Liquidity commonality and risk management

Gregor N.F. Weiß*

Juniorprofessur für Finance, Technische Universität Dortmund

Hendrik Supper

Technische Universität Dortmund

February 29, 2012

Abstract We propose to model the joint distribution of bid-ask spreads and log returns of a stock portfolio by using Autoregressive Conditional Double Poisson and GARCH processes for the marginals and vine copulas for the dependence structure. By estimating the joint multivariate distribution of both returns and bid-ask spreads from intraday data, we incorporate the measurement of commonalities in liquidity and comovements of stocks and bid-ask spreads into the forecasting of three types of liquidity-adjusted Value-at-Risk (L-IVaR). In a preliminary analysis, we document strong extreme comovements in liquidity and strong tail dependence between bid-ask spreads and log returns across the firms in our sample thus motivating our use of a vine copula model. Furthermore, the backtesting results for the L-IVaR of a portfolio consisting of five stocks listed on the NASDAQ show that the proposed models perform well in forecasting liquidity-adjusted intraday portfolio profits and losses.

Keywords: Liquidity; Commonality; Vine Copulas; liquidity-adjusted intraday Value-at-Risk.

JEL Classification Numbers: C58, C53, G12, G14.

*Address: Corresponding author; Otto-Hahn-Str. 6a, D-44227 Dortmund, Germany, telephone: +49 231 755 4608, e-mail: *gregor.weiss@tu-dortmund.de*. Support by the Collaborative Research Centers "Statistical Modeling of Nonlinear Dynamic Processes" (SFB 823) at TU Dortmund University and "Economic Risk" (SFB 649) at Humboldt University Berlin of the German Research Foundation (DFG) is gratefully acknowledged. We thank Janina Mühlnickel and Janet Gabrysch for outstanding research assistance.

1 Introduction

Liquidity risk is of major concern to both investors and portfolio managers. Especially in times of market turmoil, overall market liquidity can dry up forcing investors to exit positions at increased cost. Consequently, investors prefer assets which are either liquid (see Amihud and Mendelson, 1986; Brennan and Subrahmanyam, 1996; Liu, 2006) or are least not exposed to systematic decreases in liquidity (Acharya and Pedersen, 2005; Sadka, 2006). While the finance literature has long been concerned with liquidity risk and market microstructure (Demsetz, 1968; Stoll, 1978; Amihud and Mendelson, 1980), the recent financial crisis has renewed interest in both the modeling of liquidity risk and in the analysis of its driving factors (see e.g. Cornett et al., 2011; Dick-Nielsen et al., 2012).

Although early empirical and theoretical studies on liquidity have concentrated on individual securities, the analysis of comovements in liquidity among individual stocks has become a corner stone of the microstructure literature starting with the seminal papers by Chordia et al. (2000), Hasbrouck and Seppi (2001) and Huberman and Halka (2001). Since then, studies on the commonality in liquidity have unanimously found clear empirical evidence for strong comovements in liquidity as proxied, e.g., by the bid-ask spreads of individual stocks. The determinants driving this commonality in liquidity, however, have remained relatively unknown until the recent study by Karolyi et al. (2012). They find that commonality in liquidity is positively correlated with high market volatility. In addition, liquidity commonalities were particularly strong during the Asian crisis and the recent financial crisis underlining the necessity of taking liquidity risk into account in multivariate risk modeling.

The incorporation of liquidity risk into the measurement of market risk has been a recurring topic in the risk management literature since the onset of the Value-at-Risk (VaR) concept as a de-facto industry standard. Although standard VaR lacks a rigorous consideration of liquidity risk, several extensions to certain forms of Liquidity-Adjusted VaRs (L-VaR) have been proposed in the literature (see e.g. Berkowitz, 2000; Bangia et al., 2002;

Qi and Ng, 2009). Parallel to the study of liquidity-adjusted risk measures, the availability of tick-by-tick data has enabled researchers to measure the liquidity of stocks at an ultra-high frequency (Chordia et al., 2000; Engle, 2000). While the increase in the speed of trading and the widespread availability of transaction data have led to an increase in the importance of intraday risk analysis (Dionne et al., 2009; Gourieroux and Jasiak, 2010), the modeling of intraday stock returns and bid-ask spreads is severely hampered by the irregular spacing of such data. Econometric research has thus concentrated on deriving appropriate models for the duration between transactions (Engle and Russell, 1998), for the estimation of intraday volatility (Andersen et al., 2000, 2003) and for the estimation of intraday risk measures (Giot, 2005; Dionne et al., 2009).

The econometrics literature includes several studies which focus on modeling and forecasting multivariate intraday stock returns or multivariate bid-ask spreads (see e.g. Engle and Russell, 1998; Breymann et al., 2003; Heinen and Rengifo, 2007; Groß-Klußmann and Hautsch, 2011; Li and Poon, 2011). Up to date, however, no study has analyzed the joint distribution of intraday returns and bid-ask spreads of multiple stocks. Although stock returns are known to be leptokurtic and non-normally distributed with time-varying dependence (see e.g. Longin and Solnik, 1995), and bid-ask spreads are known to comove across markets, little is known about the dependence (and consequently possible comovements) between liquidity and returns of different stocks.

In addition, the multivariate modeling of bid-ask spreads has so far been a purely econometric exercise with no connection to the rich literature on liquidity commonality.[1] Furthermore, there exist only few studies on the estimation of VaR from intraday data (Dionne et al., 2009) and, to the best knowledge of the authors, no work on multivariate liquidity-adjusted portfolio-VaR. Our paper tries to link the research on liquidity commonality and

[1] As stated by Heinen and Rengifo (2007), extending the standard Autogressive Conditional Duration (ACD) model of Engle and Russell (1998) to more than one time series has proven to be quite difficult. Considering the numerous problems one encounters when trying to model multivariate count data, it is, however, not surprising that studies such as the one by Heinen and Rengifo (2007) focus on the econometric side of the problem.

on the estimation of multivariate portfolio-VaR from high frequency data.

Therefore, we propose a multivariate econometric model based on vine copulas for estimating and forecasting liquidity-adjusted risk measures for multivariate stock portfolios and apply the proposed model on returns and bid-ask spreads reconstructed from high frequency data. We start our analysis by performing a variety of diagnostic tests on the dependence structure between intraday stock returns and quoted bid-ask spreads for selected companies on the NASDAQ 100 in 2009.[2] These preliminary tests provide us with ample empirical evidence of not only strong linear correlation but also strong tail dependence between bid-ask spreads across individual stocks. We then turn to modeling the marginal behavior of stock returns and bid-ask spreads by the use of GARCH processes and the Autoregressive Conditional Double Poisson model, respectively (Heinen, 2003; Groß-Klußmann and Hautsch, 2011). As the dependence structure between intraday stock returns and bid-ask spreads stemming from different companies has not been analyzed before in the literature,[3] our dependence model is required to be able to capture a wide range of possible linear and non-linear dependencies. We therefore resort to the concept of vine copulas (Heinen and Valdesogo, 2008; Aas et al., 2009) to model the dependence structure between the returns and bid-ask spreads of multiple stocks. Based on our estimated multivariate model for returns and bid-ask spreads, we then forecast and backtest several types of liquidity-adjusted risk measures to illustrate the superiority of our proposed method.

The contributions of this study relative to the existing literature on liquidity commonality and liquidity risk management are significant and numerous. While previous studies on the commonality in liquidity have documented strong positive *linear correlation* between measures of liquidity, this study is the first to find empirical evidence for a strong *non-linear*

[2]Obviously, other measures of liquidity could have also been used. The majority of studies in the risk management literature, however, regularly employs the quoted bid-ask spread to measure liquidity (see e.g. Berkowitz, 2000; Bangia et al., 2002).

[3]The relationship between the returns and the bid-ask spread of the same company, on the other hand, is well documented, see e.g. Amihud and Mendelson (1986) who find that average portfolio risk-adjusted returns increase with their bid-ask spread with the slope of the return-spread relationship decreasing with the spread.

dependence in the form of significant tail dependence as well. This paper is also the first to integrate our understanding of liquidity commonality into the estimation and forecasting of liquidity-adjusted risk measures. In addition, this paper constitutes the first multivariate model for the joint distribution of the returns and bid-ask spreads of multiple stocks. Finally, our paper adds to the fastly growing literature on the use of vine copulas in risk management and asset pricing applications making full use of the vines' flexibility.

The results presented in this paper show that the proposed multivariate model for bid-ask spreads and intraday returns performs exceptionally well in forecasting liquidity-adjusted portfolio losses. While losses are adequately bounded below by our liquidity-adjusted VaR forecasts, our models are not too conservative and thus prevent investors like, e.g., banks from reserving unnecessary risk capital buffers.

The remainder of this article is structured as follows. Section 2 presents the econometric methodology. Section 3 discusses the empirical study as well as the results. Concluding remarks are given in Section 4.

2 Econometric methodology

The purpose of this section is to outline the econometric models for intraday bid-ask spreads and returns, their multivariate dependence structure as well as liquidity-adjusted intraday VaR.

We start with the models for intraday spreads and returns which form the fundamental building blocks of the VaR models described later.

2.1 The Autoregressive Conditional Double Poisson Model

Loosely speaking, liquidity can be seen as the ability of a market to allow for immediate trading even large amounts to minimal costs without causing remarkable price movements; hence, liquidity risk results from the difference between transaction and market price.

Kyle (1985) provides a formal definition and introduces three components of liquidity including tightness, depth and resiliency; according to Amihud and Mendelson (1986), a natural measure of liquidity (or rather liquidity risk) is the spread between the bid and ask prices. Bangia et al. (2002) classify liquidity risk into exogenous liquidity risk which arises from the general conditions of a market and is equal to all participants and endogenous liquidity risk which refers to the volume of individual trading positions and which is idiosyncratic.

The bid-ask spread has become a key parameter in modeling financial data and captures the costs of immediate trading, which can be explained by the liquidity suppliers' purchasing at the bid and selling at a higher ask price in order to recoup their own costs. Empirical properties like intraday seasonalities, linear dependence and commonalities are given in the context of the data description in section 3.

Statistically, bid-ask spreads belong to the class of discrete count data, because they count the number of ticks between bid and ask prices. To accurately model time series of count data measured at high frequency, there are two important aspects to consider: on the one hand the model has to cope with the issues of discreteness and serial dependence, on the other hand its estimation procedure has to be tractable for a large number of observations. Following Groß-Klußmann and Hautsch (2011), we drop the traditional approaches reviewed by Cameron and Trivedi (1998), MacDonald and Zucchini (1997) and McKenzie (2003) because of their complex estimation procedures and adopt the Autoregressive Conditional Double Poisson Model (ACDP hereafter) that belongs to the class of observation-driven models and is based on the Autoregressive Conditional Poisson Model (ACP) introduced by Rydberg and Shephard (1999).[4]

Since the Poisson distribution is a natural starting point for counts (see Heinen, 2003), the ACP assumes that the counts follow a Poisson distribution with an autoregressive mean. Extending this approach, the ACDP uses the Double Poisson distribution as proposed by

[4]Further details of the ACDP model are explained in Heinen (2003), Fokianos et al. (2009) and Ferland et al. (2006).

Efron (1986) instead and allows for a more flexible modeling of the spreads.[5] To provide a mathematical description of this model we start with some notations: let $M, N, d \in \mathbb{N}$ denote the number of trading days in the sample, the number of observations per day and the number of stocks, respectively, and let I, J, K be index sets given by $I := [1, M] \cap \mathbb{N}$, $J := [0, N-1] \cap \mathbb{N}_0$ and $K := [1, d] \cap \mathbb{N}$, where $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Further, let $N_0, N_1 \in \mathbb{N} \cup \{0\}$ with $N_0 < N_1$ and let $\mathcal{Z} := \{N_0 = n_0 < n_1 < ... < n_{N-1} \leq N_1\}$ be a partition of $[N_0, N_1]$ with equidistant points $n_j \in \mathbb{N}_0$ $(j \in J)$ in ascending order and norm $\Delta_{\mathcal{Z}}$, denoting the time grid of observation points for the sample days.[6] Let $A^k_{i,j}$ and $B^k_{i,j}$ denote the (best) bid and the (best) ask price (respectively) and $S^k_{i,j} := A^k_{i,j} - B^k_{i,j}$ the quoted bid-ask spread of stock k closest to time $n_j \in \mathcal{Z}$, realized at the i-th trading day $(i \in I, k \in K)$. Moreover, let DPois(λ, γ) be the Double Poisson distribution (Efron, 1986) and $\mathcal{F}^k_{i,j}$ the information available on the series of stock k up to and including time n_j of the i-th day.[7] Let L denote the lag operator defined by $L^b S^k_{i,j} := S^k_{i,j-b}$ for $b \in \mathbb{Z}$ and let two lag polynomials Φ_k and Ψ_k be given as $\Phi_k(L) := \sum_{m=1}^{q} \phi^k_m L^m$ and $\Psi_k(L) := \sum_{n=1}^{p} \psi^k_n L^n$, where $p, q \in \mathbb{N}_0$ refer to the order of the process and the coefficients are restricted by $\phi_m > 0$, $m = 1, ..., q$ as well as $\psi^k_n > 0$, $n = 1, ..., p$.

Then, the ACDP(p, q) process $\{S^k_{i,j}\}_{j \in J}$ is defined by

$$\begin{aligned} S^k_{i,j} | \mathcal{F}^k_{i,j-1} &\sim \text{DPois}(\lambda^k_{i,j}, \gamma_k), \\ \lambda^k_{i,j} &= c_k + \Phi_k(L) S^k_{i,j} + \Psi_k(L) \lambda^k_{i,j}, \end{aligned} \tag{1}$$

where $i \in I, k \in K$ and $c_k > 0$.[8]

The ACDP assumes the conditional distribution of the spread to be Double Poisson. The

[5] Since the (Double) Poisson distribution is defined on $\mathbb{N} \cup \{0\}$, all prices and spreads are expressed as multiples of ticks, i.e. multiples of the minimum tick size which equals 0.01 US-Dollar cent.

[6] The interval $[N_0, N_1]$ refers to the observed trading hours; $\mathcal{Z}$ is equal to each stock and trading day of the sample.

[7] Hence, $\mathcal{F}^k_{i,j}$ considers $(i-1)N + (j+1)$ spread observations (of stock k).

[8] Note that we have $L^b S^k_{i,j} = S^k_{i-1,N+j-b}$ and $L^b \lambda^k_{i,j} = \lambda^k_{i-1,N+j-b}$ in case of $j < b \leq N + j$.

conditional probability mass function of $S_{i,j}^k$ is then given by

$$\mathbb{P}(S_{i,j}^k = l|\lambda_{i,j}^k, \gamma_k) = c_k(\lambda_{i,j}^k, \gamma_k) \cdot \gamma_k^{\frac{1}{2}} \exp(-\gamma_k \lambda_{i,j}^k) \frac{\exp(-l) l^l}{l!} \left(\frac{\exp(1)\lambda_{i,j}^k}{l} \right)^{\gamma_k l}, \; l \in \mathbb{N}_0, \quad (2)$$

where Efron (1986) suggests the following expression for the constant $c_k(\lambda_{i,j}^k, \gamma_k)$:

$$\frac{1}{c_k(\lambda_{i,j}^k, \gamma_k)} = 1 + \frac{1-\gamma_k}{12\lambda_{i,j}^k \gamma_k} \left(1 + \frac{1}{\lambda_{i,j}^k \gamma_k} \right). \quad (3)$$

He deduces expressions for the mean and variance of the Double Poisson as well, and we have

$$\mathbb{E}\left[S_{i,j}^k \middle| \mathcal{F}_{i,j-1}^k\right] = \lambda_{i,j}^k \;\text{ and }\; \mathbb{V}\text{ar}\left[S_{i,j}^k \middle| \mathcal{F}_{i,j-1}^k\right] = \frac{\lambda_{i,j}^k}{\gamma_k}, \quad (4)$$

so the ACDP model can generate both conditional overdispersion (in case of $\gamma_k < 1$) and conditional underdispersion (in case of $\gamma_k > 1$).

Assuming $p = q = 1$, Heinen (2003) characterizes the unconditional distribution by yielding formulas for the unconditional mean and variance which are given by

$$\mathbb{E}\left[S_{i,j}^k\right] = \frac{c_k}{1-(\phi_1^k + \psi_1^k)} \;\text{ and }\; \mathbb{V}\text{ar}\left[S_{i,j}^k\right] = \frac{1}{\gamma_k} \cdot \frac{\mathbb{E}\left[S_{i,j}^k\right](1-(\phi_1^k+\psi_1^k)^2 + \phi_1^{k,2})}{1-(\phi_1^k+\psi_1^k)^2} \geq \mathbb{E}\left[S_{i,j}^k\right]. \quad (5)$$

Hence, the expression for the unconditional mean is identical to the mean of an ARMA process and the unconditional distribution exhibits overdispersion in the case of $\gamma_k \leq 1$, whereas $\gamma_k > \frac{1-(\phi_1+\psi_1^k)^2+\phi_1^{k,2}}{1-(\phi_1^k+\psi_1^k)^2}$ leads to underdispersion.

Furthermore, as stated by Heinen (2003) and shown by Ferland et al. (2006), the ACP model yields covariance stationary and strictly stationary solutions if $\sum_{m=1}^q \phi_i^k + \sum_{n=1}^p \psi_j^k < 1$; because of (4) this holds for the ACDP model as well.

The ACDP model is straightforwardly estimated by maximum likelihood, where the log

likelihood can be calculated for each stock $k \in K$ using (1) and is given by

$$\log \mathcal{L}(\theta, \gamma_k) = \sum_{i \in I} \sum_{j \in J} \left(\log \left(c_k(\lambda_{i,j}^k, \gamma_k) \right) + \frac{1}{2} \log \left(\gamma_k \right) - \gamma_k \lambda_{i,j}^k + \gamma_k S_{i,j}^k \left(1 + \log \left(\frac{\lambda_{i,j}^k}{S_{i,j}^k} \right) \right) \right), \tag{6}$$

where $\theta := \left(c_k, \phi_1^k, ..., \phi_q^k, \psi_1^k, ..., \psi_p^k \right)^\top$ contains all the parameters of the autoregressive conditional intensity.[9]

2.2 Modeling intraday returns with GARCH processes

In order to model intraday mid price returns we adopt the model of Giot (2005) who applies different specifications of GARCH models to deseasonalized returns to capture their intraday volatility dynamics. Before presenting the mathematical basics of this model, we extend our notations introduced in the preceding section: let $P_{i,j}^k := \frac{A_{i,j}^k + B_{i,j}^k}{2}$ denote the mid price of stock k closest to time $n_j \in \mathcal{Z}$ at the i-th day ($i \in I, j \in J, k \in K$) and $R_{i,j}^k := \log \left(P_{i,j}^k \right) - \log \left(P_{i,j-1}^k \right)$ the raw log returns associated with two adjacent mid prices. According to Giot (2005), intraday volatility can be modeled by standard volatility models after taking into account the strong intraday seasonality pattern of volatility documented by Andersen and Bollerslev (1997). Following Giot (2005) we assume a deterministic seasonality in the intraday volatility and compute the deseasonalized log returns $r_{i,j}^k$ from the raw log returns as

$$r_{i,j}^k = \frac{R_{i,j}^k}{\sqrt{\xi_j^k}}, \tag{7}$$

[9] We omitted the term $S_{i,j}^k \left(\log \left(S_{i,j}^k \right) - 1 \right) - \log \left(S_{i,j}^k ! \right)$ in the expression for the log likelihood because it is independent of θ and vanishes in the score.

where ξ_j^k is the deterministic pattern of the intraday volatility of stock k for the time interval $[n_{j-1}, n_j]$ given by the expected volatility conditioned on time-of-day and computed as

$$\xi_j^k = \frac{1}{|I|} \sum_{i \in I} \left(R_{i,j}^k\right)^2, \; j \in J, k \in K. \tag{8}$$

Having filtered the intraday seasonalities, we apply a standard GARCH$(1,1)$ model with normally distributed innovations to capture the dynamics of the intraday returns.
Hence, the model for the intraday returns used in the framework of this article is given by

$$\begin{aligned} r_{i,j}^k &= \frac{R_{i,j}^k}{\sqrt{\xi_j^k}}, \\ r_{i,j}^k &= \mu_k + \varepsilon_{i,j}^k \sqrt{h_{i,j}^k}, \\ h_{i,j}^k &= \omega_k + \alpha_1^k \left(\varepsilon_{i,j-1}^k\right)^2 h_{i,j-1}^k + \beta_1^k h_{i,j-1}^k, \end{aligned} \tag{9}$$

where μ_k is the expected return, $\varepsilon_{i,j}^k \sim$ IID $\mathcal{N}(0,1)$, $\omega_k > 0, \alpha_1^k \geq 0$ and $\beta_1^k \geq 0$ for $i \in I, j \in J, k \in K$.

2.3 Joint modeling of returns and liquidity with vine copulas

We now turn to the task of modeling the joint distribution of (deseasonalized) mid price returns $r_{i,j}^k$ and bid-ask spreads $S_{i,j}^k$ of *several* stocks.[10] The model should therefore be able to capture the dependence between the mid price returns, the commonalities between the spreads and a possible dependence between returns and spreads.

In order to model the joint distribution of returns and spreads, we follow in the footsteps of Nolte (2008) who models the price changes, transaction volumes, bid-ask spreads and intertrade durations of single stocks by the use of a parametric copula. In contrast to his work, however, we extend this idea to the modeling of the returns and spreads of multiple stocks and employ a D-vine model to deal with the increase in dimensionality and to allow

[10]Note that d stocks thus require a distributional model in $2d$ dimensions.

for a more flexible modeling.

In the statistics literature, several papers starting with Joe (1996, 1997), Bedford and Cooke (2001, 2002) and Whelan (2004) have proposed models based on hierarchical Archimedean copulas and vine copulas (sometimes referred to as pair-copula constructions, PCC) to model high-dimensional joint distributions. The seminal work by Aas et al. (2009) introduced vine copulas to the field of risk management and spurred a surge in empirical applications of vine copulas (see Chollete et al., 2009; Heinen and Valdesogo, 2008; Aas and Berg, 2009; Min and Czado, 2010, 2011). For our purpose in this study, vine copulas are especially appropriate for two reasons: first, vine copulas allow the modeling of high-dimensional distributions. Second, vine copulas are composed of bivariate building blocks (called pair-copulas) with each bivariate pair-copula capturing the conditional dependence between two variables. As the dependence structure between returns and spreads is presumably quite complex, this flexibility allows us to model each bivariate pair of returns and/or spreads separately. We quickly review the definition and major properties of vine copulas below.[11]

Consider a d-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with joint density function $f(x_1, \ldots, x_d)$ and distribution function $F(x_1, \ldots, x_d)$. According to Bedford and Cooke (2001), the joint density f can be expressed as a product of the marginal densities and a set of conditional bivariate copulas. More precisely, we have

$$f(\mathbf{x}) = \prod_{k=1}^{d} f(x_k) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,j+1|1,\ldots,j-1}(F(x_j|x_1, \ldots, x_{j-1}), F(x_{j+i}|x_1, \ldots, x_{j-1})) \tag{10}$$

with bivariate conditional copulas $c_{i,i+j|i+1,\ldots,i+j-1}$ which are referred to as *pair-copulas*. The decomposition given above is known as a *canonical* or *C-vine* and is only one of several ways the density can be split up.[12] Then, for a d-dimensional density, there exist $d!/2$ different

[11] A rigorous treatment of vine copulas and their properties can be found in Joe (1996) and Bedford and Cooke (2001, 2002).

[12] The decompositions described here are special cases of the more general *regular vines* which are described e.g. in Brechmann and Czado (2011).

C-vines depending on the initial sorting of the d random variables. As each of the $d(d-1)/2$ bivariate pair-copula can be chosen from a different parametric copula family, a vine copula model is considerably more flexibile than a traditional d-dimensional copula model.

Different compositions of the density f are given by the so-called *D-vines* defined as

$$f(\mathbf{x}) = \prod_{k=1}^{d} f(x_k) \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j|i+1,\ldots,i+j-1}(F(x_i|x_{i+1},\ldots,x_{i+j-1}), F(x_{i+j}|x_{i+1},\ldots,x_{i+j-1})). \tag{11}$$

Again, for a d-dimensional density there are $d!/2$ different possible decompositions.

The estimation of a vine's parameters via Maximum-Likelihood is straightforward and is explained in detail in Aas et al. (2009). The selection of the $d(d-1)/2$ parametric pair-copulas, on the other hand, is a much more delicate task requiring the use of either goodness-of-fit tests (Aas et al., 2009) or model selection criteria like Akaike's Information Criterion (AIC; Dißmann et al., 2012). Here, we follow Brechmann and Czado (2011) and Dißmann et al. (2012) and employ AIC for selecting the best-fitting vine copula model.

Then, let $\varepsilon_{i,j}^k$ be the residuals of the GARCH processes for the mid-price returns and let $S_{i,j}^k$ be the bid-ask spreads of the k-th stock ($k \in K$). We then follow Nikoloulopoulos et al. (2012) and model the joint distribution of the GARCH innovations instead of the returns themselves yielding the $2d$-dimensional model

$$\mathbf{X} = \left(\varepsilon_{i,j}^1, \ldots, \varepsilon_{i,j}^d, S_{i,j}^1, \ldots, S_{i,j}^d\right) \tag{12}$$

for the joint distribution of the returns and spreads.[13] The marginals are modeled as normally distributed and Double Poisson distributed, respectively, thus the spread marginals exhibit a discrete distribution which causes some complications in regard to the applicability of the theoretical framework underlying the use of copulas. A pivotal requirement within this framework is the availability of the Probability Integral Transformation Theorem (PITT) of

[13]Note that we only consider unconditional copulas in this work due to the already large number of parameters that need to be estimated. As we reestimate each model after forecasting 24 observations, the negative effect of using an unconditional model for the dependence structure should, however, be neglectable.

Fisher (1932) which states that the probability integral transform (PIT) of a random variable under its marginal distribution is distributed uniformly on $[0;1]$. The PITT, however, only holds for continuous distributions and does not apply in case of discretely distributed marginals. To overcome this difficulty we use the continuousation approach described in Heinen and Rengifo (2007) and introduce the continuoused spread variable

$$\tilde{S}_{i,j}^{k} := S_{i,j}^{k} + (U - 1)\,,$$

where U is uniformly distributed on $[0;1]$. The PIT of $\tilde{S}_{i,j}^{k}$ is given by

$$\tilde{F}_{i,j}^{k}(\tilde{S}_{i,j}^{k}) = U \cdot F_{i,j}^{k}(S_{i,j}^{k}) + (1 - U) \cdot F_{i,j}^{k}(S_{i,j}^{k} - 1) \tag{13}$$

and follows a Uniform distribution on $[0;1]$, where $\tilde{F}_{i,j}^{k}$ and $F_{i,j}^{k}$ denote the cumulative distribution functions of $\tilde{S}_{i,j}^{k}$ and $S_{i,j}^{k}$, respectively.[14] We now replace the $S_{i,j}^{k}$ in (12) by their continuoused versions and adopt the $2d$-dimensional model

$$\tilde{\mathbf{X}} = \left(\varepsilon_{i,j}^{1}, \ldots, \varepsilon_{i,j}^{d}, \tilde{S}_{i,j}^{1}, \ldots, \tilde{S}_{i,j}^{d}\right) \tag{14}$$

for the joint distribution of the returns and the spreads.[15] In this way, we capture the complete dependence structure of returns and spreads and ensure that the theoretical conditions for the use of copulas are met as well.

As a model for the dependence structure, we employ a $2d$-dimensional D-vine copula. We argue that the D-vine is the best suited vine model for our purposes as it permits us to first model the dependence structure between the bid-ask spreads and returns separately. The conditional dependence between spreads and returns are modeled during later stages of the model estimation. An example for a D-vine with four variables (spreads and returns of two stocks) is illustrated in Figure 1.

[14] A formal proof can be found in Brockwell (2007).

[15] Note that continuousation does not alter the dependence between the marginals.

— insert Figure 1 here —

As one can see from Figure 1, the specific sorting of the variables allows us to first model the dependence between the (in this case two) spreads and the returns unconditionally (the nodes on level T2). On the next level of the vine, we model the conditional dependence between the spreads and returns of the same stock given information on the spreads or returns of the remaining stock. Finally, on the last layer T4 of the vine, the cross-dependence between the returns on the first and the spreads of the second stock are modeled given the remaining two variables.

2.4 Liquidity-adjusted intraday VaR

Having dealt with the models for the intraday spreads and returns as well as their multivariate dependence structure, we now present the different risk measures which are used in our empirical study concentrating on the liquidity-adjusted intraday VaR (L-IVaR). We present three different models including the model proposed by Bangia et al. (2002), the model suggested by Heude and Wynendaele (2001) and a model based on liquidity-adjusted returns each of which uses the information provided by the models previously presented in a particular way.

2.4.1 The L-IVaR model of Bangia et al. (BDSS, 2002)

As stated in section 2.1, we are interested in modeling the exogenous liquidity risk which is a result of general market characteristics and relevant for all market participants. Since (exogenous) liquidity risk arises from the deviation of transaction prices from market prices and can be measured by the bid-ask spread, Bangia et al. (2002) propose a methodology that incorporates exogenous liquidity risk into VaR by considering the spread as well as its volatility and adding an exogenous cost of liquidity component (COL) to standard VaR.
In terms of our notations and adjusted to an intraday horizon, their model for the L-IVaR

of the k-th stock in time n_j of the i-th day takes the following form:

$$\text{L-IVaR}_{i,j}^k = P_{i,j}^k \cdot \left[\left(1 - \exp\left(\sqrt{\xi_j^k}\mu_k - z_{1-\alpha}\sqrt{h_{i,j}^k \xi_j^k} \right) \right) + \frac{1}{2}\left({}_r\bar{S}^k + s_\alpha \sigma_s \right) \right], \tag{15}$$

where $z_{1-\alpha}$ denotes the $1-\alpha$ quantile of the standard normal distribution, ${}_r\bar{S}^k$ describes the average relative spread and s_α and σ_s refer to the α quantile and standard deviation of the relative spread distribution; thereby the relative spread is defined as ${}_rS_{i,j}^k := \frac{S_{i,j}^k}{P_{i,j}^k}$ and ${}_r\bar{S}^k$ results from averaging ${}_rS_{i,j}^k$ over i and j.

Hence, the first summand in (15) is given by

$$\text{IVaR}_{i,j}^k = P_{i,j}^k \cdot \left(1 - \exp\left(\sqrt{\xi_j^k}\mu_k - z_{1-\alpha}\sqrt{h_{i,j}^k \xi_j^k} \right) \right) \tag{16}$$

and yields the standard IVaR of the intraday mid prices that result from the model for the intraday returns described in section 2.2.[16]

The second summand is defined as

$$\text{COL}_{i,j}^k = \frac{1}{2} \cdot P_{i,j}^k \cdot \left({}_r\bar{S}^k + s_\alpha \sigma_s \right) \tag{17}$$

and describes the exogenous cost of liquidity which incorporates the (exogenous) liquidity risk and considers both the relative spread and its volatility.[17]

The total risk is thus split up into price and liquidity risk and (15) takes the simple form

$$\text{L-IVaR}_{i,j}^k = \text{IVaR}_{i,j}^k + \text{COL}_{i,j}^k. \tag{18}$$

Thus, the L-IVaR proposed by Bangia et al. (2002) adjusts the mid price to price and liquidity risk in an intuitive way; in order to compute the L-IVaR at the portfolio level for

[16] Note that following Giot (2005) we use the $\sqrt{\xi_j^k}$ factor to re-introduce the seasonality component of the intraday volatility (which was removed prior to modeling the intraday returns via GARCH).

[17] ${}_r\bar{S}^k$ serves as a normalizing device providing comparability and s_α can be interpreted as a scaling factor for the volatility to cover α% of the spread situations (see Bangia et al., 2002).

an equally weighted portfolio consisting of d stocks, we can simply sum up the adjusted mid prices of the single stocks according to their weights and get

$$\text{PF-L-IVaR}_{i,j}^{d} = \frac{1}{d}\sum_{k=1}^{d} \text{L-IVaR}_{i,j}^{k}. \tag{19}$$

This model, however, faces some drawbacks: on the one hand there are problems in estimating the scaling factor, if there is no spread distribution specified, because the spread distribution is far from normal; on the other hand the way price and liquidity risk are aggregated assumes extreme events in returns and extreme events in spreads to happen simultaneously.[18] In the next section we present a model proposed due to Heude and Wynendaele that is based on the same idea as the one by Bangia et al. but which aims at rectifying the drawbacks of the latter.

2.4.2 The L-IVaR model of Heude and Wynendaele (HW, 2001)

The model approach by Heude and Wynendaele (2001) is also based on the idea of adjusting the mid price to price risk and (exogenous) liquidity risk, but takes into consideration the drawbacks of the model proposed by Bangia et al. (2002). Using our notations, the model can be summarized as follows:

$$\text{L-IVaR}_{i,j}^{k} := P_{i,j}^{k} \cdot \left[\left(1 - \left(1 - \frac{{}_r\bar{S}^k}{2}\right) \cdot \exp\left(\sqrt{\xi_j^k}\mu_k - z_{1-\alpha}\sqrt{h_{i,j}^k \xi_j^k}\right)\right) + \frac{1}{2}\left({}_rS_{i,j}^k - {}_r\,\bar{S}^k\right)\right]. \tag{20}$$

The first summand of (20) is given by

$$P_{i,j}^{k} - P_{i,j}^{k} \cdot \left(1 - \frac{{}_r\bar{S}^k}{2}\right) \cdot \exp\left(\sqrt{\xi_j^k}\mu_k - z_{1-\alpha}\sqrt{h_{i,j}^k \xi_j^k}\right) \tag{21}$$

[18] Note that Bangia et al. (2002) assume a perfect correlation between (exogenous) liquidity risk and VaR, which leads to an overestimation of the L-IVaR; Heude and Wynendaele (2001), however, evidence that most of the extreme events in spreads occur when the return is around its mean.

and describes the basic IVaR adjusted to liquidity, where the VaR concept is directly applied to a theoretical bid obtained from the mid prices adjusted by half the average spread. In this way, the joint impact of price and (exogenous) liquidity risk is captured in a single expression.[19]

The second summand is defined as

$$\frac{1}{2} \cdot P_{i,j}^{k} \cdot \left({}_{r}S_{i,j}^{k} -_{r} \bar{S}^{k}\right) \tag{22}$$

and considers the dynamic aspect of the liquidity factor by comparing the immediate relative spread with the average relative spread. If the former exceeds the later, this component will increase the VaR number (and reduce it in the opposite case).

Hence, the model proposed by Heude and Wynendaele builds on the same idea as the model suggested by Bangia et al. (2002) but improves some of its drawbacks.[20] In the next section, we present a third approach for measuring L-IVaR which is based on liquidity-adjusted returns.

2.4.3 Estimating L-IVaR from liquidity-adjusted returns (AR)

The L-IVaR models presented in the previous sections are based on certain adjustments of the mid price to introduce (exogenous) liquidity risk in VaR. The approach based on liquidity-adjusted returns, however, arises from an appropriate adjustment for the returns and is grounded on the idea of incorporating (exogenous) liquidity risk into returns and applying the standard VaR to these modified returns.

Since trading costs reduce realized returns and due to the capacity of relative spreads as

[19] This is in contrast to Bangia et al. (2002), who split up this joint impact into two components.

[20] Heude and Wynendaele (2001) extend the expression in (20) and incorporate the endogenous liquidity risk by integrating the quantities liquidated. We are interested in the exogenous liquidity risk and therefore keep the term in (20).

normalizing devices (Bangia et al., 2002), we adjust the returns in the following way:

$$ {}_{\text{adj}}r_{i,j}^{k} := \frac{R_{i,j}^{k} - \frac{1}{2}{}_{r}S_{i,j}^{k}}{\sqrt{\xi_j^k}} = r_{i,j}^{k} - \frac{{}_{r}S_{i,j}^{k}}{2\sqrt{\xi_j^k}}, \tag{23} $$

where $R_{i,j}^{k}$ and $r_{i,j}^{k}$ denote the raw and the deseasonalized return (as in section 2.2), respectively, and $i \in I, j \in J, k \in K$.[21] Hence, (23) yields the (deseasonalized) return after taking into consideration costs of liquidation, i.e. the effectively realized return.[22]

Applying the VaR concept to ${}_{\text{adj}}r_{i,j}^{k}$, we define

$$ \text{L-IVaR}_{i,j}^{k} := P_{i,j}^{k} \cdot \left[1 - \exp\left(\sqrt{\xi_j^k}\mu_{k,\text{adj}} +_{\text{adj}} q_{1-\alpha}\sqrt{{}_{\text{adj}}h_{i,j}^{k}\xi_j^k}\right)\right], \tag{24} $$

where $\mu_{k,\text{adj}}, {}_{\text{adj}}h_{i,j}^{k}$ and ${}_{\text{adj}}q_{1-\alpha}$ characterize the distribution of the liquidity-adjusted returns and denote the mean, volatility and the $1-\alpha$ quantile.[23]

In order to produce a PF-L-IVaR based on this approach that can be used as a benchmark (or rather an upper bound) for the models proposed by Bangia et al. (2002) and Heude and Wynendaele (2001), we use equation (19) and simply sum up the L-IVaRs of the individual stocks in our portfolio according to their portfolio weights.

3 Empirical study

3.1 Data

Our empirical study is based on high-frequency data from the National Association of Securities Dealers Automated Quotations (NASDAQ) provided by the database LOBSTER.

Over the years, NASDAQ has emerged as a fully computerized, web-based trading platform which provides systems that link all of the liquidity suppliers in a given stock together

[21] Amihud and Mendelson (1986) use a similar adjustment and call (23) the spread-adjusted return.

[22] We implement the adjustment prior to eliminating the intraday seasonality and attach the trading costs to the raw returns.

[23] The distribution of ${}_{\text{adj}}r_{i,j}^{k}$ integrates the dependence structure between the stocks considered.

allowing for competitive and efficient trading. Among the market participants, there are over 30 market makers (dealers) providing liquidity by always being willing to trade at their bid and ask prices posted into the NASDAQ network at all times.[24] LOBSTER is an online limit order book generating tool providing NASDAQ order book data developed at Humboldt University, Berlin. It reconstructs the full limit order book on the basis of raw message data (i.e. orders as they were submitted by market participants and recorded by the market organizers) and yields an efficient matching algorithm.

LOBSTER is connected to a storage facility containing over 5 TB of NASDAQ's ITCH messages which only include limit order messages, whereas other messages such as imbalance data events and administrative messages have been cleaned out. For arriving limit order submissions, it records order ID, limit price, quantity and trade direction in the order pool; for arriving cancellation or execution messages, the system firstly finds the corresponding limit order submission by comparing the order IDs, then records the remaining non-executed size (or deletes the order from the pool in case of a remaining size equal to zero), and finally updates the limit order book by changing the size and price on the associated ask and bid price level.[25]

In our empirical study we apply the econometric models presented in the previous section to five stocks selected from the NASDAQ 100 for the period between January and February 2009.[26] Our preliminary diagnostics on the correlation and the tail dependence between the stocks' liquidity and log returns are based on the full year of 2009.

We include the intraday data of Alexion (NASDAQ ticker symbol ALXN), Amazon (AMZN), Apple (AAPL), Baidu (BIDU) and Google (GOOG) in our sample. The selected companies allow us to model a portfolio consisting of competing firms (e.g. Baidu and Google) as well as companies possibly offering ample opportunity to diversify (Alexion and Apple).

[24] Hence, the NASDAQ Stock Market is organized as a multiple dealers market with a quote-driven system, where the order flow is fragmented across these dealers (see Chan et al. (1995)).

[25] For a more detailed description see the Technical Report on http://lobster.wiwi.hu-berlin.de/Lobster/.

[26] The NASDAQ 100 includes 100 of the largest non-financial companies listed on the NASDAQ.

In order to yield equidistantly time-spread observations, we re-sample the non-regularly time-spaced data along a pre-specified time grid, where we consider the realized bid-ask spreads and mid price returns closest to the observation points of the grid. We assume the distance between two observation points to equal five minutes and thus compute spreads and returns as end points of five minute long intervals.[27] Furthermore, we exclude the pre-market and after-market trading hours as well as the first and last three five-minute intervals of the regular market hours, because, due to the characteristics and specific conditions on the NASDAQ, these periods possess different dynamics and could bias our estimates of intraday seasonality. Hence, we consider the period from 9:45 a.m. to 3:45 p.m. and have 71 observation points per day.[28]

For the preliminary analysis of illiquidity-return commonality, we employ the data for the full year of 2009 yielding a sample size of $17,537$ intraday observations at five minute intervals. For the estimation and forecasting of L-IVaR, we use the first 20 trading days in January 2009 as our in-sample ($1,420$ observations) and the following 5 trading days (355 observations) as our out-of-sample.

We construct an equally weighted portfolio and model the five spreads and the five log returns of the stocks considered. Each spread and each return is firstly modeled univariately according to the models in 2.1 and 2.2, respectively. In the next step we model the multivariate dependence structure of this portfolio and estimate a 10-dimensional vine copula model via Maximum-Likelihood with the pair-copulas being selected using AIC as the selection criterion. Finally, we calculate forecasts for the liquidity-adjusted intraday VaR based on simulated portfolio values.

Descriptive statistics on the data sample used in our risk management application are presented in Table 1.

— insert Table 1 here —

[27]Regarding the returns this means that we choose the mid prices closest to the observation points of the grid and calculate the returns on the basis of these five-minute mid prices.

[28]Note that we need two adjacent observation points for the computation of the returns so that we get 71 observation out of 72 five-minute intervals.

Histograms of the bid-ask spreads and log mid price returns of all five stocks in our sample are presented in Figure 2.

— insert Figure 2 here —

The descriptive statistics for the bid-ask spreads given in Panel (a) of Table 1 show that the mean liquidity/illiquidity varies considerably across the five companies in our sample portfolio with mean spreads ranging from 2.2317 cents (Apple) to 25.6606 cents (Google). Furthermore, the upper panels in Figure 2 underline the notion that all spreads exhibit skewed empirical distributions. Additionally, we can see from Panel (b) in Table 1 that the five stocks possess the usual stylized characteristics of negligible mean log returns and non-normally distributed returns with a weakly skewed distribution (see also the lower panels in Figure 2).

In the next section, we try to explore in greater detail the relationship between liquidity and asset returns across the companies in our sample to provide anecdotal evidence of cross-market liquidity-return comovements.

3.2 Anecdotal evidence of liquidity-return commonality

As a simple first step, we start our empirical investigation by reporting anecdotal evidence on the relationship between liquidity and asset returns. Here, we are especially interested in documenting the dependence between the log returns on mid prices and bid-ask spreads of stocks traded on the NASDAQ. As a preliminary analysis, we plot the illiquidity against the returns on selected stocks in our sample to document the systematic relationship between the two.

— insert Figure 3 here —

Figure 3 illustrates the time variation in the log returns and bid-ask spreads of all firms in our sample. The figure depicts time series plots of the bid-ask spreads in the upper

panels (a-c) and (g-i) and corresponding plots for the log returns in the lower panels (d-f) and (j-l). To illustrate the trend of both the spreads and returns, the intraday data measured at five-minute intervals are averaged over 50 interval windows yielding trend lines highlighted in color in Figure 3. The time series plots show that, in some cases, spikes in the bid-ask spreads coincide with negative returns thus supporting the common finding of Chordia et al. (2000), Amihud (2002) and Acharya and Pedersen (2005) that liquidity comoves with contemporaneous returns. Furthermore, bid-ask spreads seem to comove as well, as can be seen from panels (a,c,h,i) around intervals 800-900. Although these first results constitute only anecdotal evidence, we can already see that the joint distribution of spreads and returns could be systematically characterized by extreme liquidity commonality, spread-return comovements as well as extreme comovements of stock returns.

Further exploring this conjecture, we estimate the linear (unconditional) correlations between returns and spreads for the 5 companies in our sample which are later used in our risk management application based on our full sample of $17,537$ observations for the full year of 2009. The results are given in Table 2.

— insert Table 2 here —

The three panels in Table 2 present the correlation coefficients between bid-ask spreads, log returns and correlations between returns and spreads, respectively. Panel (a) in Table 2 provides strong evidence of a commonality in liquidity between the five companies. With the exception of the correlations between the pharmaceutical company Alexion and the IT companies Apple and Baidu, all pairs of companies exhibit strong positive correlations between their bid-ask spreads (ranging from 0.0434 to 0.6029). Confirming the second stylized fact on financial asset returns, Panel (b) underlines the common finding that stock returns are positively correlated with results ranging from 0.2069 to 0.6698 (Apple vs. Google). These extremely high positive correlations are not surprising, however, as our sample period covers the onset and climax of the recent financial crisis in 2009. Consequently, our results

are symptomatic of the often cited "correlation breakdown" during times of financial market turmoil.

Turning to the correlations between liquidity and returns, Panel (c) in Table 2 weakly supports our first impression from Figure 3 that liquidity comoves with contemporaneous returns as the spreads and returns of the *same* company (main diagonal in bold type) are weakly negatively correlated. More interestingly, Panel (c) presents evidence for a weak correlation between the spreads and returns of *different* companies. Again, we can see that liquidity comoves with returns across firms for most firm-pairs with spread-return correlations ranging from 9.24% (Alexion-Google) to −16.75% (Apple-Baidu).

The correlation analysis in Table 2 stresses the fact that the dependence between liquidity and returns needs to be taken into account when forecasting liquidity-adjusted measures of risk by means of a joint distributional model. However, non-linear and extreme dependence between spreads and returns could be present in addition to the found linear dependence thus possibly requiring the use of flexible copula models (instead of a purely Gaussian model). We therefore complement our correlation analysis by estimating the lower and upper tail dependence between the bid-ask spreads and returns of the companies in our sample using a nonparametric estimator (see Schmidt and Stadtmüller, 2006, for details of the estimator and its asymptotic properties).[29] Results on the tail dependence between spreads and returns are presented in Tables 3 and 4.

— insert Tables 3 and 4 here —

As can be seen from Panels (a) and (b) in Tables 3 and 4, both the bid-ask spreads as well as the log returns of the stocks in our sample are characterized by strong upper and lower tail dependence. Hence, liquidity commonality is not only manifested in a linear correlation between bid-ask spreads and returns, but also in strong tail dependence between the liquidity of different assets.

[29] By using a nonparametric estimator, we circumvent the error-prone problem of choosing a parametric copula family just for estimating the tail dependence coefficients.

Panels (c) in Tables 3 and 4 report the estimated coefficients of lower and upper tail dependence between the spreads and returns of the firms in our sample. The results provide ample evidence of a strongly nonlinear relationship between liquidity and returns of not only the same asset, but also between the liquidity and returns of different firms' stocks. Interestingly, the upper tail dependence between spreads and returns seems to be more pronounced than the lower tail dependence. It thus seems that extreme increases in illiquidity are accompanied by an extreme rise in the risk premia in contemporaneous returns. Extreme increases in liquidity, on the other hand, do not coincide with extreme downward movements in returns as can be seen from the coefficients of lower tail dependence which are regularly lower than the corresponding coefficients of upper tail dependence.

This finding is also confirmed by an analysis of the time-variation of lower and upper tail dependence. The results of this analysis are presented in Figure 4 which shows the time-variation of both tail dependence coefficients estimated nonparametrically from rolling windows of 300 observations based on our full sample of $17,537$ intraday data points for all five stocks we consider.

— insert Figure 4 here —

The time series plots in Figure 4 show that upper tail dependence (red lines) between the bid-ask spreads and the log returns of the five individual stocks is systematically higher than lower tail dependence (see e.g. Panel (a) in Figure 4). In line with corresponding results on diurnal returns, both lower and upper tail dependence seems to be strongly time-varying.

In summary, both the dependence structure between the spreads and returns of individual stocks as well as the dependence between cross-market illiquidity and returns seem to be highly nonlinear and characterized by a high degree of tail dependence. Even more importantly, from an investor's point of view, neglecting the strong cross-market tail dependence between illiquidity and asset returns could lead to severely biased estimates of portfolio liquidity risk.

Our choice of a flexible vine copula model for the joint distribution of spreads and returns which accounts for possible tail dependence between the variables thus seems to be well justified.

3.3 Results on L-IVaR estimation and forecasting

Using the intraday data for our portfolio consisting of five stocks, we forecast the portfolio L-IVaR according to the three models presented above. The marginals are modeled using the desribed GARCH and ACDP models. As candidate parametric copulas, we use the Gaussian, Student's t, Clayton, Frank, Gumbel, Joe, BB1, BB8, Survival Clayton and Survival BB8 copulas. We employ in-samples with sizes of $1,420$ observations at five minute intervals and forecast the following 23/24 observations.[30] The in-sample is then shifted forward and the next 23/24 observations are forecasted using the reestimated models. From each estimated model, we simulate $1,000$ portfolio returns and bid-ask spreads from which we calculate the portfolio 5%-L-IVaR via Monte Carlo simulation.

Results on the estimated coefficients of the marginal models averaged over all 15 reestimations (5 days $\times$ 3 forecasting periods) are presented in Table 5.

— insert Table 5 here —

The average estimated coefficients show that, on average, all marginal fits include a pronounced autoregressive component as expressed by the parameters ψ_1^k and β_1^k of the ACDP and GARCH model, respectively. Moreover, the average parameter estimates show that both underdispersion (Amazon and Apple) as well as high overdispersion (e.g., Google) is fitted with the ACDP models. In unreported model diagnostics tests, we use the tests of Box and Pierce (1970) and Ljung and Box (1978) to check the hypothesis that there is no autocorrelation left in the standardized residuals of the GARCH marginals. The unreported results for the 15 reestimations show that the hypothesis of no autocorrelation cannot be

[30] Our out-of-sample consists of five days with 71 intraday observations each. Each day is split up into three forecasting periods of length 24 and 23 for the last period on that day, respectively.

rejected at the 5% level for any of the fitted marginal GARCH models. Corresponding results for Ljung-Box and Box-Pierce tests on the Pearson residuals $\frac{S^k_{i,j}-\lambda^k_{i,j}}{\sqrt{\frac{\lambda^k_{i,j}}{\gamma_k}}}$ of the ACDP models (Heinen and Rengifo, 2007) additionally confirm the good fit of the ACDP model for the bid-ask spreads.

Table 6 presents results on the percentages of parametric copula families which were selected as the first 24 of the overall 45 pair-copulas in our 15 reestimations.

— insert Table 6 here —

The percentages given in Table 6 show that the vast majority of pair-copulas are selected from the Gaussian, Clayton and Frank families. Interestingly, the Student's t copula is selected in only few instances strikingly contrasting the usual finding in multivariate market risk modeling that the Student's t copula provides the optimal fit for financial market data (see Breymann et al., 2003). Moreover, Table 6 underlines the notion that the modeling of a high-dimensional portfolio consisting of bid-ask spreads and returns requires the increased flexibility of a vine copula model as most bivariate pair-copulas are chosen from more than one parametric copula family.

Turning to the L-IVaR forecasts, Figure 5 presents a comparison between the realized liquidity-adjusted profits and losses for our 5-dimensional portfolio in the out-of-sample and corresponding simulated L-IVaR forecasts computed by the three models presented above. The three panels in Figure 5 present the comparisons of the L-IVaR forecasts computed from the BDSS model (Panel (a)), the HW model (b) and the L-IVaR computed from adjusted returns (c). Panel (d) combines all three model forecasts and realized P/L in a single plot to highlight the differences between the models' forecasts. The beginning of each of the five days in our our out-of-sample is marked with a dashed vertical line.

The plots in Figure 5 give ample evidence of the three different models' ability to predict the portfolio's L-IVaR. Although the L-IVaR estimates stay relatively close to the realized losses on the portfolio, the L-IVaR forecasts are exceeded in only 3 cases for the BDSS

model, in 4 cases for the HW model and just once for the L-IVaR model based on adjusted returns. These out-of-sample results stress the finding that our multivariate model performs exceptionally well in forecasting L-IVaR. At the same time, our forecasts are not too conservative thus preventing e.g. banks from reserving too much risk capital. Furthermore, it is interesting to note that our L-IVaR models fully capture the increased seasonal volatility of intraday spreads and returns during the first hour of daily trading as shown by the steep decrease in both portfolio losses and L-IVaR forecasts after each turn-of-the-day (dashed vertical lines).

Panel (d) in Figure 5 finally shows that all three different models for forecasting L-IVaR produce comparable results. Results on the number of L-IVaR exceedances show that the BDSS and HW models (3 and 4 exceedances, respectively) do not differ while the L-IVaR model based on liquidity-adjusted returns (1 exceedance) which we use as a benchmark seems to be slightly too conservative.

To formally test the adequacy of these out-of-sample L-IVaR forecasts, we follow the vast majority of studies in quantitative risk management and employ the test of Conditional Coverage proposed by Christoffersen and Pelletier (2004) on the forecasts of our three models.[31] The results show that both the BDSS model (p-value 0.3660, test statistic 0.4195) as well as the HW model (p-value 0.7121, test statistic 0.5135) cannot be rejected at the 10% level. As already seen from the number of L-IVaR exceedances which was too conservative, the model based on adjusted returns (p-value 0.0590, test statistic 0.5081) is rejected at the 10% level and cannot be rejected only at the 5% level.

4 Conclusion

Liquidity risk has rarely received more attention than during the past few years following the recent subprime and sovereign debt crises. In this paper, we propose a multivariate

[31] The test of Conditional Coverage tests for the correct number as well as serial independence of VaR-exceedances. Details of this likelihood ratio test and its properties can be found in Christoffersen and Pelletier (2004). The test's p-values are computed via parametric bootstrap with $5,000$ simulations.

model that incorporates the joint modeling of liquidity as well as market price risk in a flexible copula-based distributional model. We propose to model the joint distribution of bid-ask spreads and log returns of a stock portfolio by using Autoregressive Conditional Double Poisson and GARCH processes for the marginals and vine copulas for the dependence structure. Using intraday data from the NASDAQ, we incorporate the measurement of commonalities in liquidity and comovements of stocks and bid-ask spreads into the forecasting of three types of liquidity-adjusted Value-at-Risk.

To the best knowledge of the authors, this is the first study to document anecdotal evidence of strong extreme comovements in liquidity and strong tail dependence between bid-ask spreads and log returns across firms. Motivated by this finding, we employ an extremely flexible vine copula model to capture the diverse dependence structures between intraday bid-ask spreads and log returns. We then use this multivariate model to forecast the liquidity-adjusted intraday VaR of a five-dimensional portfolio.

Our findings clearly show that the proposed multivariate model for bid-ask spreads and intraday returns performs exceptionally well in forecasting liquidity-adjusted portfolio losses. While losses are adequately bounded below by our liquidity-adjusted VaR forecasts, our models are not too conservative and thus prevent investors like, e.g., banks from reserving unnecessary risk capital buffers. These results are confirmed in a formal backtesting based on the models' conditional coverage.

Further research should concentrate on generalizing the anecdotal evidence of the found nonlinear dependence between liquidity and asset returns. Furthermore, the relationship between liquidity and asset returns of different firms needs to be analyzed in greater detail. The results from our vine copula modeling underline the notion that cross-firm dependence between liquidity and asset returns needs to be taken into account when forecasting liquidity-adjusted VaRs. In a bottom-up approach to VaR modeling, the idiosyncratic and macroeconomic factors driven this form of commonality would need to be identified. We intend to address these questions in future research.

References

Aas, K., Berg D., 2009. Models for construction of multivariate dependence: a comparison study. European Journal of Finance 15, 639-659.

Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance, Mathematics and Economics 44, 182-198.

Acharya, V.V., Pedersen, L.H., 2005. Asset pricing with liquidity risk. Journal of Financial Economics 77, 375-410.

Amihud, Y., Mendelson, H., 1980. Dealership market: Market-making with inventory. Journal of Financial Economics 8, 31-53.

Amihud, Y., Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics 17, 223-249.

Amihud, Y., 2002. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. Journal of Financial Markets 5, 31-56.

Andersen, T., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance 4, 115-158.

Andersen, T., Bollerslev, T., Diebold, F., Labys, P., 2000. Exchange rate returns standardized by realized volatility are (nearly) gaussian. Multinational Finance Journal 4, 159-179.

Andersen, T., Bollerslev, T., Diebold, F., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71, 579-625.

Bangia, A., Diebold, F.X., Schuermann, T., Stroughair, J.D., 2002. Modeling liquidity risk, with implications for traditional market risk measurement and management. In: Figlewski, S., Levich, R.M. (Eds.), Risk Management: The State of the Art, Vol. 8. Kluwer Academic Publishers, Boston.

Bedford, T., Cooke, R.M., 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. Annals of Mathematics and Artificial Intelligence 32, 245-268.

Bedford, T., Cooke, R.M., 2002. Vines - a new graphical model for dependent random variables. Annals of Statistics 30, 1031-1068.

Berkowitz, J., 2000. Incorporating liquidity risk into value-at-risk models. Working paper, University of California.

Box, G.E.P., Pierce, D.A., 1970. Distribution of residual correlations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association 65, 1509-1526.

Brechmann, E.C., Czado, C., 2011. Risk Management with High-Dimensional Vine Copulas: An Analysis of the Euro Stoxx 50. Working paper, Technische Universität München.

Brennan, M.J., Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of Financial Economics 41, 441-464.

Breymann, W., Dias, A., Embrechts, P., 2003. Dependence structures for multivariate high-frequency data in finance. Quantitative Finance 3, 1-14.

Brockwell, A., 2007. Universal residuals: A multivariate transformation. Statistics and Probability Letters 77, 1473-1478.

Cameron, C., Trivedi, P., 1998. Regression Analysis of Count Data. Econometric Society Monograph No.30, Cambridge University Press.

Chan, K.C, Christie, W.G., Schultz, P.H., 1995. Market structure and the intraday pattern of bid-ask spreads for NASDAQ securities. Journal of Business 68, 35-60.

Chollete, L., Heinen, A., Valdesogo, A., 2009. Modeling International Financial Returns with a Multivariate Regime-switching Copula. Journal of Financial Econometrics 7, 437-480.

Chordia, T., Roll, R., Subrahmanyam, A., 2000. Commonality in liquidity. Journal of Financial Economics 56, 3-28.

Christoffersen, P., Pelletier, D., 2004. Backtesting Value-at-Risk: A Duration-Based Approach. Journal of Financial Econometrics 2, 84-108.

Cornett, M.M., McNutt, J.J., Strahan, P.E., Tehranian, H., 2011. Liquidity risk management and credit supply in the financial crisis. Journal of Financial Economics 101, 297-312.

Demsetz, H., 1968. The cost of transacting. Quarterly Journal of Economics 82, 33-53.

Dick-Nielsen, J., Feldhütter, P., Lando, D., 2012. Corporate bond liquidity before and after the onset of the subprime crisis. Journal of Financial Economics 103, 471-492.

Dionne, G., Duchesne, P., Pacurar, M., 2009. Intraday value at risk (IVaR) using tick-by-tick data with application to the Toronto Stock Exchange. Journal of Empirical Finance 16, 777-792.

Dißmann, J., Brechmann, E.C., Czado, C., Kurowicka, D., 2012. Selecting and estimating regular vine copulae and application to financial returns. Working paper, Technische Universität München.

Efron, B., 1986. Double exponential families and their use in generalized linear regression. Journal of the American Statistical Association 81, 709-721.

Engle, R.F., 2000. The econometrics of ultra-high frequency data. Econometrica 68, 1-22.

Engle, R.F., Russell, J.R., 1998. Autoregressive conditional duration: a new model for irregulary spaced transaction data. Econometrica 66, 1127-1162.

Ferland, R., Latour, A., Oraichi, D., 2006. Integer-valued Garch process. Journal of Time Series Analysis 27, 923-943.

Fisher, R.A., 1932. Statistical Methods for Research Workers. Oliver and Boyd, London.

Fokianos, K., Rahbek, A., Tjostheim, D., 2009. Poisson autoregression. Journal of the American Statistical Association 12, 1430-1439.

Giot, P., 2005. Market risk models for intraday data. European Journal of Finance 11, 309-324.

Gourieroux, C., Jasiak, J., 2010. Lokal likelihood density estimation and value at risk. Journal of Probability and Statistics, doi:10.1155/2010/754851.

Groß-Klußmann, A., Hautsch, N., 2011. Predicting bid-ask spreads using long memory autoregressive conditional poisson models. Working paper, Humboldt-Universität zu Berlin.

Hasbrouck, J., Seppi, D., 2001. Common factors in prices, order flows and liquidity. Journal of Financial Economics 59, 383-411

Heinen, A., 2003. Modelling time series count data: An autoregressive conditional poisson model. Working paper, University of California.

Heinen, A., Rengifo, E., 2007. Multivariate autoregressive modeling of time series count data using copulas. Journal of Empirical Finance 14, 564-583.

Heinen, A., Valdesogo, A., 2008. Asymmetric CAPM dependence for large dimensions: The canonical vine autoregressive model. Working paper, Université catholique de Louvain, Center for Operations Research and Econometrics.

Heude, A.F., Wynendaele, P. van, 2001. Integrating Liquidity Risk in a Parametric Intraday VaR Framework. Working paper, Université de Perpignan.

Huberman, G., Halka, D., 2001. Systematic liquidity. Journal of Financial Research 24, 161-178.

Joe, H., 1996. Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. In: Rüschendorf, L., Schweizer, B., Taylor, M.D. (Eds.), Distributions with Fixed Marginals and Related Topics, Institute of Mathematical Statistics, Hayward

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman and Hall, London.

Karolyi, G.A., Lee, K.-H., van Dijk, M.A., 2012. Understanding commonality in liquidity around the world. Journal of Financial Economics, in press.

Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315-1335.

Li, L., Poon, S.-H., 2011. Using Intraday Data for Estimation of Multivariate Dependence. Working paper, University of Konstanz and University of Manchester.

Liu, W., 2006. A liquidity-augmented capital asset pricing model. Journal of Financial Economics 82, 631-671.

Ljung, G.M., Box, G.E.P., 1978. On a measure of lack of fit in time series models. Biometrika 65, 297-303.

Longin, F., Solnik, B., 1995. Is the international correlation of equity returns constant: 1960-1990?. Journal of International Money and Finance 14, 3-26.

MacDonald, I., Zucchini, W., 1997. Hidden Markov and other models for discrete-valued time series. Chapman and Hall, London.

McKenzie, E., 2003. Discrete variate time series. Handbook of Statistics 21, 573-606.

Min, A., Czado, A., 2010. Bayesian Inference for Multivariate Copulas using Pair-copula Constructions. Journal of Financial Econometrics 8, 511-546.

Min, A., Czado, A., 2011. Bayesian model selection for D-vine pair-copula constructions. Canadian Journal of Statistics 39, 239-258.

Nikoloulopoulos, A., Joe, H., Li, H., 2012. Vine copulas with asymmetric tail dependence and applications to financial return data. Computational Statistics and Data Analysis, in press.

Nolte, I., 2008. Modeling a multivariate transaction process. Journal of Financial Econometrics 6, 143-170.

Qi, J., Ng, W.L., 2009. Liquidity adjusted intraday value at risk. Working paper, University of Essex.

Rydberg, T.H., Shephard, N., 1999. BIN models for trade-by-trade data. Modeling the number of trades in a fixed interval of time. Working paper, Nuffield College.

Sadka, R., 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. Journal of Financial Economics 80, 35-55.

Schmidt, R., Stadtmüller, U. 2006. Nonparametric estimation of tail dependence. Scandinavian Journal of Statistics 33, 307-335.

Stoll, H., 1978. The supply of dealer services in securities markets. Journal of Finance 33, 1133-1151.

Whelan, N., 2004. Sampling from Archimedean copulas. Quantitative Finance 4, 339-352.

Figures and Tables

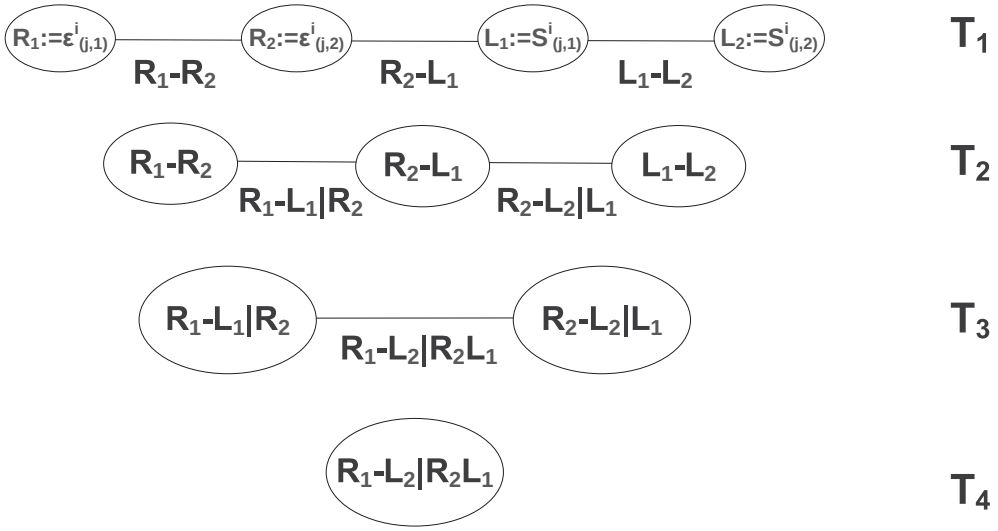

Figure 1: **Four-dimensional example of a D-vine**. The variables represent the intraday bid-ask spreads and the mid-price returns of two stock, respectively. Each edge in layers T1-T3 (nodes in layers T2-T4) corresponds to a bivariate pair-copula.

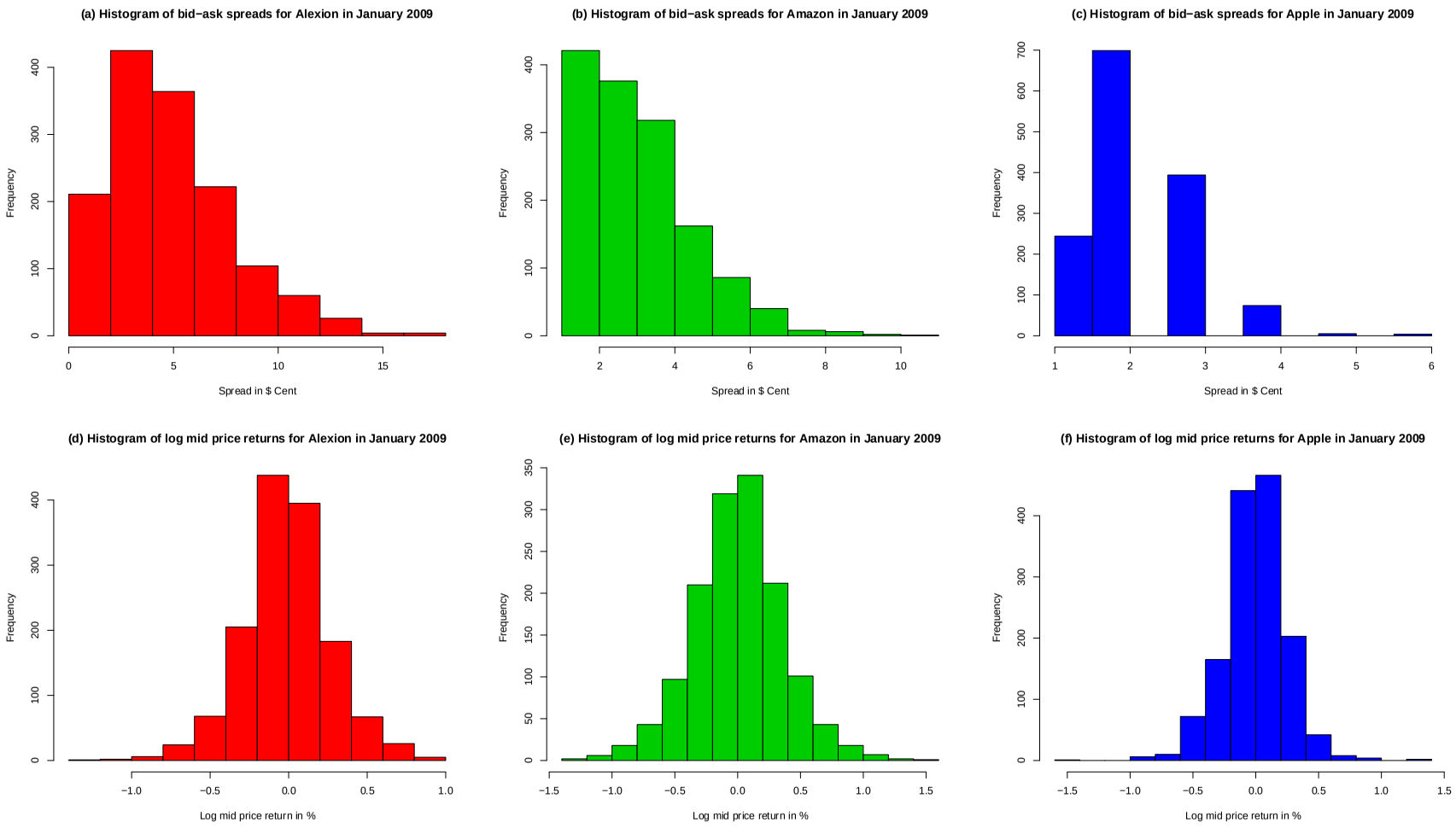

Figure 2: **Histograms of bid-ask spreads and log mid price returns**. The figure shows histograms of the bid-ask spreads and log mid price returns of the five stocks based on our first in-sample of $1,420$ intraday observations measured at five-minute intervals. The sample period is January 1, 2009 - February 5, 2009 and taken from the LOBSTER database containing tick-by-tick data for all companies listed on the NASDAQ.

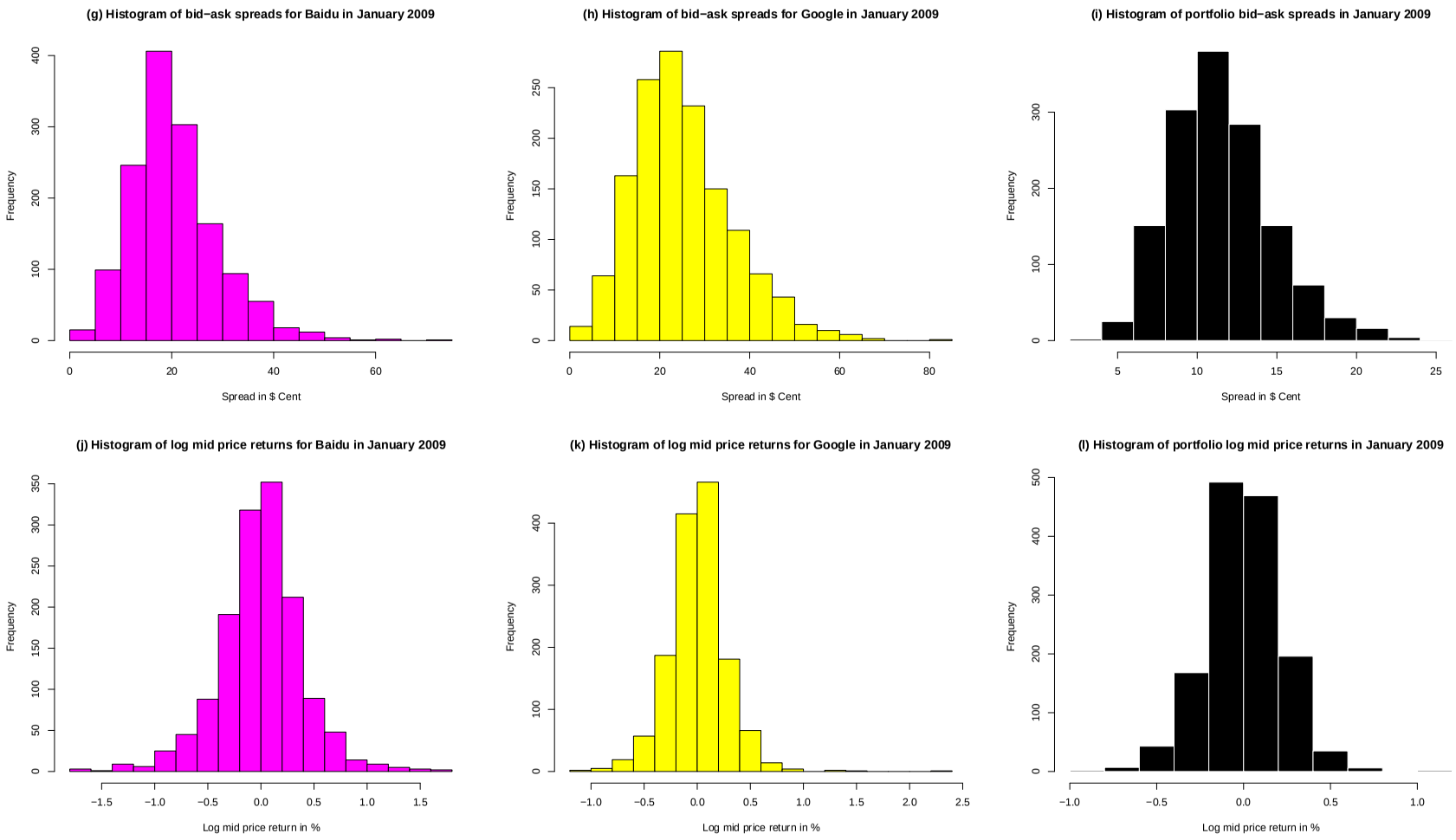

Figure 2: **Histograms of bid-ask spreads and log mid price returns (continued).**

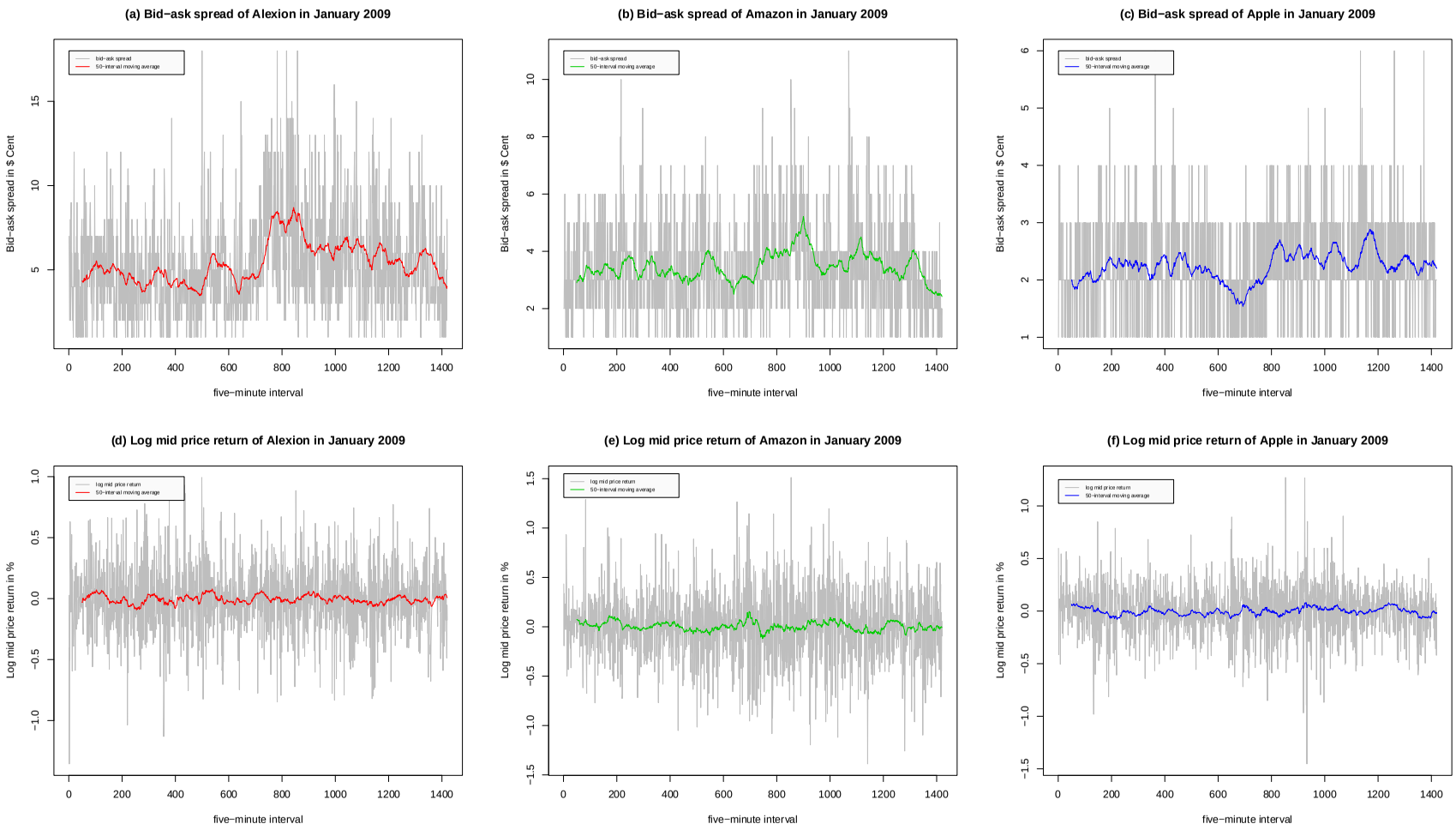

Figure 3: **Bid-ask spreads and log mid price returns**. The figure shows the bid-ask spreads and log mid price returns of the five stocks based on our first in-sample of $1,420$ intraday observations measured at five-minute intervals. Moving averages estimated from 50 intervals are printed in color. The sample period is January 1, 2009 - February 5, 2009 and taken from the LOBSTER database containing tick-by-tick data for all companies listed on the NASDAQ.

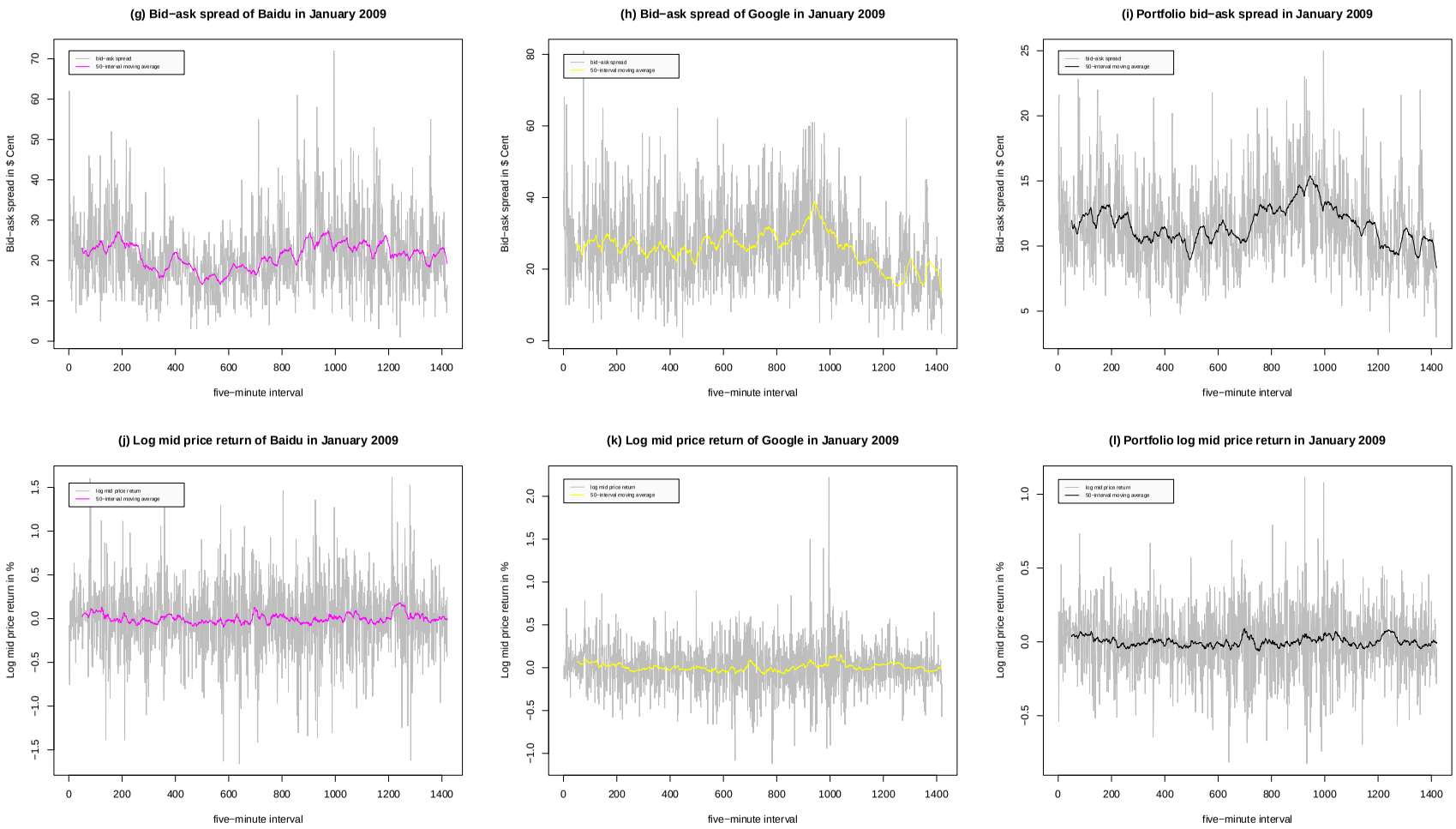

Figure 3: **Bid-ask spreads and log mid price returns (continued).**

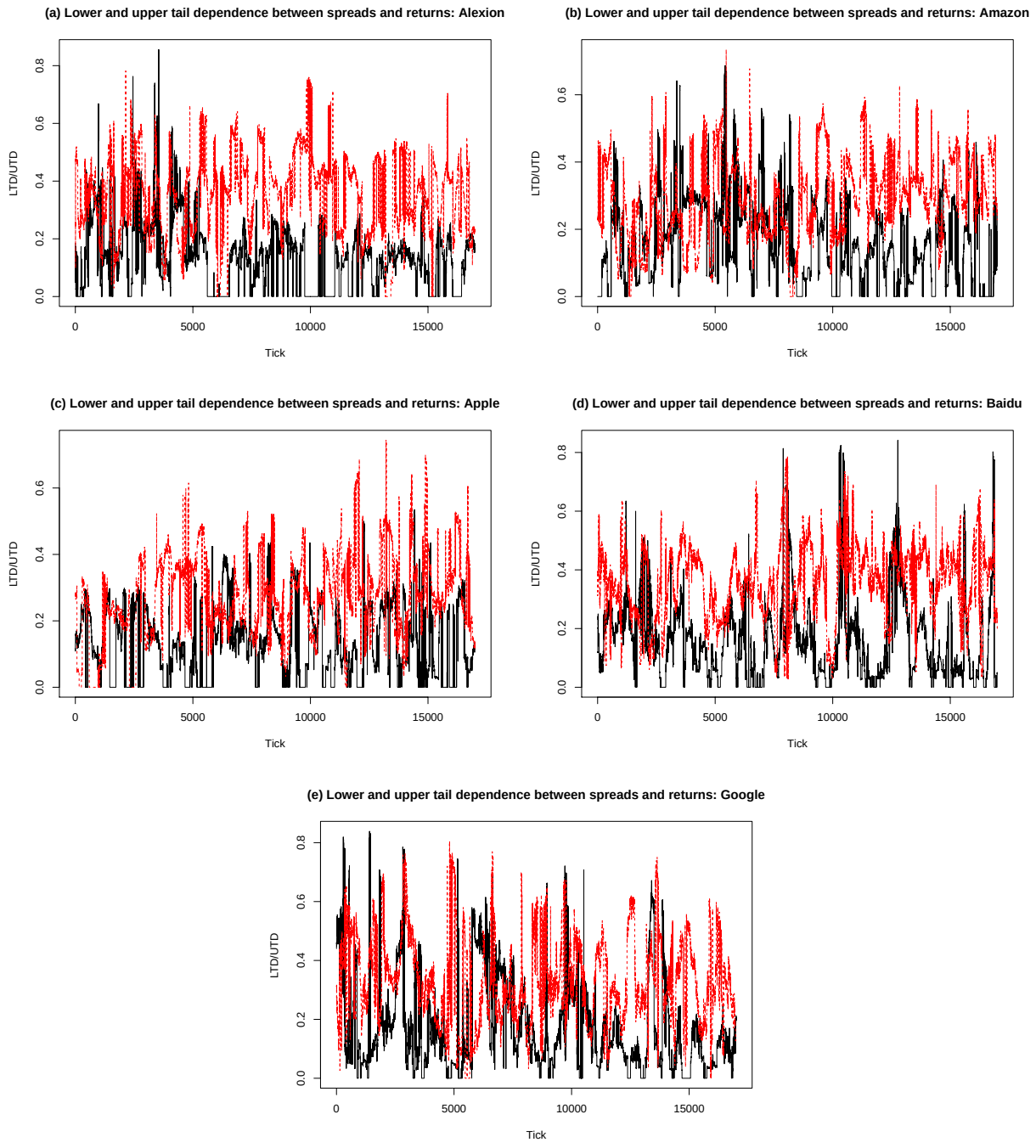

Figure 4: **Time evolution of the tail dependence between spreads and returns**. The figure shows the time evolution of the coefficients of lower (black line) and upper tail dependence (red line) between the intraday bid-ask spreads and the log returns computed from the intraday mid prices. The coefficients of lower and upper tail dependence are estimated from rolling windows of 300 data points using the nonparametric estimator of Schmidt and Stadtmüller (2006). The sample period is January 1, 2009 - December 31, 2009 and taken from the LOBSTER database containing tick-by-tick data for all companies listed on the NASDAQ.

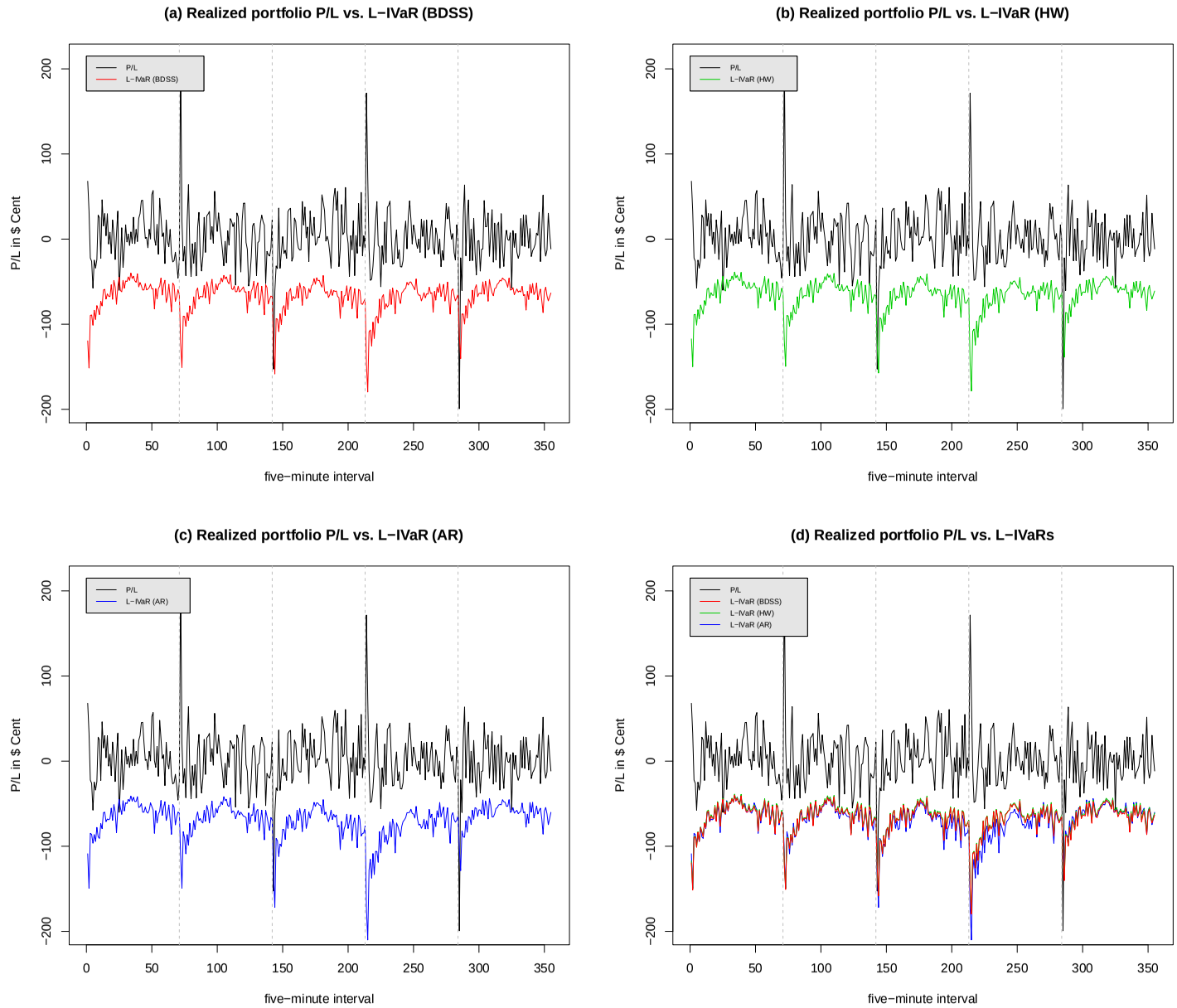

Figure 5: **Realized portfolio losses and VaR estimates**. The figure presents the realized out-of-sample profits and losses (P/L) on our sample portfolio as well as the three different IVaR model estimates (Panels (a-c)). For the ease of comparability, the realized portfolio P/L and the L-IVaR estimates of all three competing models are plotted together in Panel (d). The confidence level for the L-IVaR estimates is 5%. The L-IVaR ist estimated from in-samples of length $1,440$ in order to forecast the following out-of-sample of length 24. All models are then reestimated with both the in- and out-of-sample being shifted forward by 24 observations (5 min intervals). The complete out-of-sample covers 355 intraday observations. Each turn-of-the-day of the five days in our out-of-sample is marked with a dashed vertical line.

	Alexion	Amazon	Apple	Baidu	Google	Portfolio
Panel (a): Descriptive statistics of spreads (in US-Dollar cents)						
number of observations	1420	1420	1420	1420	1420	1420
minimum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
maximum	18.0000	11.0000	6.0000	72.0000	81.0000	25.0000
mean	5.3831	3.4345	2.2317	21.1500	25.6606	11.5719
variance	8.6861	2.4460	0.6813	73.8993	122.6417	10.0123
median	5.0000	3.0000	2.0000	20.0000	24.0000	11.2000
95% quantile	11.0000	6.0000	4.0000	37.0000	46.0000	17.2000
Jarque-Bera	280.6177	180.2079	114.4241	535.4775	172.756	121.4738
(p-value)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Panel (b): Descriptive statistics of log returns (in %)						
number of observations	1420	1420	1420	1420	1420	1420
minimum	-1.3559	-1.3888	-1.4533	-1.6586	-1.1209	-0.8256
maximum	0.9934	1.5123	1.2706	1.6152	2.2214	1.1145
mean	0.0071	0.0062	-0.0009	0.0009	0.0079	0.0014
variance	0.0008	0.0013	0.0006	0.0016	0.0007	0.0005
median	0.0000	0.010	0.0084	0.0110	0.0115	-0.0014
5% quantile	-0.4582	-0.5980	-0.4289	-0.6490	-0.4231	-0.3672
Jarque-Bera	84.2009	39.8320	330.7854	241.8356	1517.205	105.7776
(p-value)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Table 1: **Descriptive statistics of spreads and log returns.** The table presents descriptive statistics on the five-minute bid-ask spreads (in US-Dollar cents) and the log mid price returns (in %) used for estimating and forecasting L-IVaR. The sample period is January 1, 2009 - February 5, 2009 and taken from the LOBSTER database containing tick-by-tick data for all companies listed on the NASDAQ.

	Alexion	Amazon	Apple	Baidu	Google
Panel (a): Bravais-Pearson correlations - bid-ask spreads					
Alexion	1	0.2389	-0.4670	-0.4807	0.2704
Amazon		1	0.0434	0.1392	0.3298
Apple			1	0.6029	0.3252
Baidu				1	0.1857
Google					1
Panel (b): Bravais-Pearson correlations - log returns					
Alexion	1	0.2434	0.3575	0.2069	0.2503
Amazon		1	0.5607	0.5310	0.6194
Apple			1	0.5273	0.6698
Baidu				1	0.5280
Google					1
Panel (c): Correlations - bid-ask spreads vs. log returns	Alexion(SP)	Amazon(SP)	Apple(SP)	Baidu(SP)	Google(SP)
Alexion(RET)	**-0.0119**	0.0256	0.0631	0.0424	0.0924
Amazon(RET)	-0.0439	**-0.0559**	-0.0422	0.0834	-0.0666
Apple(RET)	0.0014	-0.1194	**-0.1852**	-0.0846	-0.1025
Baidu(RET)	0.0190	-0.1195	-0.1675	**-0.0669**	-0.1157
Google(RET)	-0.0314	-0.0474	-0.1207	-0.0405	**0.0075**

Table 2: **Linear correlations of bid-ask spreads and log returns.** The table presents the unconditional linear correlation coefficients estimated from the daily average bid-ask spreads and log returns computed from mid prices. The daily averages are computed from the 71 daily intraday measurements of bid-ask spreads and mid prices at 5 min intervals. The abbreviations (SP) and (RET) denote bid-ask spreads and log returns, respectively. Correlation coefficients between the spreads and returns of the same company (main diagonal) in Panel (c) are printed in bold type. The sample period is January 1, 2009 - December 31, 2009 and taken from the LOBSTER database containing tick-by-tick data for all companies listed on the NASDAQ.

	Alexion	Amazon	Apple	Baidu	Google
Panel (a): Lower tail dependence - bid-ask spreads					
Alexion	1	0.7616	0.0000	0.0000	0.0938
Amazon		1	0.2647	0.2000	0.7857
Apple			1	0.2857	0.5204
Baidu				1	0.3247
Google					1
Panel (b): Lower tail dependence - log returns					
Alexion	1	0.5283	0.0000	0.4762	0.5000
Amazon		1	0.0000	0.0000	0.5002
Apple			1	0.8156	0.8315
Baidu				1	0.7632
Google					1
Panel (c): Lower tail dependence - spreads vs. returns	Alexion(SP)	Amazon(SP)	Apple(SP)	Baidu(SP)	Google(SP)
Alexion(RET)	**0.0004**	0.0385	0.0000	0.0370	0.0000
Amazon(RET)	0.0455	**0.1111**	0.0435	0.0000	0.0526
Apple(RET)	0.4444	0.3056	**0.2740**	0.3056	0.3165
Baidu(RET)	0.4595	0.3587	0.2537	**0.2131**	0.2821
Google(RET)	0.2466	0.1176	0.0000	0.1364	**0.0021**

Table 3: **Lower tail dependence between bid-ask spreads and log returns.** The table presents the coefficients of lower tail dependence estimated nonparametrically from the daily average bid-ask spreads and log returns computed from mid prices. The daily averages are computed from the 71 daily intraday measurements of bid-ask spreads and mid prices at 5 min intervals. The abbreviations (SP) and (RET) denote bid-ask spreads and log returns, respectively. The lower tail dependence between the spreads and returns of the same company (main diagonal) in Panel (c) are printed in bold type. The sample period is January 1, 2009 - December 31, 2009 and taken from the LOBSTER database containing tick-by-tick data for all companies listed on the NASDAQ.

	Alexion	Amazon	Apple	Baidu	Google
Panel (a): Upper tail dependence - bid-ask spreads					
Alexion	1	0.3438	0.0000	0.0000	0.4831
Amazon		1	0.3956	0.3188	0.7966
Apple			1	0.6535	0.4000
Baidu				1	0.4815
Google					1
Panel (b): Upper tail dependence - log returns					
Alexion	1	0.5351	0.4787	0.4141	0.5738
Amazon		1	0.7063	0.6457	0.0000
Apple			1	0.8298	0.3333
Baidu				1	0.8519
Google					1
Panel (c): Upper tail dependence - spreads vs. returns	Alexion(SP)	Amazon(SP)	Apple(SP)	Baidu(SP)	Google(SP)
Alexion(RET)	**0.2917**	0.3418	0.2000	0.2500	0.3108
Amazon(RET)	0.4118	**0.2381**	0.3176	0.0000	0.3373
Apple(RET)	0.5124	0.2973	**0.0000**	0.0645	0.3776
Baidu(RET)	0.4149	0.6516	0.4082	**0.2308**	0.3108
Google(RET)	0.4737	0.2794	0.0714	0.0000	**0.3690**

Table 4: **Upper tail dependence between bid-ask spreads and log returns.** The table presents the coefficients of upper tail dependence estimated nonparametrically from the daily average bid-ask spreads and log returns computed from mid prices. The daily averages are computed from the 71 daily intraday measurements of bid-ask spreads and mid prices at 5 min intervals. The abbreviations (SP) and (RET) denote bid-ask spreads and log returns, respectively. The upper tail dependence between the spreads and returns of the same company (main diagonal) in Panel (c) are printed in bold type. The sample period is January 1, 2009 - December 31, 2009 and taken from the LOBSTER database containing tick-by-tick data for all companies listed on the NASDAQ.

Parameter	Alexion	Amazon	Apple	Baidu	Google
Panel (a): Mean of estimates for the ACDP(1,1) models					
$\boldsymbol{\gamma_k}$	0.7002	1.5352	3.3679	0.3203	0.2553
$\boldsymbol{c}$	0.3184	0.2784	0.1395	2.5031	0.7898
$\boldsymbol{\phi_1^k}$	0.1284	0.1367	0.0820	0.1879	0.1675
$\boldsymbol{\psi_1^k}$	0.8133	0.7816	0.8563	0.6927	0.7999
Panel (b): Mean of estimates for the GARCH(1,1) models					
$\boldsymbol{\mu_k}$	-0.0175	0.0191	0.0078	0.0031	0.0233
$\boldsymbol{\omega_k}$	0.0847	0.0188	0.0095	0.0288	0.0119
$\boldsymbol{\alpha_1^k}$	0.0258	0.0369	0.0286	0.0291	0.0316
$\boldsymbol{\beta_1^k}$	0.8903	0.9443	0.9617	0.9422	0.9564

Table 5: **Mean parameter estimates for the ACDP and GARCH model.** The table presents the mean estimates of the marginal spread and return models for each stock in the portfolio, where the means result from averaging the parameter estimates over 15 reestimations. Panel (a) shows the results for the spread models (given by ACDP(1,1)) and panel (b) yields the outcomes for the return models (given by GARCH(1,1)). In the ACDP(1,1) we have $S_{i,j}^k|\mathcal{F}_{i,j-1}^k \sim \text{DPois}(\lambda_{i,j}^k, \gamma_k)$, where the mean equation is given by $\lambda_{i,j}^k = c_k + \phi_1^k S_{i,j}^k + \psi_1^k \lambda_{i,j}^k$; hence, c_k, ϕ^k and ψ^k characterize the autoregressive mean and the γ_k parameter accounts for overdispersion ($\gamma_k < 1$) as well as underdispersion ($\gamma_k > 1$) in the conditional (and unconditional) distribution. The GARCH(1,1) is used to model the deseasonalized log mid price returns and given by $r_{i,j}^k = \mu_k + \varepsilon_{i,j}^k \sqrt{h_{i,j}^k}$, where $r_{i,j}^k = \frac{R_{i,j}^k}{\sqrt{\xi_j^k}}$ and $h_{i,j}^k = \omega_k + \alpha_1^k \left(\varepsilon_{i,j-1}^k\right)^2 h_{i,j-1}^k + \beta_1^k h_{i,j-1}^k$; thus, μ_k yields the mean of the (deseasonalized) return and the ω_k, α_1^k and β_1^k parameters characterize its conditional variance.

Pair		*Conditioning variables*		*Parametric copula families*							
				Gaussian	t	Clayton	Frank	Gumbel	Joe	Surv. Clayt.	Surv. BB8
ALXN(S)	AMZN(S)			**60%**	0%	**13%**	**13%**	**7%**	0%	**7%**	0%
AMZN(S)	AAPL(S)			**60%**	0%	**13%**	**20%**	0%	0%	0%	0%
AAPL(S)	BIDU(S)			**33%**	**7%**	**33%**	**7%**	0%	0%	0%	0%
BIDU(S)	GOOG(S)			**20%**	**7%**	**20%**	**20%**	**13%**	0%	**7%**	0%
GOOG(S)	ALXN(R)			**20%**	0%	**7%**	**33%**	**7%**	0%	0%	0%
ALXN(R)	AMZN(R)			**20%**	0%	**7%**	**47%**	**7%**	0%	**7%**	0%
AMZN(R)	AAPL(R)			**27%**	0%	**13%**	**53%**	0%	0%	0%	0%
AAPL(R)	BIDU(R)			**27%**	0%	**7%**	**47%**	**7%**	0%	**7%**	0%
BIDU(R)	GOOG(R)			**7%**	0%	**27%**	**47%**	**7%**	**7%**	0%	0%
ALXN(S)	AAPL(S)	AMZN(S)		**20%**	0%	**27%**	**40%**	0%	0%	**7%**	0%
AMZN(S)	BIDU(S)	AAPL(S)		**40%**	**7%**	**7%**	**33%**	0%	0%	0%	0%
AAPL(S)	GOOG(S)	BIDU(S)		**33%**	0%	**20%**	**27%**	0%	0%	**7%**	0%
BIDU(S)	ALXN(R)	GOOG(S)		**27%**	0%	**13%**	**47%**	0%	0%	0%	0%
GOOG(S)	AMZN(R)	ALXN(R)		**40%**	0%	0%	**27%**	0%	0%	**13%**	0%
ALXN(R)	AAPL(R)	AMZN(R)		**47%**	0%	**13%**	**27%**	0%	0%	0%	0%
AMZN(R)	BIDU(R)	AAPL(R)		**33%**	0%	**27%**	**40%**	0%	0%	0%	0%
AAPL(R)	GOOG(R)	BIDU(R)		**40%**	0%	**7%**	**27%**	0%	0%	**13%**	0%
ALXN(S)	BIDU(S)	AMZN(S)	AAPL(S)	**40%**	0%	**13%**	**27%**	**7%**	0%	0%	0%
AMZN(S)	GOOG(S)	AAPL(S)	BIDU(S)	**47%**	0%	**27%**	**20%**	0%	0%	0%	0%
AAPL(S)	ALXN(R)	BIDU(S)	GOOG(S)	**27%**	0%	**13%**	**27%**	**7%**	0%	0%	0%
BIDU(S)	AMZN(R)	GOOG(S)	ALXN(R)	**27%**	0%	**13%**	**33%**	**7%**	**7%**	0%	0%
GOOG(S)	AAPL(R)	ALXN(R)	AMZN(R)	**20%**	**7%**	**13%**	**47%**	0%	0%	**13%**	0%
ALXN(R)	BIDU(R)	AMZN(R)	AAPL(R)	**27%**	0%	**20%**	**27%**	**7%**	**7%**	**7%**	0%
AMZN(R)	GOOG(R)	AAPL(R)	BIDU(R)	**27%**	0%	**13%**	**20%**	**20%**	**7%**	0%	0%

Table 6: **Selected parametric pair-copulas.** The table presents average results on the selected bivariate parametric pair-copulas for the first 24 pair-copulas conditional on up to two random variables. The results are given in per cent of the 15 reestimations used for forecasting the L-IVaR at 355 intraday intervals. The copulas used in the estimation include the Gaussian, Student's t, Clayton, Frank, Gumbel, Joe, Survival Clayton and Surival BB8 parametric copula families which are presented and other parametric families with negligible percentages not shown (e.g., BB1, BB8 copulas). The parametric copulas are selected based on a sequential selection using the (bivariate) Akaike's Information Criterion (AIC) as the selection criterion to be minimized (see Dißmann et al., 2012, for details of this procedure). Missing percentages to one are due to rounding or to negligible results for the BB1 and BB8 copula families. Spreads are denoted with (S), returns are denoted with (R).