesogo (2009), and the autoregressive dependence parameter of Patton (2006). The autoregressive dependence of Patton (2006) assumes that the dependence parameter follows an ARMA(1,10) process, and the determination of the forcing (or updating) variable and the choice of number of lags are key issues to deal with. To avoid specifying the parameter changing scheme, Giacomini, Hardle, and Spokoiny (2009) use local change point analysis to detect intervals of local homogeneity for estimation of the copula parameter. In a more recent paper, Creal, Koopman, and Lucas (2013) GAS model addresses the problem of choosing a forcing variable in Patton (2006) by using the lagged score which then incorporates information of the entire density. All these models face varying degrees of the same problem: when too much data from the distant past is used, the estimated parameters could be biased. However using only recent data would result in too few observations for inference. While the local change point analysis method attempts to circumvent this problem, it is only effective if sufficient post-break data is available.

Second, the availability of high-quality tick-by-tick data in financial time series has led to much innovation and research in the last decade of financial econometrics, with the development of realized volatility of Barndorff-Nielsen and Shephard (2002), realized covariances and other realized measures based on the theory of quadratic variation. The realized correlation [Barndorff-Nielsen and Shephard (2004), henceforth BNS] is derived from dividing realized covariances by the realized volatilities of the assets. We also wish to harness intraday data in estimating correlations and other non-linear dependence measures for use in copulas for two simple reasons: (1) they provide information on the most current status about the relationship between different assets of interest; (2) they provide a large quantity of observations for econometric inference.

The idea of using intraday data for copula dependence estimation has been considered before, but was focused on sampling frequency that is shorter than 1 day for use in estimating dependence across a large number of days (i.e., not within the same day). For example, Dias and Embrechts (2004) do not model the time-varying dependence parameter explicitly, they apply the method of change-point analysis on the copula dependence parameter. Breymann, Dias, and Embrechts (2003) consider FX data and find that the empirical marginals of the residuals can be rejected for ellipticity for frequencies of 8 h and shorter. In a different and more recent approach, Fengler and Okhrin (2012) makes use of Hoeffding's lemma to match the realized covariance to a copula dependence parameter using methods-of-moments and term this the "realized copula." Salvatierra and Patton (2013) extend the GAS model of Creal, Koopman, and Lucas (2013) by adding an additional regressor of realized measures in the autoregression equation.

Here we use daily returns in the SCOMDY model but with daily time-varying dependence parameter estimated using the same SCOMDY model on the intraday data within the day. Our simulation studies indicate that the dependence parameter is stable under time aggregation and is biased for non-Gaussian copulas. However despite the bias, we show that such a method gives us a useful estimate of the daily dependence. We apply the method during the crisis period between 2006 and 2008 where our dependence estimator picks up the crisis clearly and dramatically as compared to other methods. An empirical horse race also shows clear evidence that the intraday dependence estimator beats existing time-varying copula methods by providing more accurate out-of-sample value-at-risk (VaR) forecasts.

The rest of this paper is organized as follows: Section 2 discusses the class of semiparametric copula-based multivariate dynamic models, introduces the concept of intraday dependence and discusses its properties. Section 3 describes two simulation experiments that reinforces the discussions in Section 2. Section 4

begins the empirical study, where we estimate the intraday dependence measures for three stock pairs and discuss some of its estimation issues. The dependence measures are then forecasted for the 1-step ahead VaR estimation. Section 5 concludes.

2 High frequency data and semiparametric copula

Assume the returns processes y_a and y_b of assets a and b are i.i.d. Let F_{ab} be the joint density function between the returns. Sklar's theorem (Sklar 1959) states that F_{ab} can be decomposed using into the individual marginal densities and a copula,

$$F_{ab}(y_a, y_b) = C(F_a(y_a), F_b(y_b); \theta),$$

where F_a and F_b are continuous marginals and C is a copula with dependence parameter θ . The copula is thus a dependence function between the two asset returns, and various multivariate densities can be described by specifying the copula function. For risk management purposes, an important copula concept is tail dependence,

where the upper and lower tail dependence are given by $\tau_U = \lim_{v \rightarrow 1} \Pr(F_a(y_a) > v | F_b(y_b) > v) = \lim_{v \rightarrow 1} \frac{1 - 2v + C(v, v)}{1 - v}$

and $\tau_L = \lim_{v \rightarrow 0} \Pr(F_a(y_a) \leq v | F_b(y_b) \leq v) = \lim_{v \rightarrow 0} \frac{C(v, v)}{v}$ respectively. These describe the dependence in the

upper right quadrant and lower left quadrant respectively, or the probabilities that extreme events occur simultaneously. Gaussian copulas have no tail dependence and are often inadequate for modeling financial returns, where the lower tail dependence is typically observed to be larger than upper tail dependences. The Clayton and survival Gumbel copulas are more suitable in this respect.¹

Financial returns are however not i.i.d and typically exhibit inter-temporal dependence. The SCOMDY model allows for the dynamics (conditional mean and conditional variance) of the individual return series to be modeled, and the dependence of the scale-free residuals to be then specified by a copula. Furthermore, rather than pre-specifying the parametric distribution of the marginals, the empirical marginal distribution functions are used. This reduces the risk of misspecifying the marginals in the estimation of dependence but comes at a cost of decreasing efficiency of the estimation (as compared to when the parametric forms of the marginals are known).² Chen and Fan (2006) showed that in this two-step procedure, the asymptotic distribution of $\hat{\theta}$ is unaffected by the initial step of GARCH estimation.

In this paper, we estimate the time-varying daily dependence parameter from intraday data within the day by applying the same SCOMDY model on all intraday returns in day i :

$$F_{ab}^m(\eta_{a,i,j}, \eta_{b,i,j}) = C(F_a(\eta_{a,i,j}), F_b(\eta_{b,i,j}); \theta_i^m) \quad (1)$$

where θ_i^m is the intraday dependence for day i , $\eta_{a,i,j}$ and $\eta_{b,i,j}$ are the intraday residuals from a parametric model for day i and intraday period j . The superscript m denotes the sampling frequency at which the estimation takes place. The marginals F_a and F_b can be estimated by the empirical CDF $F_{n,a}$ and $F_{n,b}$ respectively.

We posit that exploiting intraday data allows us to observe the dynamics of the latent conditional dependence. This approach is much less sensitive to structural breaks in long time series as only information within the day is used. There are however some important issues to contend with. Except for the Gaussian copula, copula

¹ For detailed discussion on copulas, refer to Nelsen (1999) and Joe (1997). A brief description of the form of some commonly used copulas is given in the Web Appendix, Section E.

² See Genest, Ghoudi, and Rivest (1995) which derives the asymptotic properties of the canonical maximum likelihood (CML) estimator. The term 'canonical' is used here as it is no longer a standard maximum likelihood estimation due to the use of empirical marginals.

distributions such as Gumbel and Clayton are not closed under aggregation, i.e., the distribution of returns at daily level and at intradaily level are not the same. However, we show that by passing returns through a GARCH filter in the SCOMDY framework, the scale-free residuals retain their dependence structure with relative stability. It is thus imperative for our purpose to consider the properties of GARCH residuals. Another issue is that high-frequency returns do not capture information during non-trading hours, unlike the case when daily open-to-open or close-to-close returns are used. Due to these issues which we will discuss in the following subsection, there is biasness in the intraday dependence estimator. We find however that it remains a useful method of estimating daily dependence, and consider using an adjustment factor for the estimator. Intraday seasonality and other practical estimation issues also have to be addressed and we do this in Section 4.1.1.

2.1 The intraday dependence estimator

Assume a univariate ARMA(1,1)-GARCH(p,q) process is used to model the conditional mean and variance of intraday returns $y_{k,i,j}$, $k=1, \dots, K$, $i=1, \dots, t$, $j=1, \dots, 1/m$, for m regular time-spaced sampling intervals within a day, $0 < m \leq 1$. Let us denote $\eta_{1,i}^m, \dots, \eta_{K,i}^m$, $0 \leq i \leq t$ as the vector of GARCH residuals of K assets in day i , where t is the number of days in the sample period. For each day i and sampling frequency m , the intraday return at time j is given by

$$y_{k,i,j} = c_{k,i}^m + a_{k,i}^m y_{k,i,j-1} + \epsilon_{k,i,j} + b_{k,i}^m \epsilon_{k,i,j-1}, \quad (2)$$

$$\epsilon_{k,i,j} = \sigma_{k,i,j} \eta_{k,i,j}, \quad (3)$$

$$\sigma_{k,i,j}^2 = \omega_{k,i}^m + \sum_{q_1=1}^q \beta_{k,i,q_1}^m \sigma_{k,i,j-q_1}^2 + \sum_{p_1=1}^p \alpha_{k,i,p_1}^m y_{k,i,j-p_1}^2, \quad (4)$$

where $k=1, \dots, K$, $i=1, \dots, t$, $j=1, \dots, 1/m$ and $\omega_{k,i}^m > 0$, $\beta_{k,i,q_1}^m > 0$, $\alpha_{k,i,p_1}^m > 0$. The parameters $c_{k,i}^m$, $a_{k,i}^m$, $b_{k,i}^m$, $\omega_{k,i}^m$, β_{k,i,q_1}^m and α_{k,i,p_1}^m for $q_1=1, \dots, q$, $p_1=1, \dots, p$ are different for each asset k , each day i and are a function of m , hence the corresponding subscripts and superscript. The copula function of the intraday residuals is given by

$$F_K^m(\eta_{1,i,j}, \dots, \eta_{K,i,j}) = C_i(F_{1,i}(\eta_{1,i,j}), \dots, F_{K,i}(\eta_{K,i,j}); \theta_i^m), \quad (5)$$

where θ_i^m is the intraday dependence on day i for the K assets. The marginal distributions $F_{k,i}(\eta_{k,i,j})$ can be estimated by the empirical marginal distributions of the residuals for each day i ³

$$\hat{F}_{k,i}(x) = \frac{m}{1+m} \sum_{2 \leq j \leq 1/m+1} I(\hat{\eta}_{k,i,j} \leq x), \quad -\infty < x < \infty, \quad 0 < m \leq 1. \quad (6)$$

where $\hat{\eta}$ are the estimated standardized residuals from (3).

We now consider the constant dependence estimator, θ , for the period. The constant dependence between K assets is given by θ under a SCOMDY model,

$$F_K(\eta_{1,i}, \dots, \eta_{K,i}) = C(F_1(\eta_{1,i}), \dots, F_K(\eta_{K,i}); \theta),$$

where $\eta_{1,i}, \dots, \eta_{K,i}$ are the vector of residuals from an ARMA(1,1)-GARCH(p,q) process at a daily sampling frequency: The marginal distribution function $F_k(x)$ can be estimated in a similar fashion as (6).

³ Note that the counter for j begins from 2 here as in Berkes and Horvath (2003) as the first estimated residual tends to be biased due to the initial values assumed for the conditional mean and variance. Other papers however include the first estimated residual, i.e., j begins from 1.

The copula CML estimators $\hat{\theta}$ and $\hat{\theta}_i^m$ are given by

$$\hat{\theta} = \arg \max_{\theta} \frac{1}{t} \sum_{i=1}^t \log c(\hat{F}_1(\hat{\eta}_{1,i}), \dots, \hat{F}_K(\hat{\eta}_{K,i}); \theta),$$

and

$$\hat{\theta}_i^m = \arg \max_{\theta_i^m} \frac{m}{1+m} \sum_{j=1}^{1/m+1} \log c_i(\hat{F}_{1,i}(\hat{\eta}_{1,i,j}), \dots, \hat{F}_{K,i}(\hat{\eta}_{K,i,j}); \theta_i^m)$$

respectively, where $c(x_1, \dots, x_K; \theta)$ is the copula density function.

We have now defined our intraday dependence estimator $\hat{\theta}_i^m$. Consider the returns of two assets k_1 and k_2 and assume that dependence is non time-varying between the days $i=0$ to $i=t$ and is given by θ . What we are interested in is how $\hat{\theta}_i^m$ relates to $\hat{\theta}$. We consider two cases: (i) gaussian copulas and (ii) non-gaussian copulas. We find that under certain circumstances in the second case, $\hat{\theta}_i^m$ would be biased.

(i) Gaussian case:

For simplicity, assume that the returns follow a GARCH(1,1) process.⁴ GARCH(1,1) models are compatible with continuous time processes and can be viewed as a discretised form of continuous stochastic models.⁵ Berkes and Horvath (2003) proved that the asymptotic behavior of the empirical process of squared GARCH residuals weakly converges to a Gaussian process. This result was used in Chan et al. (2009) to establish the CML consistency using weighted approximation of the empirical residuals and to show that the limiting distribution of $\hat{\theta}_i$ is independent of the GARCH filtering.

Consider two assets k_1 and k_2 , let $v_{1,i} = \Phi^{-1}(F(\eta_{k_1,i}))$ and $v_{2,i} = \Phi^{-1}(F(\eta_{k_2,i}))$, where Φ denotes the standard normal cumulative distribution. In the Gaussian copula, the copula dependence parameter is simply the linear correlation between $v_{1,i}$ and $v_{2,i}$. Thus

$$\hat{\theta} = \frac{t^{-1} \sum_{i=1}^t v_{1,i} v_{2,i} - \bar{v}_1 \bar{v}_2}{\left[t^{-1} \sum_{i=1}^t v_{1,i}^2 - \bar{v}_1^2 \right] \left[t^{-1} \sum_{i=1}^t v_{2,i}^2 - \bar{v}_2^2 \right]} \quad (7)$$

where $\bar{v}_k = t^{-1} \sum_{i=1}^t v_{k,i}$. Using intraday data, let $v_{1,i,j} = \Phi^{-1}(F(\eta_{k_1,i,j}))$ and $v_{2,i,j} = \Phi^{-1}(F(\eta_{k_2,i,j}))$, then the Gaussian copula dependence is given by

$$\hat{\theta}_i^m = \frac{\left(1 + \frac{1}{m}\right)^{-1} \sum_{j=1}^{1+1/m} v_{1,i,j} v_{2,i,j} - \bar{v}_{1,i} \bar{v}_{2,i}}{\left[\left(1 + \frac{1}{m}\right)^{-1} \sum_{j=1}^{1+1/m} v_{1,i,j}^2 - \bar{v}_{1,i}^2 \right] \left[\left(1 + \frac{1}{m}\right)^{-1} \sum_{j=1}^{1+1/m} v_{2,i,j}^2 - \bar{v}_{2,i}^2 \right]} \quad (8)$$

where $\bar{v}_{k,i} = \left(1 + \frac{1}{m}\right)^{-1} \sum_{j=1}^{1+1/m} v_{k,i,j}$. Using Theorem C.1 (in the Web Appendix), derived under assumptions about the distribution of GARCH residuals and using the results of the empirical process of GARCH residuals

⁴ The simple GARCH model for estimating the conditional variance is commonly used to model daily returns, see e.g., Kim, Malz, and Mina (1999), Christoffersen and Diebold (2006) and Giacomini, Hardle, and Spokoiny (2009). These papers pointed out that for daily equity data, conditional means are dominated by conditional variances, hence a zero-mean assumption is reasonable. We also checked our results by including a conditional mean ARMA(1,1) model. The conditional means are small and overall results remain unchanged. Conditional means matter more for intraday data than for daily data, but are still dominated by conditional variances, as noted by Andersen and Bollerslev (1997). We consider the use of ARMA(1,1) to model conditional means but find that the assumption of a zero conditional mean is adequate. We discuss this issue further in Section 4.1.1.

⁵ Nelson (1990) showed that a sequence of discrete time GARCH(1,1) processes with i.i.d. normal innovations converges in distribution to Itô processes as the length of the discrete time intervals goes to zero.

in Berkes and Horvath (2003), $\lim_{1/m \rightarrow \infty} \sup_{0 \leq x < \infty} |\hat{F}_{k,i} - \tilde{F}_k(x)| = 0$ asymptotically. It is then straightforward to see that both $\hat{\theta}_i^m$ and $\hat{\theta}$ converges in probability to θ as $1/m \rightarrow \infty$ or $t \rightarrow \infty$.

(ii) Non-Gaussian case:

Theorem C.1 is based on the assumption that the GARCH residuals at both frequencies have a distribution with zero mean and unit variance, which implies a strong GARCH process [see Drost and Nijman (1993)]. This assumption is often made by researchers without regard to sampling frequency used. However, strong GARCH processes are not closed under aggregation and only weak GARCH processes are time-aggregating (for weak GARCH, the estimated σ_t^2 is not the conditional variance but the best linear predictor of the squared residuals, i.e., $E(y_{t+1}^2 - \sigma_t^2)y_{t-i}^r = 0$, $i \geq 0$, $r = 0, 1, 2$). This means that if the data generating process is weak GARCH(1,1) at a certain sampling frequency, then the weak GARCH(1,1) will be the data generating process for any other sampling frequency (Alexander and Lazar 2005).

In a seminal paper, Drost and Nijman (1993) defined the strong, semi-strong and weak GARCH, where strong GARCH assumes the residuals to be i.i.d. and have a distribution with zero mean and unit variance [e.g., $\sim N(0, 1)$]. Semi-strong GARCH assumes that the residuals are uncorrelated while for weak GARCH, the estimated σ_t^2 is not the conditional variance but the best linear predictor of the squared residuals, i.e., $E(y_{t+1}^2 - \sigma_t^2)y_{t-i}^r = 0$, $i \geq 0$, $r = 0, 1, 2$. Denote the parameter vector of the true GARCH process by $\gamma_{k,0} = (\omega_k, \beta_k, \alpha_k)^T$, and the parameter vector of the intradaily GARCH process by $\gamma_k^m = (\omega_k^m, \beta_k^m, \alpha_k^m)^T$. With these definitions, Drost and Nijman (1993) showed that GARCH models in the strong or semi-strong sense are not closed under aggregation and hence $\gamma_k^m \neq \gamma_{k,0}$. Only weak GARCH models are time-aggregating. This means that if the data generating process is weak GARCH(1,1) at a certain sampling frequency, then the weak GARCH(1,1) will be the data generating process for any other sampling frequency (Alexander and Lazar 2005).

Nelson (1990) showed that as sampling frequency increases, the GARCH residuals approach a limiting Student's- t distribution with $2+4\tau/\alpha^2$ degrees of freedom, where $\tau = \lim_{\Delta t \rightarrow 0} \Delta t^2(1-\beta^*-\alpha^*)$. Alexander and Lazar (2005) derive the continuous limit of the weak GARCH(1,1) as a stochastic volatility model with price-volatility correlation that is related to the skewness and kurtosis of the returns density. The limit reduces to the results of Nelson (1990) if the returns density is normal. Similarly, Drost and Werker (1996)⁶ show that the implicit assumption of an underlying continuous model implies the presence of heavy tails, and consistently find that leptokurtosis is less pronounced in aggregated series. This means that for any copula model which captures tail dependences, $\hat{\theta}_i^m$ will be larger than $\hat{\theta}$ as sampling frequency increases.

Despite the biasness in the dependence estimate, the intraday dependence estimator remains a good estimator for time-varying dependence for several reasons. First, simulations in Drost and Nijman (1992) suggest that the QML estimates are close to the true parameters even if the model is weak GARCH. Second, Nelson (1990) and Alexander and Lazar (2005) showed that as sampling frequency increases, the estimated GARCH variance will offer a good approximation for the variance of the true process, even if the GARCH model is misspecified. Both of these arguments imply that the effects of time aggregation of GARCH is limited.

Finally, in reality the sum of intraday returns is often not equivalent to daily returns, i.e., $y_{k,i} \neq \sum_{j=1}^{1/m+1} y_{k,i,j}$. This is because intraday returns do not take into account of overnight returns when trading is not 24 hours, for example, in stock markets. While there are no intraday observations in non-trading hours, stock prices do shift due to international news releases and other information made available to the investor. This means that dependence between assets in non-trading hours is not taken into account by the intraday dependence estimator and it will be downward-biased relative to $\hat{\theta}$.

There is also Epps effect to deal with at high sampling frequencies, where dependence is biased to zero as sampling intervals approaches zero. This phenomenon is due to non-synchronous trading of assets, but can usually be dealt with by using lower frequency returns (e.g., 1 or 5 min returns) rather than the highest sampling frequency. However, by using lower frequency returns, the number of observations within the single

⁶ Drost and Werker (1996) also showed that both diffusion and jump-diffusion processes can be approximated by weak GARCH models sampled at any discrete time frequency, with the kurtosis of GARCH jump-diffusion processes being larger than that of GARCH diffusion processes.

trading day becomes limited (for example, if 1 min returns are used, there are approximately 390 observation points within the day) and the dependence estimate incurs finite sampling biasness.

To deal with the multiple sources of biasness simultaneously, we consider the use of a rescaled intraday dependence [similar to that used in Martens, Chang, and Taylor (2002) and Wiggins (1991) for volatility] in our empirical section for VaR estimation (see Section 4.3). We denote the rescaling or adjustment factor by a_f and estimate it as the average of the dependence ratios $\hat{\theta}_i^m / \hat{\theta}$ in the initial in-sample period. The adjusted intraday estimator in the out-of-sample periods is then given by $\hat{\theta}_i^m / a_f$, where a_f is assumed to be constant.

3 Simulations

In this section, we study two cases to observe the effect of time aggregation on the intraday dependence estimator proposed in Section 2. In the first case, returns are simulated under a SCOMDY model. In the second case we have a more neutral and realistic setting by using a bivariate Normal Inverse Gaussian (NIG) Lévy process, where only the true linear dependence is known.

For our first simulation case, we generate 1000 random samples of size 50,000 from the Gaussian, Clayton and survival Gumbel (s. Gumbel) copulas with true copula parameters set at $\theta_{\text{gaussian}}=0.7$, $\theta_{\text{clayton}}=2.5$ and $\theta_{\text{s.gumbel}}=3.0$. To obtain the univariate returns series, we assume a zero conditional mean and GARCH(1,1) parameters $\omega_1=\omega_2=1\times 10^{-3}$, $\alpha_{1,1}=\alpha_{2,1}=0.05$ and $\beta_{1,1}=\beta_{2,1}=0.90$, where the error terms are simulated from the respective copulas above and normal marginals are assumed. The simulated returns are aggregated by $k=1, 2, 5, 10, 15, 20, 30, 60$ and 90 times to form lower frequency returns samples. The GARCH and copula dependence parameters are then re-estimated at each aggregation level.

Table 1 shows estimates of the mean and standard deviation of the pseudo maximum likelihood estimator for the Gaussian, Clayton and s. Gumbel. (The complete results with the estimated GARCH parameters are given in the Web Appendix Section C.) As discussed in Section 2, the dependence parameter θ is relatively stable for the Gaussian copula while for fat-tailed copulas, dependence declines as the number of times of aggregation increases. This is due to the decreasing kurtosis of the residuals as aggregation times increases.

For the second case, we repeat the simulation using a fat-tailed bivariate NIG Lévy process, a sub-class of the more general class of hyperbolic Lévy processes. This class of processes are closed under convolution unlike GARCH processes. The NIG Lévy process was introduced by Barndorff-Nielsen (1995) as a realistic model for log stock price returns. It exhibits non-Gaussian semi-heavy tails and allows for skewness. We hence choose to simulate the NIG process as it is a continuous model, it generates jumps which results in fat-tailed distributions, is closed under convolution and has been found to be a realistic model for stock returns. With the exception of the Gaussian case, the exact relationship between the copula dependence parameter and the NIG processes is unknown to us, but generally, a larger tail index and scale parameter in the process specification would result in greater copula dependence. The bivariate process is given by

Table 1 Estimation results for copula dependence parameters under reestimation at 1, 2, 5, 10, 15, 20, 30, 60 and 90 times of aggregation.

Copula	True	Aggregation by k times								
		1	2	5	10	15	20	30	60	90
Gaussian	0.7	0.700	0.698	0.697	0.696	0.695	0.695	0.695	0.693	0.693
		(0.157)	(0.141)	(0.084)	(0.063)	(0.084)	(0.094)	(0.084)	(0.072)	(0.089)
Clayton	2.5	2.500	1.992	1.728	1.643	1.615	1.601	1.589	1.580	1.563
		(22.0)	(13.9)	(11.6)	(10.2)	(9.93)	(10.4)	(10.2)	(10.8)	(10.7)
s. Gumbel	3.0	2.993	2.845	2.759	2.730	2.720	2.717	2.712	2.701	2.702
		(5.43)	(8.41)	(7.70)	(7.39)	(8.09)	(7.92)	(7.35)	(7.61)	(7.51)

Numbers in brackets are sample variance $\times$ no. of sample observations after aggregation.

$$X = \mu + Z\Gamma\beta + \sqrt{Z\Gamma^{1/2}}Y,$$

where $\beta \in \mathbb{R}^2$, $Z \sim IG[\delta^2, \alpha^2 - \beta^T \Gamma \beta]$, $IG[\chi, \psi]$, $\alpha, \delta, \chi, \psi > 0$. $\Gamma \in \mathbb{R}^{d \times d}$ is a positive definite structure matrix that controls the correlations between the variables. IG denotes the inverse Gaussian distribution which has the pdf

$$f(z) = \left(\frac{\chi}{2\pi z^3} \right)^{1/2} e^{\sqrt{\chi\psi}} \exp \left[-\frac{1}{2}(\chi z^{-1} + \psi z) \right], z > 0,$$

and $Y \sim N_d[0, I]$. The characteristic function of the NIG process is given by

$$\phi(u; \alpha, \beta, \delta, \mu) = \exp \{ \delta(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + iu)^2}) + i\mu u \},$$

where $u \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\delta > 0$, $0 \leq |\beta| \leq \alpha$, and the $\alpha, \beta, \delta, \mu$ parameters can be interpreted as the tail index, symmetry, scale and location parameters, respectively (Øigard et al. 2005). Table 2 gives the results of the estimated dependence under aggregation of the simulated NIG process.⁷ The Gaussian copula dependence is accurate at approximately 0.5 throughout the aggregations. While we do not have a true theoretical value of θ for the survival Gumbel and Clayton copulas, we observe that there is a greater degree of stability in the NIG simulations as compared to the GARCH simulations. There is still a downward bias for survival Gumbel and a relatively greater downward bias for Clayton copula dependence as the number of times of aggregation increases.

The two simulation cases show that for copulas with tail dependence, there is an upward bias in the estimation of θ_i as sampling frequency increases (and sampling intervals get smaller). This is in line with our discussion in Section 2.1. Overall, the results are promising in that despite the bias, the intraday dependence estimator is informative about the true dependence of the data generating process.

Table 2 Estimation results for Gaussian, Clayton and s. Gumbel copula dependence when a heavy-tailed bivariate NIG process is simulated.

Copula	True	Aggregation by k times								
		1	2	5	10	15	20	30	60	90
Gaussian	0.50	0.493	0.497	0.500	0.500	0.500	0.500	0.500	0.499	0.498
		(37.6)	(6.92)	(0.664)	(0.579)	(0.551)	(0.570)	(0.601)	(0.578)	(0.555)
Clayton	–	0.727	0.706	0.685	0.676	0.674	0.672	0.675	0.677	0.683
		(43.5)	(22.3)	(10.5)	(6.86)	(5.67)	(5.85)	(4.43)	(3.68)	(3.65)
s. Gumbel	–	1.472	1.457	1.444	1.438	1.436	1.436	1.436	1.439	1.440
		(23.8)	(6.73)	(1.48)	(1.40)	(1.30)	(1.31)	(1.40)	(1.35)	(1.30)

Numbers in brackets are sample variance $\times$ no. of sample observations after aggregation.

4 Empirical study

In this section, we estimate VaR, an often-used risk measure in financial institutions and provide empirical evidence that the intraday dependence estimator provides more accurate out-of-sample VaR forecasts

⁷ The following parameters are used: $\alpha = \begin{pmatrix} 1.892 \\ 1.892 \end{pmatrix}$, $\beta = \begin{pmatrix} 0.0092 \\ -0.0263 \end{pmatrix}$, $\delta = \begin{pmatrix} 0.9906 & 0.5 \\ 0.5 & 1.0098 \end{pmatrix}$ and $\mu = \begin{pmatrix} -0.0045 \\ -0.0043 \end{pmatrix}$. The resulting simulated process has an empirical mean = $\begin{pmatrix} -0.0059 \\ -0.0120 \end{pmatrix}$, covariance = $\begin{pmatrix} 0.3473 & 0.1753 \\ 0.1753 & 0.3541 \end{pmatrix}$, skewness = $\begin{pmatrix} -0.0071 \\ -0.0386 \end{pmatrix}$, and excess kurtosis = $\begin{pmatrix} 2.7598 \\ 2.7617 \end{pmatrix}$. These values are close to those obtained empirically for our dataset.

than the constant, rolling window and autoregressive dependence estimators. We should emphasize that our focus is not on copula choice, but on comparing the performance of our time-varying dependence estimator with other commonly-used time-varying dependence estimators.

4.1 Data

For our empirical study, we use the NYSE TAQ data of three stocks – Citigroup (c), JP Morgan (jpm) and IBM (ibm) between 2 January 2006 and 31 December 2008. The data is cleaned using the algorithm described in Barndorff-Nielsen et al. (2008), and we include only observations for each trading day from 9.30 to 16.00. We use quote data for our study as it contains a larger number of observations as compared to trade data, thus reducing the effect of asynchronous trading on the estimations. Finally, we use the calendar time sampling and the last-tick interpolation method to obtain regularly spaced return observations. Figure 1 plots the daily and 1-min intraday log returns time series of the three stocks. We can see that for both daily and intraday data, the volatility of returns increases sharply from mid-2007 onwards due to the subprime crisis.

4.1.1 Intraday seasonality adjustment and other issues

An important characteristic of high frequency data is the presence of intraday seasonality. Wood, McInish, and Ord (1985) and Harris (1986) were the first to note this periodic U-shape pattern in return volatility during the trading day while Andersen and Bollerslev (1997) show that it is important to take into account intraday seasonality before applying standard volatility models to intraday data. They propose the Flexible Fourier Form (FFF) for estimating the seasonality pattern. Martens, Chang, and Taylor (2002) show that FFF adjusted GARCH(1,1) forecasts intraday volatility better than other seasonality correction methods. We hence adopt FFF for seasonality adjustment.

To illustrate the resulting intraday dependence estimate for both with and without seasonality adjustment, Figure 2 compares the BNS realized correlation⁸ and the Gaussian copula intraday dependence of Citigroup and JP Morgan returns in the period 2006–2008 using 1-min sampling frequency. The dotted line is when the FFF seasonality adjustment is applied while the dash-dot line is when no seasonality adjustment is made. The graph shows the intraday Gaussian dependence and the realized correlation to be closely matched. This means that dependence is preserved after the univariate ARMA-GARCH filtration, as was observed by Dias and Embrechts (2004).

Furthermore, we find that the inclusion of seasonal adjustment gives a dependence parameter that is very close to that obtained without seasonal adjustment. We also find that assuming a zero conditional mean gives a dependence parameter that is close to that when the conditional mean is modeled by an ARMA process (which is used in Figure 2).

We use FFF to adjust for intraday seasonality of the returns before passing them through a conditional mean and volatility model.⁹ We find, however, that the seasonality correction removes the extreme observations of each day (usually at the start and end of day), resulting in an intraday residual empirical distribution that has thinner tails. Figure 3 shows the scatter plots of the intraday residuals for the c-jpm stock pair for three different days: 3 January 2006, 1 August 2007 and 13 October 2008. The residual plots on the left side do not have deseasonalization adjustment while those on the right side are when the returns are first deseasonalized.

⁸ The realized covariance matrix that was used to compute realized correlation has been subsampled to reduce the effects of market microstructure noise.

⁹ For FFF as described in Andersen and Bollerslev (1997), we use $P=2$ and $J=0$ for c and jpm and $P=4$ and $J=0$ for ibm, and find that this provides an adequate fit. Figure 1 in the Web Appendix shows that the estimated average intraday periodic patterns over the sample period using FFF provide a close approximation to the intraday absolute returns.

Figure 1 Return series of c, jpm and ibm (from top to bottom) at daily (left) and 1-min (right) frequencies.

A first observation is that the dependence of the stock pair is clearly time-varying. On 3rd January 2006, the scatter plots are relatively elliptical with little tail dependence. On 1 August 2007, which is close to the start of the credit crisis, the dependence of the stock-pair has increased sharply. 13 October 2008, about the middle of the credit crisis, shows a reduction in linear dependence as compared to 1 August 2007, but still exhibits lower tail dependence. These dependence variations over time would have not been captured if we had used daily data though the period (see Figure 4).

Figure 2 Plot of the Citigroup-JP Morgan realized correlation (blue dotted line) and the intraday dependence of the Gaussian copula using 1-min sampling frequency without pre-applying seasonality adjustment (black solid line with *) and when FFF seasonality adjustment is applied (red solid line).

The second observation is that the shapes of the scatter plots are similar when deseasonalization is applied (plots on right) and when it is not (plots on left), hence resulting in the dependence parameter estimates for deseasonalized and non-deseasonalized returns being rather close, as noted in Figure 2. However, the density of tail observation is slightly thinner after deseasonalization as it removes the effect of the beginning and end of day returns, which tend to exhibit the largest intraday volatilities. Since these extreme tail observations are important for copula tail estimation, we decide to not apply deseasonalization to the returns before passing through the ARMA-GARCH filter.

Andersen and Bollerslev (1997), Martens, Chang, and Taylor (2002) and other papers apply the ARMA-GARCH filtration to high frequency returns across multiple days and find a large bias in the estimated model parameters induced by the periodicity effect when seasonality is not first adjusted for. In this paper, however, the conditional mean and variance model is applied only to a single day's intraday return and reestimated everyday. We find that this provides some form of adjustment for the intraday seasonality pattern. Figure 5 shows the average intraday conditional variance of the stocks *c*, *jpm* and *ibm* captured by GARCH(1,1).¹⁰ They exhibit the signature lopsided U-shape where volatility is the highest at the beginning of the trading day, lowest at midday, and picks up towards trading day close. Hence standardising returns using a GARCH(1,1) is capable of removing the intradaily seasonality effect. An issue that arises with such an approach is that the GARCH variance looks non-stationary. However, we checked the estimated GARCH parameters and find them to fulfil the stationarity conditions.

We next turn to the issue of using ARMA to model the conditional mean. As with many earlier research papers, we find that the inclusion of the conditional mean has a negligible effect as compared to the conditional variance. The magnitude of the conditional mean of intraday data is very small but matter more than in the case of daily data. Andersen and Bollerslev (1997) find that allowing for an MA term in the conditional mean can account for the economically minor first order autocorrelation in returns, and the MA term is negative at sampling frequencies of less than 5 min. The negative MA term is sometimes explained by the effect

¹⁰ Similar graphs for *jpm* and *ibm* are found in the Web Appendix Section A.

Figure 3 Scatter plots of 1-min residuals for c-jpm returns pair for days 3 January 2006 (top), 1 August 2007 (middle) and 13 October 2008 (bottom), respectively. Plots of left side use seasonality-unadjusted returns while plots on right side use deseasonalized returns.

Figure 4 Daily residual scatter plot of stock pair c-jpm for the period between 3 January 2006 and 31 December 2008.

Figure 5 Average intraday variance of c using GARCH (1,1) effectively captures the intradaily seasonality effects. Graphs for other stocks are given in the Web Appendix.

of bid-ask bounce in high frequency data due to dealers trading around the spread producing a slight mean reversion in returns.¹¹

We model the conditional mean for intraday returns using an ARMA process and find the estimated dependence measures to be very close to the case where the conditional mean is assumed to be zero. The residual distributions are similar but slightly more centred when the conditional mean is filtered out. Since our aim is not to forecast intraday returns, and the Ljung-Box statistics of the residuals do not detect significant autocorrelations after a GARCH filtration, we do not filter out the conditional mean for intraday returns.¹²

¹¹ However, for the highly liquid stocks that we are using, the effect of the bid-ask bounce is small, since we are not considering tick-by-tick returns.

¹² We also recognize that EGARCH or similar models that account for leverage effects may be more appropriate for daily equity data. We find, however, by plots of the squared residuals against the lagged residuals, that the leverage effect is not prominent in intraday data.

4.1.2 Descriptive statistics

Descriptive statistics of the return series and GARCH residuals at daily and 1-min sampling frequencies are given in Table 3. The descriptive statistics indicate that intraday returns tend to have thinner tails than daily returns. However, the GARCH residuals display the reverse characteristic [as found by Nelson (1990) and Drost and Werker (1996)]. Finally, the Ljung-Box statistics indicate that there is no significant autocorrelation in the residuals at both daily and intradaily frequencies.

4.1.3 Effect of sampling frequency

To investigate the effects of sampling frequency on the intraday dependence estimator when empirical data is used, we estimate intraday dependence of the stock pairs c-jpm, c-ibm and jpm-ibm at sampling frequencies ranging from 1 to 30 min. Figure 6 shows the average intraday dependence at the different sampling frequencies for year 2006 (left) and year 2008 (right) for the stock-pairs using the Gaussian, survival Gumbel and Clayton copulas. We use these three copulas since they are commonly used in empirical applications.

The graphs show that the parameters are fairly stable at different sampling frequencies but show a consistent slight upward trend as the sampling frequency decreases. (Note that the Gaussian dependence parameter is also increasing slightly but this increase is less obvious in the graphs due to scale effect.) While Epps effects may be present at high sampling frequencies, we postulate that this upward bias is mainly due to the

Table 3 Descriptive statistics of daily returns and univariate GARCH residuals (top) and average descriptive statistics at 1 min sampling frequency (bottom).

	c		jpm		ibm	
	Returns	Residuals	Returns	Residuals	Returns	Residuals
Daily frequency						
Mean	−0.001	−0.059	0.000	−0.002	0.000	0.026
Median	0.000	−0.067	0.000	−0.003	0.001	0.050
Std. Dev.	0.021	1.009	0.024	1.007	0.015	0.999
Skew	0.035	0.004	−0.028	0.147	−0.141	−0.156
Kurtosis	9.853	4.674	9.866	4.662	6.657	4.220
Min	−0.121	−3.914	−0.143	−3.825	−0.057	−3.904
Max	0.116	4.315	0.119	4.198	0.087	4.614
LB (10)	35.44 (0.00)	10.72 (0.38)	16.10 (0.10)	14.77 (0.14)	28.65 (0.00)	16.06 (0.10)
LB (25)	55.70 (0.00)	21.11 (0.69)	64.14 (0.00)	26.92 (0.36)	61.79 (0.00)	33.00 (0.13)
1 Min frequency						
Mean	0.000	−0.007	0.000	0.002	0.000	0.006
Median	0.000	−0.004	0.000	−0.002	0.000	0.001
Std. Dev.	0.001	1.054	0.001	1.035	0.001	1.047
Skew	0.056	0.042	0.031	0.005	0.077	−0.013
Kurtosis	6.729	6.596	6.969	6.007	7.986	6.639
Min	−0.005	−4.367	−0.005	−4.086	−0.003	−4.318
Max	0.005	4.420	0.005	4.098	0.003	4.152
LB (10)	16.08 (0.27)	10.74 (0.47)	16.41 (0.27)	10.69 (0.47)	18.06 (0.22)	11.51 (0.42)
LB (25)	33.90 (0.29)	24.82 (0.52)	35.21 (0.26)	25.72 (0.48)	36.25 (0.24)	26.38 (0.46)

Numbers in brackets indicate the p-value of the Ljung-Box statistic with the null hypothesis of no autocorrelation. ‘c’, ‘jpm’ and ‘ibm’ are short for Citigroup, JP Morgan and IBM, respectively.

Figure 6 Intraday dependence signature plots of c-jpm (top), c-ibm (middle) and jpm-ibm (bottom) for year 2006 and 2008 using the Gaussian (solid lines), Clayton (dashed lines) and s. Gumbel (x-lines) copulas.

effect of finite sampling.¹³ At 1-min sampling frequency, there are 390 observations for estimation while at 30 min sampling frequency, there are only 12 observations for the estimation of the dependence parameter. We now see that there are two potential effects that overlap when we use actual empirical data: at too low sampling frequencies we have fewer data points and face a positive bias due to finite sampling, while at very high sampling frequencies, we face a positive bias due to the distribution of the residuals tending to fatter tails, as described in Nelson (1990). As a compromise of these two effects in the data, we use the 1-min sampling frequency. Naturally an interesting research question would be the determination of an optimal sampling frequency but we do not delve further into this issue here.

4.1.4 Copula choice

Given that there are hundreds of copula specifications, it is impossible to consider all of them. To reduce the estimation load, we estimate seven different copulas [Gaussian, Clayton, survival Gumbel, Plackett, Frank,

¹³ See simulation results in Section C.1 in the Web Appendix.

Student's- t and symmetrized Joe-Clayton (SJC)] using the daily GARCH residuals of the three stock pairs (c-jpm, c-ibm, jpm-ibm) and pick the three copulas that give the three largest average log likelihoods. The Gaussian, Plackett and Frank copulas do not exhibit tail dependence, while the Clayton and survival Gumbel copulas have lower tail dependence. The Student's- t and SJC copulas have both upper and lower tail dependence. Table 4 gives the average log likelihood estimates for the three stock pairs. Finally, we always include the Gaussian copula as a basis of comparison, which gives four copulas for each stock pair.

Our treatment here on the issue of copula choice is rather simplistic as it is not the focus of this paper. This issue is however important and non-trivial, with an ongoing strand of research on copula selection and goodness-of-fit testing (see, e.g., Diks, Panchenko, and van Dijk (2010) and Diks et al. (2014) that optimizes copula choice using out-of-sample forecasting criterion) and we refer the reader to these related literature.

4.2 Comparing the intraday dependence estimator with time-varying dependence estimators

We compare the intraday dependence with the constant unconditional dependence estimator and two other time-varying dependence estimators: the rolling window estimator and the autoregressive estimator. The constant unconditional dependence estimator estimates the copula parameter over the whole time series and one constant estimate over the whole period is thus obtained. The rolling window estimator estimates the copula parameter over a rolling window of 250 observations (approximately 1 year). After the first 250 observations, the window incorporates one new observation and drops the oldest observation at each incremental time point. The autoregressive estimator is adapted from that of Patton (2006) and estimated using maximum likelihood. For the specification of the Gaussian and SJC time varying copulas, we use the specification of ARMA(1,10) as in Patton (2006). As noted in the paper, a key difficulty is in identifying the forcing (or updating) variable. For the other time varying copula models, we use the mean absolute difference between the GARCH residuals u_t and v_t as forcing variable with evolution equation for θ given by

$$\theta_t = \Lambda_2 \left(\beta_0 + \beta_1 \cdot \theta_{t-1} + \beta_2 \cdot \frac{1}{10} \sum_{j=1}^{10} |u_{t-j} - v_{t-j}| \right), \quad (9)$$

where $\Lambda_2(x) = x^2 + 1$ which keeps the parameter above 1 at all times. This is suitable for the Clayton, Gumbel, Student's- t 's v parameter where the parameter must lie above 1.¹⁴ For the time-varying Plackett copula, the same forcing variable was used and $\Lambda_2(x) = x$. Since for the Plackett copula, $\theta > 0$ and $\theta \neq 0$, penalties to the likelihood were imposed to deal with these restrictions.

Figure 7 shows the estimated dependence measures of c-jpm using different copulas with the constant estimator (solid horizontal lines), intraday dependence estimator (dotted lines), rolling window estimator (dashed lines) and autoregressive estimator (solid lines). Overall, the graphs show that dependence between stocks is dynamic, and that the time-varying copula parameters tend to trend together. This emphasizes the phenomenon of time-varying dependence. In general, the autoregressive estimator hovers around the unconditional constant estimator while the rolling window estimator displays a smooth trend. The intraday

Table 4 Average log likelihood of different copulas over intraday dependence estimated for the period 2006–2008.

	Gaussian	Clayton	s. Gumbel	Plackett	Frank	Student's- t	SJC
c-jpm	0.4335	0.3762	0.4446	0.4414	0.4064	0.4723	0.4599
c-ibm	0.1336	0.1111	0.1335	0.1396	0.1357	0.1430	0.1359
jpm-ibm	0.1226	0.1166	0.1392	0.1463	0.1355	0.1548	0.1385

Bold font indicates it is one of the best three log likelihood estimates and is selected for further estimations.

¹⁴ Clayton and survival Gumbel copulas do not allow for negative dependence and have a parameter boundary value for θ of 1 and 2, respectively, while the degrees of freedom for a Student's- t copula should lie above 2 if we assume finite variance.

Figure 7 Constant and time varying dependence estimators Co (black solid lines), RW (black dashed-dotted lines), AR (blue solid lines) and ID (red dotted lines) for c-jpm between 2006 and 2008. Top row: Gaussian and s. Gumbel copulas. Left middle and bottom rows: Student's- t copula (ρ and ν); right middle and bottom rows: SJC copula (upper and lower tail parameters).

dependence estimates tends to be lower than other estimators during the calm periods, but becomes greater than other estimators during the subprime crisis period. The intraday dependence estimator exhibits steep increases from the second quarter of 2007 to about last quarter of 2008 (coinciding with the subprime crisis). A large drop in dependence is observed in November 2008 when Citigroup experienced a one week crash in its stock prices. The rolling window estimator does effectively capture the increasing dependence but exhibits a lag. The autoregressive estimator moves much in tandem with the intraday dependence measure in capturing short-term trends but on a long term basis moves largely about the constant unconditional estimator and does not capture the increasing dependence over the period. Finally, the intraday dependence estimator tends to lie below the other estimators but increases dramatically during the crisis period. This illustrates how it effectively captures the increase in risk quickly during periods when stocks are exhibiting large

negative co-movements. Similar observations can be made for c-ibm and jpm-ibm stock pairs (see Figures 10 and 11 in the Appendix).¹⁵

From Figure 7, we can also see that the use of an asymmetric copula, the SJC copula, allows us to identify that the lower tail dependence increases more steeply than that of the upper tail dependence during the crisis. There are more sharp peaks in the lower tail dependence during the crisis. A further comment is that the degrees of freedom parameter of the Student's- t copula seem to jump a lot between 1 and above 10 (see bottom left of Figure 7). At above 10 degrees of freedom, the Student's- t copula approaches that of a normal distribution and we see that such occurrences are more common in 2006 and the first half of 2007. Due to the large scale between 1 and 10, the degree of freedom plots appears very jumpy, but illustrates varying degree of fat-tailedness that changes each day (realized measures are in general rather volatile).

4.3 Using intraday dependence for VaR estimation

We now consider using the intraday dependence estimator for VaR estimation. We use data from 2 January 2006 to 31 December 2006 for estimation and forecast the period between 2 January 2007 to 31 December 2008. The out-of-sample forecast period is particularly challenging given that it consists mainly of the credit crisis, whereas the initial in-sample period consists only of the calm market period. However we wish to observe how the intraday dependence performs in relation to other time varying dependence measures in times of extreme market fluctuations. Since VaR is the next period's forecasted loss with a certain (extreme) probability, its estimation using copulas requires a forecast of the next day's dependence.

4.3.1 Forecasting intraday dependence

Figure 8 shows the sample autocorrelations (SACFs) of the intraday dependences in a Gaussian, s. Gumbel and SJC copula for the c-jpm stock pair.¹⁶ It shows the slow decay of the dependence measures, a property also exhibited by realized volatility. The strong persistence is suggestive of long memory behavior of the intraday dependence. Long memory processes can be modeled parametrically by fractionally integrated processes

Figure 8 Sample autocorrelations of the intraday dependences for the c-jpm stock pair using the Gaussian (solid line), s. Gumbel (dash-dot line) and SJC (line with dot) copulas.

¹⁵ For the Student's- t copula in c-ibm stock pair, convergence for the autoregressive estimator could not be obtained for the ν (degrees of freedom) parameter. This illustrates a shortcoming of the autoregressive estimator in that such a parametric form of the dynamics is suitable for only certain copula dependences.

¹⁶ The Student's- t copula is not included here as the Student's- t correlation is similar to that of the Gaussian copula dependence.

such as the ARFIMA model or by the heteroscedastic autoregressive (HAR) model introduced by Corsi (2009) and Corsi et al. (2008).

We use a HAR model for forecasting intraday dependence due to its simplicity in estimation and good performance in modeling realized volatility and correlations [see for example Corsi (2009), Audrino and Corsi (2010), Vortelinos (2010)]. The form is a stochastic additive cascade of three different components corresponding to three different time horizons:

$$\hat{\theta}_{t+1}^{(d)} = a + b^{(d)}\theta_t^{(d)} + b^{(w)}\theta_t^{(w)} + b^{(m)}\theta_t^{(m)}, \quad (10)$$

where $\theta_t^{(d)}$, $\theta_t^{(w)}$ and $\theta_t^{(m)}$ are the daily, weekly and monthly dependences. For our purpose, the weekly and monthly dependences are computed by taking averages of intraday dependences of the past 5 and 20 days respectively. Estimation of (10) is simply made via OLS.

4.3.2 VaR estimation and evaluation

We assume an equally-weighted portfolio and estimate the VaR of the portfolio using the constant parameter estimator (CO), rolling-window estimator (RW), autoregressive estimator (AR), intraday dependence (ID) estimator and the adjusted intraday dependence (ID*) estimator. As described in Section 2, we assume that the adjustment or rescaling factor a_f remains constant throughout the out-of-sample period. We estimate $ID_t^* = ID_t / \bar{a}_f$, where $\bar{a}_f$ is the mean of $a_{f,t} = ID_t / CO_t$ estimated in the initial in-sample period, i.e., January–December 2007.

For the constant dependence estimator (CO), we estimate the copula dependence of the initial in-sample period and keep it constant over the whole forecast period. For the rolling-window dependence estimator (RW), we use the last 250 days to estimate the dependence and assume this is the next period's best forecast. For the autoregressive estimator (AR), we use the last 250 days to update the parameters of the model and forecast the next period's dependence. Finally, we use the HAR model (described in previous subsection) to forecast both the intraday dependence estimator (ID) and adjusted intraday dependence estimator (ID*).

We simulate 50,000 observations for each day t using the forecasted dependence to obtain the joint cumulative distribution of uniform marginals $\{u_1, u_2\}$. The residuals are obtained by $\eta_{k,t} = F_N^{-1}(u_k)$, where the empirical probability distribution of the last 250 days is used to invert the simulated cdf by linear interpolation. The returns of each asset, $X_{k,t}$ are obtained by $X_{k,t} = \sigma_{k,t} \eta_{k,t}$, where $\sigma_{k,t}$ is forecasted using univariate GARCH(1,1) parameters estimated in-sample. The returns of portfolio, P_t are then computed by

$$P_t = \sum_{k=1}^K w_k X_{k,t},$$

where w_k are the fixed portfolio weights and $\sum_{k=1}^K w_k = 1$ (for our purpose, we use $K=2$, $w_1=w_2=0.5$). VaR is obtained by taking the empirical quantiles. We estimate the 1, 5 and 10% VaR (also referred to in practice as the 99%, 95% and 90% VaR), but also include the 0.5, 2, 3, 4, 7% quantiles to have a more exact characterization of the lower tail distribution.

For VaR evaluation, we compute what is popularly known as “Hit Rate” which is simply the percentage of exceedances (% ex). This number should preferably be as close to the quantile of interest as possible. The Kupiec Test (Ku) is a popular VaR backtest used to check for VaR adequacy and it considers the unconditional coverage, i.e., the number of violations. The test statistic takes the form

$$Ku = 2 \log \left(\left(\frac{1-\hat{p}}{1-p} \right)^{N - \sum \mathbb{I}_{y_t < \widehat{VaR}_t^p}} \left(\frac{\hat{p}}{p} \right)^{\sum \mathbb{I}_{y_t < \widehat{VaR}_t^p}} \right),$$

where N is sample size, p is the quantile of interest, $\hat{p} = \frac{1}{N} \sum \mathbb{I}_{y_t < \widehat{VaR}_t^p}$, y_t is daily return and $\mathbb{I}$ is the indicator function. The statistic is asymptotically χ^2 -distributed with 1 degrees of freedom with the null hypothesis that the VaR model is adequate. A rejection of the statistic hence indicates inadequate coverage.

We also use the Dynamic Quantile (DQ) test introduced by Engle and Manganelli (2004) to check if the VaR model is prone to “violation clustering,” an undesirable property for VaR models. An ideal VaR model should not only fulfill the regulatory Hit Rate, but also have Hits (H_t) that are uncorrelated with the variables in the information set $\mathcal{F}_{t-1}$. To conduct the test, we regress the variable, $\tilde{H}_t = \mathbb{I}_{y_t < \widehat{\text{VaR}}_t^p} - p$ against past lagged $\tilde{H}_t$ and the contemporaneous VaR forecast,

$$\tilde{H}_t = \lambda_0 + \sum_{i=1}^h \lambda_i \tilde{H}_{t-i} + \lambda_{h+1} \widehat{\text{VaR}}_t^p + \mu_t = X\lambda + \mu_t. \quad (11)$$

The DQ test statistic is then given by $\frac{\hat{\lambda}' X' X \hat{\lambda}}{p(1-p)} \sim \chi^2(h+2)$, where $\hat{\lambda}$ is estimated using OLS. We use $h=4$ lags as in Engle and Manganelli (2004).

To check for VaR accuracy, we use the tick loss function $TL_{t,p}$ of Komunjer (2005),

$$TL_{t,p} = (p - \mathbb{I}_{y_t < \widehat{\text{VaR}}_t^p})(y_t - \widehat{\text{VaR}}_t^p). \quad (12)$$

This loss function considers the distance of the estimated VaR from the true quantile. To compare the different estimators M , we use the Diebold-Mariano test with loss function (12) and adopt the constant estimator as the base model (M_0). The test statistic for equal predictive accuracy is given by

$$DM_{TL} = \frac{\bar{d}}{(\hat{LRV}/N)^{1/2}},$$

where $\bar{d} = \frac{1}{N} \sum d_t$, $d_t = TL_{t,p,M} - TL_{t,p,M_0}$ and $\hat{LRV} = \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j$, $\gamma_j = \text{cov}(d_t, d_{t-j})$ is the long run variance that takes into account of serial correlation. The statistic DM_{TL} is asymptotically $N(0,1)$ distributed. A significant value implies the rejection of the null of equal predictive accuracy, and a negative value indicates that model M has better predictive accuracy than base model M_0 .

Tables 5–7 tabulate the VaR evaluation for the three stock pairs. As the estimation of the autoregressive estimator for the degrees of freedom, ν , in the Student- t copula faces convergence issues, we leave out the autoregressive estimator for the Student's- t copulas in our results. The p-values for the Kupiec Test (Ku) and Dynamic Quantile test (DQ) are reported in the tables. Bold numbers indicate a rejection of the test at 1% significance level. A rejection of the Ku means that the VaR model produces a hit rate that is either too high or too low, while a rejection of the DQ means that the VaR model either produces a hit rate that is unacceptable or that the hits cannot be rejected for no autocorrelation.

Unsurprisingly, the CO estimator performs poorly since it is estimated only using the in-sample calm period data and the out-of-sample forecast period is an extremely turbulent period. In almost all cases, it fails the Ku and DQ tests. A plot of the VaR estimates show that the CO estimator underestimates VaR in the second half of 2007 and for all of 2008. The RW estimator improves the VaR performance by constantly updating dependence estimates and it has an acceptable hit rate at the more extreme quantiles (i.e., 0.5 and 1%). From the DM_{TL} statistic, it is significantly more accurate than the CO estimator at almost all p . The AR estimator gives a similar performance to the RW estimator.

From the DM_{TL} statistics, the ID estimator gives the most accurate VaR estimates, but produces too many exceedances as compared to the RW and AR estimators. The use of ID* estimator corrects this problem, at the cost of slightly reduced accuracy as compared to its non-adjusted counterpart. ID* consistently gives the exceedances that are closest to the desired hit rate at all quantile levels. With the exception of the 7 and 10% quantile using SJC copula for the c-jpm pair, the accuracy of the adjusted intraday dependence estimator is comparable to the RW and AR estimators. At the 10% quantile for c-jpm using the SJC copula, the VaR estimate using ID* is significantly less accurate than using the CO estimator. This is likely due to the fact that the ID estimate is rather high for c-jpm and after correction, the ID* estimate often hits the upper bound in the dependence of the SJC copula. This problem is not faced for the same copula when other stock pairs with lower dependence are used.

We plot the 10% quantile using SJC copula for the c-jpm stock pair in Figure 9 (see bottom half of Figure 9), and observe that in this case, the ID* gives a much higher (more negative) VaR than all the other estimators. While it is less accurate than the CO estimator (that severely underestimates VaR), it is the only estimator that

Table 5 VaR evaluation for stock-pair c-jpm using the constant (CO), rolling window (RW), autoregressive estimator (AR) and the intraday dependence (ID) estimators for the Gaussian, s. Gumbel, Student's-t and SJC copulas.

p		0.005	0.01	0.02	0.03	0.04	0.05	0.07	0.10	0.005	0.01	0.02	0.03	0.04	0.05	0.07	0.10
		Gaussian															
Co	ex	0.026	0.056	0.095	0.113	0.121	0.123	0.155	0.197	0.016	0.052	0.087	0.107	0.121	0.123	0.155	0.195
	Ku	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	DQ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TL	0.012	0.020	0.048	0.056	0.066	0.085	0.107	0.137	0.008	0.016	0.036	0.050	0.062	0.076	0.105	0.137
RW	ex	0.062	0.050	0.000	0.002	0.007	0.001	0.000	0.008	0.388	0.220	0.023	0.018	0.021	0.014	0.004	0.008
	Ku	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.221	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	DQ	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.221	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TL	-1.474	-2.305	-2.670	-2.610	-2.603	-2.410	-2.225	-1.752	-0.966	-1.987	-2.331	-2.385	-2.466	-2.343	-2.199	-1.671
AR	ex	0.010	0.020	0.040	0.056	0.068	0.082	0.101	0.135	0.008	0.016	0.032	0.050	0.062	0.076	0.101	0.139
	Ku	0.167	0.050	0.005	0.002	0.004	0.003	0.009	0.012	0.388	0.220	0.081	0.018	0.021	0.014	0.009	0.005
	DQ	0.081	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.220	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TL	-1.507	-2.307	-2.698	-2.677	-2.635	-2.464	-2.287	-1.770	-0.956	-1.999	-2.346	-2.405	-2.468	-2.341	-2.190	-1.676
ID	ex	0.016	0.024	0.052	0.064	0.078	0.091	0.119	0.145	0.008	0.020	0.044	0.062	0.070	0.091	0.115	0.147
	Ku	0.006	0.008	0.000	0.000	0.000	0.000	0.000	0.001	0.388	0.050	0.001	0.000	0.002	0.000	0.000	0.001
	DQ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.224	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TL	-1.444	-2.435	-2.798	-2.764	-2.775	-2.535	-2.333	-1.878	-1.026	-2.081	-2.391	-2.534	-2.593	-2.409	-2.249	-1.801
ID*	ex	0.008	0.020	0.034	0.052	0.062	0.074	0.095	0.131	0.008	0.016	0.032	0.048	0.062	0.074	0.097	0.137
	Ku	0.388	0.050	0.044	0.010	0.021	0.023	0.033	0.025	0.388	0.220	0.081	0.032	0.021	0.023	0.022	0.008
	DQ	0.215	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.226	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TL	-1.256	-2.135	-2.544	-2.348	-2.221	-2.261	-2.202	-1.612	-1.003	-2.012	-2.360	-2.425	-2.482	-2.371	-2.220	-1.709
		Student's-t															
Co	ex	0.026	0.052	0.093	0.113	0.121	0.125	0.155	0.197	0.020	0.052	0.093	0.113	0.121	0.135	0.175	0.203
	Ku	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	DQ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TL	0.008	0.018	0.044	0.054	0.066	0.085	0.107	0.139	0.008	0.016	0.038	0.052	0.062	0.080	0.105	0.141
RW	ex	0.388	0.109	0.001	0.005	0.007	0.001	0.002	0.005	0.388	0.220	0.011	0.010	0.021	0.005	0.004	0.004
	Ku	0.229	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.218	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	DQ	0.229	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.218	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TL	-1.396	-2.226	-2.515	-2.569	-2.596	-2.446	-2.230	-1.739	-1.154	-2.089	-2.474	-2.509	-2.582	-2.501	-2.347	-1.977
AR	ex	-	-	-	-	-	-	-	-	0.008	0.018	0.038	0.048	0.064	0.080	0.103	0.139
	Ku	-	-	-	-	-	-	-	-	0.388	0.109	0.011	0.032	0.013	0.005	0.006	0.005
	DQ	-	-	-	-	-	-	-	-	0.222	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TL	-	-	-	-	-	-	-	-	-1.175	-2.092	-2.498	-2.539	-2.605	-2.511	-2.321	-1.952
ID	ex	0.012	0.022	0.052	0.064	0.078	0.091	0.119	0.147	0.012	0.020	0.048	0.064	0.076	0.091	0.121	0.145
	Ku	0.062	0.021	0.000	0.000	0.000	0.000	0.000	0.001	0.062	0.050	0.000	0.000	0.000	0.000	0.000	0.001
	DQ	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	TL	-1.432	-2.325	-2.598	-2.723	-2.780	-2.520	-2.288	-1.833	-1.237	-2.210	-2.573	-2.733	-2.852	-2.676	-2.499	-2.210
ID*	ex	0.008	0.016	0.032	0.048	0.062	0.072	0.095	0.133	0.006	0.016	0.028	0.036	0.048	0.052	0.058	0.078
	Ku	0.388	0.220	0.081	0.032	0.021	0.037	0.033	0.018	0.766	0.220	0.236	0.460	0.391	0.863	0.264	0.082
	DQ	0.212	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.290	0.006	0.000	0.000	0.000	0.000	0.000	0.000
	TL	-1.242	-2.083	-2.470	-2.355	-2.257	-2.275	-2.151	-1.562	-1.112	-1.652	-1.617	-1.443	-1.070	-0.517	0.435	1.737

The exceedance rate and the p-values for the Kupiec (Ku) and Dynamic Quantile (DQ) tests are given. TL indicates the Diebold-Mariano statistic for predictive accuracy using the tick loss function, where the base model is that using the CO estimator.

Table 6 VaR evaluation for stock-pair c-ibm using the constant (CO), rolling window (RW), autoregressive (AR) and intraday dependence (ID) estimators for the Gaussian, Plackett, Student's-*t* and SJC copulas.

P	0.005	0.01	0.02	0.03	0.04	0.05	0.07	0.10	0.005	0.01	0.02	0.03	0.04	0.05	0.07	0.10
Gaussian																
CO	ex	0.024	0.046	0.080	0.099	0.111	0.123	0.147	0.195	0.036	0.052	0.080	0.101	0.111	0.121	0.143
	Ku	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	DQ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RW	ex	0.008	0.016	0.038	0.058	0.064	0.083	0.107	0.133	0.010	0.018	0.044	0.058	0.064	0.083	0.099
	Ku	0.388	0.220	0.011	0.001	0.013	0.002	0.000	0.018	0.167	0.109	0.001	0.001	0.013	0.002	0.015
	DQ	0.052	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000
AR	TL	-0.488	-1.935	-2.446	-2.554	-2.686	-2.572	-2.303	-2.175	-1.943	-2.462	-2.738	-2.652	-2.699	-2.559	-2.208
	ex	0.012	0.014	0.040	0.058	0.064	0.083	0.105	0.135	0.008	0.020	0.040	0.060	0.064	0.083	0.101
	Ku	0.062	0.405	0.005	0.001	0.013	0.002	0.004	0.012	0.388	0.050	0.005	0.001	0.013	0.002	0.009
	DQ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.053	0.000	0.000	0.000	0.000	0.000	0.000
ID	TL	0.541	-0.887	-1.762	-2.007	-2.237	-2.155	-1.940	-1.857	-1.931	-2.438	-2.729	-2.654	-2.675	-2.531	-2.185
	ex	0.010	0.016	0.042	0.060	0.066	0.089	0.113	0.143	0.014	0.022	0.046	0.060	0.074	0.089	0.113
	Ku	0.167	0.220	0.002	0.001	0.007	0.000	0.000	0.000	0.020	0.021	0.000	0.001	0.001	0.000	0.000
	DQ	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ID*	TL	-0.587	-2.026	-2.458	-2.595	-2.772	-2.618	-2.323	-2.211	-1.909	-2.452	-2.690	-2.671	-2.732	-2.549	-2.203
	ex	0.006	0.014	0.032	0.052	0.064	0.080	0.097	0.125	0.010	0.018	0.040	0.056	0.064	0.080	0.095
	Ku	0.766	0.405	0.081	0.010	0.013	0.005	0.022	0.068	0.167	0.109	0.005	0.002	0.013	0.005	0.033
	DQ	0.341	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000
	TL	-0.327	-1.732	-2.254	-2.421	-2.529	-2.435	-2.156	-1.991	-1.871	-2.373	-2.664	-2.616	-2.659	-2.525	-2.175
Student's-t																
CO	ex	0.012	0.044	0.080	0.093	0.111	0.123	0.147	0.193	0.012	0.038	0.076	0.091	0.111	0.125	0.157
	Ku	0.062	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.062	0.000	0.000	0.000	0.000	0.000	0.000
	DQ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RW	ex	0.006	0.014	0.036	0.056	0.066	0.085	0.105	0.137	0.006	0.012	0.028	0.056	0.064	0.083	0.113
	Ku	0.766	0.405	0.023	0.002	0.007	0.001	0.004	0.008	0.766	0.673	0.236	0.002	0.013	0.002	0.018
	DQ	0.342	0.008	0.000	0.000	0.000	0.000	0.000	0.000	0.332	0.014	0.000	0.000	0.000	0.000	0.000
AR	TL	-0.041	-1.698	-2.456	-2.452	-2.591	-2.545	-2.240	-2.084	0.389	-1.299	-2.294	-2.456	-2.647	-2.656	-2.419
	ex	-	-	-	-	-	-	-	-	0.006	0.010	0.036	0.056	0.066	0.083	0.107
	Ku	-	-	-	-	-	-	-	-	0.766	0.989	0.023	0.000	0.007	0.002	0.002
	DQ	-	-	-	-	-	-	-	-	0.321	0.266	0.000	0.000	0.000	0.000	0.000
ID	TL	-	-	-	-	-	-	-	-	0.452	-1.310	-2.263	-2.375	-2.546	-2.564	-2.238
	ex	0.008	0.014	0.042	0.060	0.068	0.091	0.113	0.145	0.006	0.016	0.042	0.058	0.072	0.091	0.115
	Ku	0.388	0.405	0.002	0.001	0.004	0.000	0.000	0.001	0.766	0.220	0.002	0.001	0.001	0.000	0.001
	DQ	0.050	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.335	0.000	0.000	0.000	0.000	0.000	0.000
ID*	TL	-0.034	-1.767	-2.466	-2.506	-2.683	-2.595	-2.265	-2.137	0.287	-1.419	-2.324	-2.464	-2.690	-2.657	-2.434
	ex	0.004	0.012	0.032	0.054	0.062	0.080	0.101	0.127	0.004	0.010	0.030	0.052	0.062	0.085	0.105
	Ku	0.736	0.673	0.081	0.005	0.021	0.005	0.009	0.050	0.736	0.989	0.142	0.010	0.021	0.001	0.004
	DQ	0.824	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.827	0.110	0.000	0.000	0.000	0.000	0.000
	TL	0.372	-1.450	-2.199	-2.301	-2.418	-2.366	-2.030	-1.838	0.626	-1.157	-2.141	-2.358	-2.536	-2.515	-2.263

The exceedance rate and the p-values for the Kupiec (Ku) and Dynamic Quantile (DQ) tests are given. TL indicates the Diebold-Mariano statistic for predictive accuracy using the tick loss function, where the base model is that using the CO estimator.

Table 7 VaR evaluation for stock-pair jpm-ibm using the constant (CO), rolling window (RW), autoregressive (AR) and realized dependence (ID) estimators for the Gaussian, s. Gumbel, Plackett and SJC copulas.

	p	0.005	0.01	0.02	0.03	0.04	0.05	0.07	0.10	0.005	0.01	0.02	0.03	0.04	0.05	0.07	0.10
Gaussian																	
CO	ex	0.030	0.046	0.066	0.087	0.107	0.123	0.151	0.181	0.024	0.030	0.054	0.078	0.093	0.111	0.147	0.181
	Ku	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RW	DQ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	ex	0.014	0.018	0.028	0.046	0.066	0.078	0.111	0.127	0.006	0.014	0.022	0.038	0.056	0.072	0.105	0.129
AR	Ku	0.020	0.109	0.236	0.055	0.007	0.009	0.001	0.050	0.766	0.405	0.768	0.325	0.090	0.037	0.004	0.036
	DQ	0.001	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.539	0.066	0.007	0.000	0.000	0.000	0.000	0.000
ID	TL	-1.867	-1.862	-2.001	-2.111	-2.062	-2.000	-1.974	-2.014	-1.013	-0.919	-1.053	-1.283	-1.508	-1.636	-1.708	-1.948
	ex	0.014	0.018	0.026	0.046	0.062	0.078	0.107	0.127	0.006	0.014	0.024	0.040	0.056	0.078	0.109	0.129
ID*	Ku	0.020	0.109	0.370	0.055	0.021	0.009	0.002	0.050	0.766	0.405	0.549	0.221	0.090	0.009	0.001	0.036
	DQ	0.001	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.550	0.017	0.001	0.000	0.000	0.000	0.000	0.000
ID	TL	-1.869	-1.830	-1.994	-2.101	-2.052	-1.963	-1.939	-1.953	-1.028	-0.989	-1.051	-1.219	-1.405	-1.517	-1.621	-1.890
	ex	0.014	0.016	0.030	0.050	0.066	0.076	0.109	0.129	0.006	0.014	0.026	0.046	0.062	0.078	0.109	0.133
ID	Ku	0.020	0.220	0.142	0.018	0.007	0.014	0.001	0.036	0.766	0.405	0.370	0.055	0.021	0.009	0.001	0.018
	DQ	0.001	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.528	0.066	0.001	0.000	0.000	0.000	0.000	0.000
ID*	TL	-1.896	-1.940	-2.094	-2.190	-2.134	-2.081	-2.058	-2.110	-1.008	-0.915	-1.170	-1.334	-1.513	-1.641	-1.742	-2.042
	ex	0.010	0.014	0.022	0.038	0.054	0.066	0.103	0.121	0.006	0.010	0.022	0.034	0.052	0.068	0.103	0.127
ID*	Ku	0.167	0.405	0.768	0.325	0.136	0.125	0.006	0.123	0.766	0.989	0.768	0.624	0.200	0.085	0.006	0.050
	DQ	0.018	0.069	0.007	0.000	0.000	0.000	0.000	0.000	0.520	0.255	0.006	0.000	0.000	0.000	0.000	0.000
ID	TL	-1.756	-1.649	-1.794	-1.895	-1.859	-1.837	-1.854	-1.860	-0.915	-0.843	-0.971	-1.204	-1.378	-1.572	-1.680	-1.924
Plackett																	
CO	ex	0.036	0.054	0.074	0.089	0.107	0.119	0.147	0.179	0.028	0.036	0.060	0.085	0.097	0.119	0.155	0.185
	Ku	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RW	DQ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	ex	0.014	0.018	0.030	0.046	0.064	0.074	0.107	0.123	0.008	0.014	0.024	0.042	0.060	0.072	0.107	0.131
AR	Ku	0.020	0.109	0.236	0.055	0.013	0.023	0.002	0.092	0.388	0.405	0.549	0.144	0.036	0.037	0.002	0.025
	DQ	0.001	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.158	0.066	0.001	0.000	0.000	0.000	0.000	0.000
ID	TL	-2.143	-2.215	-2.259	-2.239	-2.125	-1.976	-1.844	-1.807	-1.149	-1.138	-1.418	-1.759	-1.883	-1.931	-2.105	-2.237
	ex	0.016	0.018	0.028	0.046	0.068	0.076	0.107	0.125	0.010	0.014	0.024	0.048	0.064	0.078	0.111	0.131
ID	Ku	0.006	0.109	0.236	0.055	0.004	0.014	0.002	0.068	0.167	0.405	0.549	0.032	0.013	0.009	0.001	0.025
	DQ	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.085	0.061	0.001	0.000	0.000	0.000	0.000	0.000
ID	TL	-2.114	-2.168	-2.225	-2.196	-2.042	-1.870	-1.719	-1.707	-1.136	-1.122	-1.393	-1.708	-1.708	-1.721	-1.853	-1.984
	ex	0.014	0.018	0.032	0.052	0.068	0.078	0.109	0.127	0.010	0.014	0.026	0.048	0.068	0.080	0.111	0.135
ID*	Ku	0.020	0.109	0.081	0.010	0.004	0.009	0.001	0.050	0.167	0.405	0.370	0.032	0.004	0.005	0.001	0.012
	DQ	0.001	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.017	0.070	0.000	0.000	0.000	0.000	0.000	0.000
ID*	TL	-2.156	-2.259	-2.273	-2.231	-2.119	-1.967	-1.852	-1.845	-1.260	-1.287	-1.682	-1.909	-1.944	-1.986	-2.151	-2.313
	ex	0.014	0.016	0.026	0.046	0.056	0.070	0.105	0.121	0.006	0.010	0.022	0.034	0.054	0.066	0.103	0.125
ID*	Ku	0.020	0.220	0.370	0.055	0.090	0.056	0.004	0.123	0.766	0.989	0.768	0.624	0.136	0.125	0.006	0.068
	DQ	0.001	0.014	0.001	0.000	0.000	0.000	0.000	0.000	0.513	0.259	0.007	0.000	0.000	0.000	0.000	0.000
ID	TL	-2.109	-2.144	-2.149	-2.127	-2.030	-1.904	-1.801	-1.731	-1.054	-1.054	-1.277	-1.576	-1.656	-1.769	-1.923	-2.000

The exceedance rate and the p-values for the Kupiec (Ku) and Dynamic Quantile (DQ) tests are given. TL indicates the Diebold-Mariano statistic for predictive accuracy using the tick loss function, where the base model is that using the CO estimator.

Figure 9 1 and 10% VaR for the c-jpm stock pair estimated using the SJC copula. Dotted green lines are the portfolio returns. The constant dependence estimator (Co) is given by blue dashed lines, rolling window estimator (RW) by green solid lines, autoregressive estimator (AR) by magenta dashed-dot lines, intraday dependence estimator (ID) by solid black lines and adjusted intraday dependence estimator (ID*) by red dashed lines.

satisfies VaR adequacy at all quantiles. In all the rest of the cases, the differences in VaR estimates produced by the different estimators are rather small. The VaR produced by ID* tends to be marginally more negative than other estimators, which gives it the best performance in terms of VaR adequacy, and fair performance in terms of VaR accuracy. Partly due to its better hit rate performance, the ID* estimator tends to give the best DQ statistics and passes most of the DQ tests at 0.5% and 1% quantiles. The overall DQ results are poor for most cases, given the challenging time period.¹⁷

The reduction of accuracy of VaR using ID* as compared to ID could come from the adjustment factor, a_t , which could be time-varying. Furthermore, for copulas like the Gaussian and SJC, the dependence parameters have upper bounds that are often hit after applying the adjustment factor. Nevertheless, our results indicate that ID* gives the best results for VaR adequacy and in situations when accuracy is compromised, it gives a more conservative estimate than necessary, which is important for risk management purposes.

5 Conclusion

In this paper, we propose a new method for estimating the daily time-varying copula dependence parameter by using high frequency intraday data. The estimator is an ex-post estimate of the day's latent dependence

¹⁷ Note that even when other time periods are used, most models in Kuester, Mittnik, and Paolella (2006) do not pass the DQ test, and for the few that do pass the test (GARCH with skewed Student's- t , Mixed GED and EVT), they do so only at a certain quantile (mostly at 1%) and none of the models pass the test at all three quantiles (1%, 2.5% and 5%).

based on intraday high frequency data and we term this “intraday dependence.” While many non-gaussian copulas distributions are generally not stable under aggregation, we find that by using the scale-free GARCH residuals to estimate copula dependence in SCOMDY framework, the dependence is relatively stable under aggregation but is biased. We discuss the sources of the biasness (aggregation, not including overnight returns, Epps effects, and finite sampling), and consider the use of an adjustment or rescaling factor on the intraday dependence estimator.

By empirical application using three stock pairs in the period preceding and during the credit crisis of 2007, as with findings of Van den Goorbergh, Genest, and Werker (2005), Rodriguez (2007) and Patton (2006), we find evidence that the copula dependence parameter varies over time. We discuss the issue of intraday-seasonality and find that the GARCH(1,1) filter provides a fair correction of this effect. For VaR application, we find that intraday dependence exhibits the long-memory property akin to realized volatility and realized correlation, and forecast intraday dependence using the HAR model of Corsi (2009). This forecasted dependence was then used to estimate the 1-day ahead VaR during the crisis period.

We find that the intraday dependence estimator gives a more accurate VaR estimation compared to other commonly used methods such as the rolling window estimator and Patton (2006) autoregressive estimator, but often produces too many exceedances. The adjusted intraday dependence estimator gives a much superior performance in terms of VaR adequacy at all quantiles, at the cost of some accuracy. The loss of accuracy is severe when the corrected dependence hits the bounds of dependence in copulas such as Gaussian and SJC, and the VaR is overestimated (more negative). Despite this, VaR adequacy remains superior, and for risk management purposes, VaR overestimation is preferable to underestimation. Other than this, the adjusted intraday dependence estimator delivers good VaR accuracy.

Our research indicates the usefulness of using intraday data in capturing the daily latent dynamic dependence, and this can be especially important for risk management purposes. There are several areas to extend this research for the future. First, the various sources of biasness could be more exactly characterized, and issue about the optimal sampling frequency in the presence of such biases would then be better studied. Second, a potential usefulness of intraday dependence is that it allows for the selection of copulas to be changed daily, such that the most suitable copula can be adopted each day. This advantage can be harnessed for more sophisticated portfolio and risk management systems.

Acknowledgments: We would like to thank two anonymous referees, Bruce Mizrach (the Editor), Winfried Pohlmeier, Joachim Grammig and Yongwoong Lee for their insightful comments and suggestions, as well as Heikki Seppala for checking the mathematical proofs in the Web Appendix Section B. The research leading to these results has received funding from the European Community’s Seventh Framework Programme FP7-PEOPLE-ITN-2008 under grant agreement number PITN-GA-2009-237984. The funding is gratefully acknowledged.

Appendices

Figure 10 Constant and time varying dependence estimators Co (black solid lines), RW (black dashed-dotted lines), AR (blue solid lines) and ID (red dotted lines) for c-ibm between 2006 and 2008. Top row: Gaussian and Plackett copulas. Left middle and bottom rows: Student's- t copula (ρ and ν); right middle and bottom rows: SJC copula (upper and lower tail parameters).

Figure 11 Constant and time varying dependence estimators Co (black solid lines), RW (black dashed-dotted lines), AR (blue solid lines) and ID (red dotted lines) for jpm-ibm between 2006 and 2008. Top row: Gaussian and s. Gumbel copulas. Left middle: Plackett copula; right middle and bottom rows: SJC copula (upper and lower tail parameters).

References

- Alexander, C., and E. Lazar. 2005. "On the Continuous Limit of GARCH." Discussion Papers in Finance, ICMA Centre, University of Reading.
- Andersen, T. G., and T. Bollerslev. 1997. "Intraday Periodicity and Volatility Persistence in Financial Markets." *Journal of Empirical Finance* 4: 115–158.
- Audrino, F., and F. Corsi. 2010. "Modeling Tick-by-Tick Realized Correlations." *Computational Statistics and Data Analysis* 54: 2372–2382.
- Barndorff-Nielsen, O. E. 1995. "Normal Inverse Gaussian Processes and the Modeling of Stock Returns." Research report no. 200, Dept of Theoretical Statistics, Inst. of Mathematics, Univ. of Aarhus, Denmark.
- Barndorff-Nielsen, O. E., and N. Shephard. 2002. "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models." *Journal of the Royal Statistical Society, Series B* 64: 253–280.

- Barndorff-Nielsen, O., and N. Shephard. 2004. "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression and Correlation in Financial Economics." *Econometrica* 72: 885–925.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard. 2008. "Realised Kernels in Practice: Trades and Quotes." *Econometrics Journal* 4: 1–32.
- Berkes, I., and L. Horvath. 2003. "Limit Results for The Empirical Process of Squared Residuals in GARCH Models." *Stochastic Processes and their Applications* 105: 271–298.
- Breymann, W., A. Dias, and P. Embrechts. 2003. "Dependence Structures for Multivariate High-frequency Data in Finance." *Quantitative Finance* 3 (1): 1–14.
- Chan, N.-H., J. Chen, X. Chen, Y. Fan, and L. Peng. 2009. "Statistical Inference for Multivariate Residual Copula of GARCH Models." *Statistica Sinica* 19: 53–70.
- Chen, X., and Y. Fan. 2006. "Estimation and Model Selection of Semiparametric Copula-Based Multivariate Dynamic Models Under Copula Misspecification." *Journal of Econometrics* 135: 125–154.
- Chollete, L., A. Heinen, and A. Valdesogo. 2009. "Modeling International Financial Returns with a Multivariate Regime-Switching Copula." *Journal of Financial Econometrics* 7 (4): 437–480.
- Christoffersen, P., and F. Diebold. 2006. "Financial Asset Returns, Direction-of-Change Forecasting, and Volatility Dynamics." *Management Science* 52: 1273–1287.
- Corsi, F. 2009. "A Simple Approximate Long-Memory." *Journal of Financial Econometrics* 7 (2): 174–196.
- Corsi, F., S. Mittnik, C. Pigorsch, and U. Pigorsch. 2008. "The Volatility of Realized Volatility." *Econometric Review* 27 (1): 46–78.
- Creal, D., S. J. Koopman, and A. Lucas. 2013. "Generalized Autoregressive Score Models with Applications." *Journal of Applied Econometrics* 28: 777–795.
- Dias, A., and P. Embrechts. 2004. "Dynamic Copula Models for Multivariate High-frequency Data in Finance." Working paper, Department of Mathematics, ETH Zurich.
- Diks, C., V. Panchenko, and D. van Dijk. 2010. "Out-of-sample Comparison of Copula Specifications in Multivariate Density Forecasts." *Journal of Economic Dynamics & Control* 34: 1596–1609.
- Diks, C., V. Panchenko, O. Sokolinskiy, and D. van Dijk. 2014. "Comparing the Accuracy of Multivariate Density Forecasts in Selected Regions of the Copula Support." *Journal of Economic Dynamics & Control* 48: 79–94.
- Drost, F. C., and T. E. Nijman. 1992. "Temporal Aggregation of GARCH Processes." Discussion Papers 9066 and 9240, Tilburg University.
- Drost, F. C., and T. E. Nijman. 1993. "Temporal Aggregation of Garch Processes." *Econometrica* 61 (4): 909–927.
- Drost, F. C., and B. J. M. Werker. 1996. "Closing the GARCH gap: Continuous Time GARCH Modelling." *Journal of Econometrics* 474: 31–57.
- Engle, R. F., and S. Manganelli. 2004. "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles." *Journal of Business & Economic Statistics* 22 (4): 367–381.
- Fengler, M. R., and O. Okhrin. 2012. "Realized Copula." Discussion paper, Humboldt-Universität zu Berlin.
- Genest, C., K. Ghoudi, and L. P. Rivest. 1995. "A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions." *Biometrika* 82 (3): 543–552.
- Giacomini, E., W. Hardle, and V. Spokoiny. 2009. "Inhomogeneous Dependence Modeling with Time-Varying Copulae." *Journal of Business & Economic Statistics* 27 (2): 224–234.
- Harris, L. 1986. "A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns." *Journal of Financial Economics* 16: 99–117.
- Joe, H. 1997. *Multivariate Models and Dependence Concepts*. Chapman & Hall.
- Kim, J., A. M. Malz, and J. Mina. 1999. "RiskMetrics Technical Document." Technical report, New York: RiskMetrics Group.
- Komunjer, I. 2005. "Quasi-Maximum Likelihood Estimation for Conditional Quantiles." *Journal of Econometrics* 128 (1): 137–164.
- Kuester, K., S. Mittnik, and M. S. Paoletta. 2006. "Value-at-Risk Prediction: A Comparison of Alternative Strategies." *Journal of Financial Econometrics* 4 (1): 53–89.
- Martens, M., Y.-C. Chang, and S. J. Taylor. 2002. "A Comparison of Seasonal Adjustment Methods When Forecasting Intraday Volatility." *The Journal of Financial Research* XXV: 283–299.
- Nelsen, R. B. 1999. *An Introduction to Copulas*. New York: Springer-Verlag.
- Nelson, D. B. 1990. "ARCH Models as Diffusion Approximations." *Journal of Econometrics* 45: 7–38.
- Øigard, T. A., A. Hanssen, R. E. Hansen, and F. Godtliebsen. 2005. "EM-estimation and Modeling of Heavy-Tailed Processes with the Multivariate Normal Inverse Gaussian distribution." *Signal Processing* 85: 1655–1673.
- Patton, A. J. 2001. "Modelling Time-varying Exchange Rate Dependence Using the Conditional Copula." Discussion Paper, University of California, San Diego.
- Patton, A. 2006. "Modelling Asymmetric Exchange Rate Dependence." *International Economic Review* 47 (2): 527–556.
- Rodriguez, J. C. 2007. "Measuring Financial Contagion: A Copula Approach." *Journal of Empirical Finance* 14: 401–423.
- Salvatierra, I. D., and A. J. Patton. 2013. "Dynamic Copula Models and High Frequency Data." Working paper, Duke University.
- Sklar, A. 1959. "Fonctions de répartition à n dimensions et leurs marges." *Publ. Inst. Statist. Univ. Paris* 8: 229–231.
- Van den Goorbergh, R. W. J., C. Genest, and B. J. M. Werker. 2005. "Bivariate Option Pricing Using Dynamic Copula Models." *Insurance: Mathematics and Economics* 37: 101–114.

- Vortelinos, D. I. 2010. "The Properties of Realized Correlation: Evidence from the French, German and Greek Equity Markets." *The Quarterly Review of Economics and Finance* 50 (3): 273–290.
- Wiggins, J. B. 1991. "Empirical Tests of the Bias and Efficiency of the Extreme-Value Variance Estimator for Common Stocks." *The Journal of Business* 64 (3): 417–432.
- Wood, R. A., T. H. McInish, and J. K. Ord. 1985. "An Investigation of Transaction Data for NYSE Stocks." *Journal of Finance* 25: 723–739.

Supplemental Material: The online version of this article (DOI: 10.1515/snde-2013-0123) offers supplementary material, available to authorized users.