Overnight News and Daily Equity Trading Risk Limits

Katja Ahoniemi[a,*], Ana-Maria Fuertes[b]
Jose Olmo[c]

[a] *Aalto University School of Economics, Finland*
[b] *Faculty of Finance, Cass Business School, City University London, U.K.*
[c] *Centro Universitario de la Defensa, Zaragoza, Spain*

Preliminary version: March 21, 2012 – This version: May 21, 2012.

Abstract

This paper compares different approaches for incorporating overnight information into one-day-ahead equity trading Value-at-Risk (VaR) limits. We develop the continuous-time theory behind a novel bivariate modeling approach where conditional mean and (co)variance forecasts are constructed separately for the daytime and overnight returns. An alternative, widely-used approach consists of bundling together the overnight squared return and the open-to-close realized volatility. In a third approach, the VaR limits are generated at the market open, when the overnight return is known and can be included in the conditioning information set. Our empirical application endorses the last approach for the S&P 500 index, and bivariate modeling for the Russell 2000 index. The contrasting findings for the two indices shed light on the extent to which the open price efficiently reflects news accumulated during non-trading hours. Price discovery at the open is less efficient for small capitalization, thinly traded stocks. [145 *words*]

Keywords: Equity index; Overnight; Price discovery; Realized volatility; Value-at-Risk.

JEL Classification: C52; C53; G15.

[*] This research was initiated when the first author was visiting Cass Business School in the fall of 2011. We thank Petri Jylha, Simone Manganelli and seminar participants at Leeds University Business School, Durham Business School, and the Helsinki Centre for Economic Research for useful comments. Katja Ahoniemi acknowledges the OP-Pohjola Group Research Foundation and Jose Olmo the Spanish Government through project MICINN ECO2011-22650 for financial support. Corresponding author: e-mail: a.fuertes@city.ac.uk. T:+44 (0)20 7040 0186; F: +44 (0)20 7040 8881.

Electronic copy available at: http://ssrn.com/abstract=2065017

1 Introduction

The world's major stock exchanges are open for a limited number of hours each day, five days a week. Hence, although investors receive information on a continuous basis, they are able to act immediately on that information for only a part of the day in each of those exchanges. The overnight 'surprise' or close-to-open return reflects information accumulated during non-trading hours, such as trades and prices from stock markets in other time zones, and news releases such as earnings announcements. Existing market microstructure research documents that information flow during trading (daytime) hours is greater than information flow during non-trading (overnight) hours.[1] However, the increasing globalization of securities markets and electronic trading systems (such as ECNs) are likely to add to the importance of overnight information flow, as events from around the globe can trigger investor reactions in all markets.[2] The events that occur during overnight non-trading hours will typically be impounded into prices very rapidly in the morning as an exchange opens for a new trading day.

The primary objective of this study is to improve the forecasting of daily return tail behavior defined in terms of Value-at-Risk (VaR). VaR has become a standard risk management tool for setting day-by-day loss limits of equity trading desks, and is widely employed by banks and hedge funds. Through one-day-ahead VaR forecasts, the risk management team of the bank monitors in real time that the equity (and possibly FOREX, fixed income, and derivatives) trading desk stays within predefined risk limits. Besides modeling techniques, tail risk forecastability also relies on a useful information set. There is considerable evidence from existing studies that the dynamics of the return process differ during non-trading and trading hours (Lockwood and Lin, 1990; Hasbrouck, 1991; Masulis and Ng, 1995; Martens, 2002; Cliff et al., 2008; Andersen et al., 2011). This has been acknowledged in various theoretical finance models (Oldfeld and Rogalski, 1980; Hong and Wang, 2000).[3] Assuming an absence

[1] See George and Hwang (2001), Ronen (1997), Jones et al. (1994), and French and Roll (1986), inter alios.

[2] De Goojier et al. (2009) exploit full information in the intra-day stock price patterns in foreign markets during non-trading hours in a home market to predict the opening of an index in the home market.

[3] In the paper we use the terms 'daytime', 'open-to-close' and 'active' period interchangeably to refer to the regular trading period between 9:30am and 4:00pm EST. Likewise, we use the terms 'overnight', 'close-to-open' and 'inactive' interchangeably to refer to the period outside regular trading hours, from 4:00pm to 9:30am EST.

Electronic copy available at: http://ssrn.com/abstract=2065017

of trading during the overnight period, however, there is no consensus as yet in the literature on how to incorporate the overnight 'surprise' for daily equity tail risk forecasting.

This paper makes two contributions. First, it develops the continuous-time theory to motivate a bivariate modeling and forecasting approach for tail behavior of the daily return that explicitly conceptualizes each 24-hour period as comprising two segments: overnight and daytime. During the overnight segment, equity markets are closed for trading, although news releases do not cease. A trader can, at the close of day $t-1$, regard the volatility of the 24-hour day ahead (from close to close) as comprising two random, but forecastable volatilities pertaining to the overnight and daytime segments, respectively, and a forecastable covariance between overnight and daytime returns. Theoretically, it is possible to argue that market microstructure effects such as *price staleness* and *news spillover* can induce a non-zero covariance between daytime and overnight returns which may be valuable in setting the day t equity VaR limits at the close of day $t-1$. Exploiting this covariance is precluded if the risk manager builds his forecast instead at the open of day t, when the return corresponding to the overnight segment from previous close to open is observed.

Our second main contribution is to compare empirically three VaR forecasting approaches that differ in how they exploit overnight information. The goal is to produce evidence on their relative accuracy in predicting the tail risk behavior of the close-to-close return distribution. The explicit question to deal with is: which approach produces more appropriate daily VaR limits for equity trading desks for assets where no overnight trading is available? In the first approach, called *bivariate*, conditional mean forecasts are derived from an AR model for the daytime (overnight) return which is allowed to depend not only on its own past values, but also on the previously observed overnight (daytime) return. In order to obtain the conditional variance forecasts, we put together three components: the overnight volatility forecast obtained from a GARCH model fitted to overnight returns, the daytime volatility forecast from a long-memory model fitted to the realized variance, and the covariance forecast from a dynamic conditional correlation (DCC) type modeling approach. In a competing approach, referred to as *bundling*, we follow the literature in bundling together the overnight and daytime periods into

single return and realized volatility measures. More specifically, a conventional AR model is fitted to close-to-close returns, and a long-memory model to overnight-adjusted realized variances. In the third approach, called *ex post*, we condition the tail risk forecast on the information available at market open. Since the overnight return is now observed ex post, it can be included as an explanatory variable in the models that forecast the conditional mean and daytime realized variance.[4]

Although the first (bivariate) and third (ex post) approaches described above are novel as such, close versions can be found in the literature. For instance, regarding the bivariate formulation, Andersen et al. (2011) approximate the daytime return variation through a long-memory HAR model, while the overnight return variation is captured by a GARCH-type model. Key differences between the latter study and ours are that Andersen et al. (2011) ascribe a role to discrete jumps (which we assume away, for simplicity), whereas the overnight-intraday covariance is not allowed to play any role. Moreover, the forecasting objective in Andersen et al. (2011) is the second moment (instead of tail risk), and they adopt the sum of the squared overnight return and the realized variance as their proxy for the true volatility. Using this proxy and the Mincer-Zarnowitz R^2 criterion in a volatility forecasting framework, they conclude that there are benefits to separately modeling the daytime and overnight volatility components. Regarding the ex post approach, existing studies predict the variance of the daily return process on the basis of long-memory models fitted to the realized variance from open-to-close (e.g. Wu, 2011; Corsi et al., 2008; Thomakos and Wang, 2003; Andersen et al., 2001), which ignores overnight effects. In contrast, we do not ignore the overnight return (observed at market open) but include it in the conditioning information set for the open-to-close return modeling.

The most commonly adopted approach thus far in the high-frequency finance literature is what we call *bundling* above, namely, modeling the long memory properties of the overnight-adjusted realized variance (RV). In essence, the latter is constructed by bundling the squared overnight return and the sum of open-to-close intraday (e.g. 5-minute) squared returns. Various overnight adjustments have

[4]It is important to note the very short (1 day ahead) forecast horizon for VaR when it is used as a 'real time' risk management control tool by banks. It is possible that a longer (weekly or monthly) forecast horizon may reduce the importance of explicitly modeling the overnight return volatility.

been used which include simply adding the squared close-to-open return as one more summand in the realized variance (Bollerslev et al, 2009; de Pooter et al, 2008; Becker et al, 2007; Martens, 2002; Blair et al, 2001), and weighted aggregation of the squared close-to-open return and the realized variance following Hansen and Lunde's (2005) procedure. Overnight-adjusted RV measures are often modeled through long memory specifications (e.g. Giot and Laurent, 2004; Koopman et al., 2005; Angelidis and Degiannakis, 2008a; Martens et al., 2009; Fuertes and Olmo, 2012). These studies have in common the implicit assumption that the overnight squared return is part of the same process that generates the daytime realized volatility which precludes, by construction, separate modeling efforts for the two segments of the day. The bundling also precludes exploiting any dynamic pattern in the covariance between intraday and overnight returns.

Our empirical application is based on S&P 500 and Russell 2000 index market data over the period November 12, 1997 to September 31, 2011. The use of index data is key to the notion that there can be a forecastable covariance between overnight and daytime returns. The index opening price is unlikely to reflect all overnight news because of an infrequent trading problem. This problem is more pronounced for small cap stocks. Index data lags actual developments, especially at the market open, since it takes some time before each of the constituent stocks begins trading. Thus, the first published index quote of each day will contain numerous stale prices, i.e. transaction prices from the previous day.[5] The fact that the first index quote is stale means that its value does not fully reflect all overnight news. Moreover, overnight news may not be rapidly impounded into prices at the market open since the initial part of the trading day is characterized by very rapid bid-ask oscillations, particularly for small cap stocks, which are noisy and unlikely to be related to fundamental news (Hasbrouck, 2012). Overnight news spill over into the trading day, possibly inducing a non-zero correlation between the intraday and overnight innovations. In other words, some of the overnight news are still reflected

[5] Stoll and Whaley (1990a) document empirically that in the 1980s, it took around five to six minutes for large stocks to open for trading on the NYSE. Spurlin et al. (2008) show that for S&P 500 stocks over the period 1996-2001, the average opening time had fallen to around 3.5 minutes. With the advent of electronic trading, the average time to open in today's markets is likely to be much shorter, but large cap stocks will open faster than small cap stocks.

in the prices observed during the daytime trading period, implying ex ante correlation between the overnight and daytime returns.[6,7]

We compare the equity tail risk forecasts generated by the three aforementioned approaches – bivariate, bundling and ex post – within a dynamic quantile backtesting framework. More specifically, accuracy of the one-day-ahead VaR forecasts is gauged through Engle and Manganelli's (2004) dynamic quantile test and a probit-based version with improved statistical properties. We find that the VaR backtesting results for the Russell 2000 index endorse the bivariate approach. In contrast, for the S&P 500 portfolio the best tail risk forecasts are obtained by waiting to observe the overnight return at the market opening. The bundling approach, which is commonplace in extant literature, appears effectively inferior to the other two approaches introduced in this study. These results are robust to the inclusion and exclusion of the recent financial crisis in the out-of-sample period.

The differences across the two indices, the large cap S&P 500 and the small cap Russell 2000, shed light on price discovery mechanisms. Small-cap stocks are typically more thinly traded and less liquid than large cap stocks. Barclay and Hendershott (2008) show that the opening price conveys more fundamental information for high trading volume stocks. Accordingly, for the Russell 2000, there is less value in observing the opening quote, and more efficient forecasts for day t can be generated at the close of day $t-1$, which allows for modeling the overnight return process and also the covariance of daytime and overnight returns. For the S&P 500, in contrast, it is worthwhile to wait until the market opens, so that the overnight return based on previous close and open prices can be exploited

[6] Various studies deal with the fact that not all equity index constituents traded in NYSE immediately begin trading at the market open. Ahoniemi and Lanne (2012) and Chan et al. (1991) both wait until 5 minutes of trading has elapsed before calculating an overnight return. Masulis and Shivakumar (2002) and Noureldin et al. (2011) wait 15 minutes, and Lin et al. (1994) wait for a full 30 minutes. Stoll and Whaley (1990b) omit the first two 5-minute returns of each trading day from their analysis of index returns, and Hecq et al. (2012) omit the first 5 minutes in their stock-level analysis.

[7] Our VaR modeling approach can be generalized to any market where trading ceases for certain periods, which still applies to the majority of securities markets worldwide, including the U.S. futures market which is closed to trading over weekends. There is a parallel, smaller empirical literature on daily volatility forecasting that focuses on the futures (instead of cash) market for which trading data is also available overnight through electronic trading platforms. Taylor (2007) exploits per-hour returns from 12:00am to 8:30am (E-mini S&P500 futures), 8:30am to 3:15pm (regular S&P 500 futures), and 3:15pm to 12:00am (E-mini S&P 500 futures). Martens (2002) uses intra-day and intra-night S&P 500 futures prices observed at the 5 min and 15 min frequencies, respectively.

in the conditioning information set to produce the day-ahead equity VaR limits.

This study continues as follows. Section 2 sets out the risk management framework of interest whereas Section 3 presents three competing downside tail risk forecasting methods that differ in how the overnight return is treated. Section 4 discusses the framework for VaR forecast evaluation. Section 5 discusses the empirical evidence, and a final section concludes.

2 Risk Management Framework

The one-day 1% Value-at-Risk (VaR) of a given portfolio is the maximum dollar-amount loss that should be exceeded only 1% of the time. Such a measure has been used by banks since the 1990s as a one-dimensional snapshot of the downside risk of the close-to-close return (or profit and loss) distribution. Formally, the daily task of risk managers by which VaR limits are set for separate trading desks (e.g. equity, FX, fixed income) can be described as predicting a quantile of the conditional close-to-close return distribution. Commercial banks routinely calculate the 1-day VaR in order to anticipate the maximum loss over the next day with a 99% probability. Real time risk-monitoring using VaR has the advantage over other approaches (e.g. setting risk limits in the form of notional limits and/or stop-loss limits) that VaR limits are comparable across assets and forward-looking in nature. The widespread use of VaR as an internal measure of risk by banks was formally acknowledged by the Basel Accord which introduced in 1996 a VaR-based capital requirement framework for positions held for trading intent (BCBS, 1996). Subsequently, the so-called 'trading book' in which risk weights to capture market risk are based on banks' own VaR models came to account for a substantial part of banks' total risk-weighted assets.

VaR-based risk limits can be obtained using non-parametric (e.g. historical simulation), semi-parametric (e.g. CAViaR) and parametric (e.g. GARCH) techniques depending on the number of assumptions made on the daily return distribution.[8] The most widely-used approach remains the parametric VaR that hinges on the assumption that returns belong to a location-scale density implying,

[8] For recent surveys on VaR modeling, see Kuester et al. (2006) and Angelidis and Degiannakis (2008b).

in turn, that the VaR is an affine function of the volatility (Giot and Laurent, 2004; Giacomini and Komunjer, 2005; Clements et al., 2008; Brownlees and Gallo, 2010; Fuertes and Olmo, 2012).

The goal is to forecast out-of-sample the α-quantile of the conditional distribution of the daily return process, $VaR_{t,\alpha} \equiv F_t^{-1}(\alpha)$, with $F_t^{-1}(\cdot)$ the inverse of the conditional distribution of r_t given $\Im_{t-1}$, i.e. $P(r_t \leq VaR_{t,\alpha}|\Im_{t-1}) = \alpha$ where $\Im_{t-1}$ denotes the conditioning information set.[9] Our model specification of choice for the daily close-to-close return process is location-scale, which can be written as $r_t = \mu_{t|t-1} + \sqrt{IV_{t|t-1}}\epsilon_t$, with ϵ_t the corresponding standardized innovation. By assuming independence between $IV_{t|t-1}$ and ϵ_t, the VaR prediction at nominal coverage level α can be expressed

$$VaR_{t|t-1,\alpha} = \mu_{t|t-1} + \sqrt{IV_{t|t-1}}F_\epsilon^{-1}(\alpha) \tag{1}$$

where the first term $\mu_{t|t-1}$ is a forecast of the conditional mean of the close-to-close return (or profit and loss) distribution; the second term $IV_{t|t-1}$ is a forecast of the conditional variance of the same process; and $F_\epsilon^{-1}(\alpha)$ is the α-quantile of the skewed Student-t distributed standardized innovation ϵ_t.

3 Overnight Information for VaR-Based Risk Monitoring

The first subsection below presents a novel theoretical framework to account for the role of overnight information in estimating the notional volatility of the daily close-to-close return process. Section 3.2 outlines various model-free estimators that have been widely in the high-frequency literature for measuring ex post the daily integrated variance. Section 3.3 presents an approach for forecasting the daily integrated variance one-day-ahead that directly builds on the framework laid out in Section 3.1. Section 3.4 describes how the various mean and variance forecasts are transformed into VaR forecasts.

[9] Our interest is in long trading positions. For short trading positions one would analyze instead the right tail, i.e. $F_t^{-1}(1-\alpha)$. Commercial banks are required to report VaR at confidence level 99% to regulators but most banks adopt the 95% level for internal backtesting. We consider both $\alpha = \{0.01, 0.05\}$.

3.1 Overnight and Daytime Integrated Variance Processes

The diffusion of the log-price process, asumming as it is usual in the finance literature that it belongs to the class of semimartingales, is characterized by the stochastic differential equation

$$dp_t = \mu_t dt + \sigma_t dW_t \quad 0 \leq t \leq T \tag{2}$$

with σ_t the instantaneous or spot volatility process which is stationary and independent of the Brownian motion W_t, and μ_t the instantaneous deterministic drift term. It is also assumed, as in Hansen and Lunde (2005) inter alios, that discrete jumps have a negligible effect on the price diffusion.

To establish notation, let O_t and C_t represent, respectively, the market opening and closing times of day t. The time period from C_{t-1} to C_t represents the entire 24 hr day t comprising an *overnight* period when the market is closed (from C_{t-1} to O_t) and an official trading period from O_t to C_t, hereafter referred to as *daytime*; likewise, day $t+1$ spans the period from C_t to C_{t+1} and so forth. The daily return is denoted $r_t \equiv p_{C_t} - p_{C_{t-1}}$ where prices are measured (in logarithms) at market closing time, i.e. $p_t \equiv p_{C_t}$. From equation (2) it follows that the *integrated variance* is given by

$$IV_t \equiv V(r_t) = \int_{t-1}^{t} \sigma_s^2 ds \tag{3}$$

which represents the true or notional variability of the daily return process on day t. In order to capture the potentially different dynamics of the overnight and daytime components of the continuous time log-price process, the stochastic differential equation (2) can be written as

$$dp_t = \{\mu_{o,t} dt + \sigma_{o,t} dW_{o,t}\} 1(C_{t-1} < t \leq O_t) + \{\mu_{d,t} dt + \sigma_{d,t} dW_{d,t}\} 1(O_t < t \leq C_t) \quad 0 \leq t \leq T \tag{4}$$

where the deterministic sequences $\mu_{o,t}$ and $\mu_{d,t}$ are spot *overnight* and *daytime* drifts, respectively, and $\sigma_{o,t}$ and $\sigma_{d,t}$ are spot volatility processes. The randomness in the overnight and daytime diffusions of the log-price process is dictated, respectively, by the Brownian motions $W_{o,t}$ and $W_{d,t}$.

At this point it is convenient to introduce the sigma-algebra $\Im_{t-1}$ in order to define the probability space underlying our bivariate modeling framework. $\Im_{t-1}$ is the union of the different sigma-algebras

generated by each set of high-frequency (HF) or intraday prices available up to day $t-1$ end; unless otherwise noted, the latter refers to the market closing time C_{t-1}. More formally, we have $\Im_{t-1} \equiv \Im_{HF,1} \cup \Im_{HF,2} \cup \ldots \cup \Im_{HF,t-1}$ with $\Im_{HF,j}$ the sigma-algebra obtained from the historical intraday prices observed on day j from C_{j-1} to C_j. Conditional on the sigma-algebra $\Im_{t-1}$, the Brownian motions in model (4) can be correlated, that is, $E[dW_{o,s} dW_{d,s^*} | \Im_{t-1}] = \rho_{ss^*}$ with $C_{t-1} < s < O_t < s^* < C_t$.[10]

From equation (4), it follows that the daily return process $r_t \equiv p_{C_t} - p_{C_{t-1}}$ can be expressed as $r_t = r_{o,t} + r_{d,t}$ where $r_{o,t} \equiv p_{O_t} - p_{C_{t-1}}$ and $r_{d,t} \equiv p_{C_t} - p_{O_t}$ are, respectively, the overnight and daytime returns. These two processes can be formalized, respectively, as

$$r_{o,t} = \int_{C_{t-1}}^{O_t} \mu_{o,s} ds + \int_{C_{t-1}}^{O_t} \sigma_{o,s} dW_{o,s}, \tag{5a}$$

and

$$r_{d,t} = \int_{O_t}^{C_t} \mu_{d,s^*} ds^* + \int_{O_t}^{C_t} \sigma_{d,s^*} dW_{d,s^*}. \tag{5b}$$

For notational simplicity we define $\mu_{o,t} \equiv \int_{C_{t-1}}^{O_t} \mu_{o,s} ds$ as the integrated deterministic drift driving the *overnight* return, and $\varepsilon_{o,t} \equiv \int_{C_{t-1}}^{O_t} \sigma_{o,s} dW_{o,s}$ as its random component; likewise, at *daytime*.

Proposition. *The daily integrated variance of the return process r_t is the sum of the integrated variances of the overnight and daytime components, $r_{o,t}$ and $r_{d,t}$, which can be formalized as*

$$IV_t = IV_{o,t} + IV_{d,t} \tag{6}$$

with $IV_{o,t} \equiv V(r_{o,t}) = \int_{C_{t-1}}^{O_t} \sigma_{o,s}^2 ds$ and $IV_{d,t} \equiv V(r_{d,t}) = \int_{O_t}^{C_t} \sigma_{d,s^*}^2 ds^*$.

The proof of this proposition is in Appendix A.

The decomposition formalized in (6) is exploited in the next section for constructing ex post realized volatility measures (i.e., estimates of IV_t) based on high frequency prices. Note that, although strictly

[10] Conditional on the sigma-algebra $\Im_{O_t}$ comprising high-frequency information up to the market opening, the Brownian motions $W_{o,t}$ and $W_{d,t}$ are uncorrelated. This result follows from the Law of Iterated Expectations: $E[dW_{o,s} dW_{d,s^*}] = E[E[dW_{o,s} dW_{d,s^*} | \Im_{O_t}]] = E[dW_{o,s} E[dW_{d,s^*} | \Im_{O_t}]] = 0$, since $E[dW_{d,s^*} | \Im_{O_t}] = 0$ for all $s^* \in [O_t, C_t]$, by definition of a Brownian motion.

speaking the term integrated volatility refers to the square root of IV_t, in this paper we interchangeably use both integrated variance and volatility to refer to IV_t.

3.2 Overnight-Adjusted Realized Variance Estimators

This section discusses the ex post estimation of the integrated variance, IV_t, using high frequency (intraday) prices corresponding to day t. The object of interest, IV_t, can be cast as a time-varying parameter and the estimation approaches outlined below are model-free or nonparametric.

Among the model-free ex post measures of IV_t available, the most popular is the *realized variance* computed as the sum of the M squared returns corresponding to equal-length (e.g., 5-minute) intervals

$$RV_t = \sum_{j=1}^{M} r_{j,t}^2, \tag{7}$$

where $r_{t,j} \equiv p_{t,j} - p_{t,j-1}$ is the jth intraday return on day t. Under certain conditions (e.g., absence of jumps) it can be shown that $RV_t \to IV_t$ asymptotically as $M \to \infty$. More often than not, this estimator has been computed from returns spanning the open to close period since it is typically deployed in the context of individual stocks or equity cash indices that are only traded for part of the day (i.e. daytime period). It can be shown that the open-to-close RV_t can underestimate in practice (finite M) the true notional daily volatility IV_t, as defined in (3), because it does not span a full 24-hour period.[11] The existing empirical literature on equity market volatility has taken different stances on how to incorporate overnight information into a daily realized volatility estimator.[12]

An ad hoc solution consists of scaling the RV_t measure upwards so that it spans a 24 hr period

$$RV_t^{SC} = \frac{\sum_{t=1}^{T} r_{d,t}^2 + \sum_{t=1}^{T} r_{o,t}^2}{\sum_{t=1}^{T} r_{d,t}^2} RV_t, \tag{8}$$

[11] This criticism extends to alternative high-frequency measures of daily risk including the *realized bipower variation* that also converges in probability to the integrated variance, and *kernel-based realized variance* estimators such as those proposed by Barndorff-Nielsen et al. (2008) to mitigate market microstructure noise. The latter is especially an issue at ultra-high frequencies finer than the 5-minute. See McAleer and Medeiros (2008) for a survey.

[12] Some papers model directly the open-to-close return alongside the realized variance computed from intraday squared returns entirely ignoring the overnight period (Brownlees and Gallo, 2010; Noureldin et al., 2011; Chevallier et al., 2011; Shephard and Sheppard, 2010; Liu and Maheu, 2009; Fuertes et al., 2009).

which has been adopted by Martens (2002), Fleming et al. (2003), Koopman et al. (2005), and Angelidis and Degiannakis (2008a) among others.

A different approach is to treat the overnight return as one more squared summand ($+ON$) alongside the M intraday returns from market open to close as follows

$$RV_t^{+ON} = r_{o,t}^2 + r_{1,t}^2 + r_{2,t}^2 + ... + r_{M,t}^2 = r_{o,t}^2 + RV_t, \tag{9}$$

which has been followed, for instance, by Martens et al. (2009), Bollerslev et al., (2009), de Pooter et al. (2008), Becker et al. (2007), Giot and Laurent (2004), Martens (2002) and Blair et al. (2001).

More formally, in an attempt to make the resulting realized variance estimator more accurate, Hansen and Lunde (2005) derive weights for the overnight return and the sum of squared intraday returns. In particular, their overnight-adjusted realized variance is given by

$$RV_t^{HL} = \omega_1^* r_{o,t}^2 + \omega_2^* RV_t, \tag{10}$$

where ω_1^* and ω_2^* are the solution to the optimization problem $\min_{\omega_1,\omega_2} \text{var}(RV_t^{HL})$, s.t. $\omega_1\mu_1 + \omega_2\mu_2 = \mu_0$ where $\mu_1 \equiv E[r_{o,t}^2]$, $\mu_2 \equiv E[RV_t]$ and $\mu_0 \equiv E[IV_t]$, $t = 1, ..., T$. Thus their procedure aggregates the overnight return and the sum of intraday returns (all in squared form) with weights defined as those that produce the most efficient realized variance estimator. The weighting scheme put forward by Hansen and Lunde (2005) offers an attractive alternative to naive combinations such as that given by $(\omega_1, \omega_2) = (1, 0)$ that obviates the information coming from the overnight return, and combination (9) given by $(\omega_1, \omega_2) = (1, 1)$ that effectively amounts to treating the 17.5 hr overnight return (from 4:00pm to 9:30am) in the same way as each of the 5-minute intra-day returns.[13]

3.3 Conditional Daily Mean and Variance Forecasts

We now discuss two modeling approaches for the integrated mean and variance components of the daily return process that differ in the conditioning information set. The models either exploit the

[13] Ahoniemi and Lanne (2012) compare the above realized variance estimators using Patton's (2011) testing framework by adopting the daily squared return as unbiased proxy for the true daily volatility. Their findings suggest for the S&P 500 index that the weighting scheme of Hansen and Lunde (2005) provides the most accurate RV estimator.

available (high-frequency) information up to the closing of day $t-1$, the sigma-algebra $\Im_{t-1}$, or further information up to the opening of day t, the sigma-algebra $\Im_{O_t}$.

Let us assume that the overnight and daytime return data generating processes (DGP) are

$$r_{o,t} = \mu_{o,t|t-1} + \varepsilon_{o,t}, \tag{11a}$$

with $\mu_{o,t|t-1} \equiv E[r_{o,t}|\ \Im_{t-1}]$ the conditional mean of the overnight return, and

$$r_{d,t} = \mu_{d,t|t-1} + \varepsilon_{d,t}, \tag{11b}$$

with $\mu_{d,t|t-1} \equiv E[r_{d,t}\ |\ \Im_{t-1}]$ the conditional mean of the daytime return. The conditional variances of $r_{o,t}$ and $r_{d,t}$ are, respectively, $IV_{o,t|t-1} \equiv E[\varepsilon^2_{o,t}|\ \Im_{t-1}]$ and $IV_{d,t|t-1} \equiv E[\varepsilon^2_{d,t}|\ \Im_{t-1}]$. In line with the continuous-time framework presented in Section 3.1, it is important to note that $Cov(r_{o,t}, r_{d,t}|\ \Im_{t-1}) \equiv E[\varepsilon_{o,t}\varepsilon_{d,t}|\ \Im_{t-1}] \equiv IV_{o,d,t|t-1}$ can be different from zero. This is because by generating the forecast at time $t-1$ (i.e., given $\Im_{t-1}$) both $r_{o,t}$ and $r_{d,t}$ are random variables. Further, by simple algebra it can be shown that the conditional covariance $IV_{o,d,t|t-1} = \int_{C_{t-1}}^{O_t}\int_{O_t}^{C_t} E[\sigma_{o,s}\sigma_{d,s^*}dW_{o,s}dW_{d,s^*}|\ \Im_{t-1}]$ is, under the minimal regularity assumptions of independence between the spot volatility processes and the corresponding Brownian motions, equal to $\int_{C_{t-1}}^{O_t}\int_{O_t}^{C_t} E[\sigma_{o,s}|\ \Im_{t-1}]E[\sigma_{d,s^*}|\Im_{t-1}]\rho_{ss^*}ds^*ds$ where ρ_{ss^*} is the spot correlation process as defined above after equation (4).

Finally, noting that $r_t = r_{o,t} + r_{d,t}$, by simple algebra the DGP of daily returns can be expressed as $r_t = \mu_{t|t-1} + \varepsilon_t$ with $\varepsilon_t = \varepsilon_{o,t} + \varepsilon_{d,t}$. The *conditional mean* $\mu_{t|t-1} \equiv E[r_t|\ \Im_{t-1}]$ is given by

$$\mu_{t|t-1} = \mu_{o,t|t-1} + \mu_{d,t|t-1}, \tag{12}$$

and the *conditional variance* process $IV_{t|t-1} \equiv V(r_t|\ \Im_{t-1}) = E[\varepsilon^2_t|\ \Im_{t-1}]$ is decomposed as

$$IV_{t|t-1} = IV_{o,t|t-1} + IV_{d,t|t-1} + 2IV_{o,d,t|t-1}. \tag{13}$$

In order to make the above expressions operational, i.e. to obtain forecasts $\widehat{\mu}_{t|t-1}$ and $\widehat{IV}_{t|t-1}$, three data histories denoted $\{r_{o,1}, r_{o,2}, ..., r_{o,t-1}\}$, $\{r_{d,0}, r_{d,1}, ..., r_{d,t-1}\}$ and $\{IV_1, IV_2, ..., IV_{t-1}\}$ can be used as conditioning set $\Im_{t-1}$. Whereas the first two (return) histories are derived from observed

market prices, this is not the case for the third (variance) history because the IV process is latent. However, the latter can be proxied ex-post by the realized variance using intraday (e.g., 5-minute) returns), that is, by the time-series $\{RV_1, RV_2, ..., RV_{t-1}\}$.[14]

A risk manager can generate the day-ahead VaR limits either at the market close time C_{t-1} or at the next day market open time O_t; the choice of the latter over the former has interesting implications. The conditional mean forecast of the daily return (or profit and loss) distribution constructed at O_t is given by $E[r_t|\Im_{O_t}] = r_{o,t} + \mu_{d,t|O_t}$, and can differ from $E[r_t|\Im_{t-1}]$ in (12). For the daily conditional variance of the return distribution, the risk manager has $IV_{t|O_t} = IV_{d,t|O_t}$, that can be different from $IV_{t|t-1}$ in (13). Interestingly, if the risk managers produces the equity VaR limits for day t at time O_t, he loses any forecasting power coming from the conditional covariance between overnight and daytime returns, and instead relies on the information content of the actual overnight previous close-to-open return $r_{o,t}$. This conditioning approach (on $\Im_{O_t}$) implicitly hinges on the assumption that the market is efficient and p_{O_t} reflects all available overnight news. This may not be the case if, for instance, in the context of stock index data, the open price suffers from staleness. The staleness is exacerbated when the proportion of small-cap, low-trading activity companies in the index is larger.

In order to materialize $IV_{t|t-1}$ in (13) as a forecast, we propose a bivariate modeling approach for the overnight and intraday segments of the day. In both segments, we allow for asymmetry or the statistical 'leverage effect' in the conditional variance equation and asymmetry and leptokurtosis in the innovations. For the latter we adopt a skewed Student-t density following previous studies (e.g., Angelidis and Degiannakis, 2008a; Giot and Laurent, 2004).[15] Given the three time series forming the information set, $\Im_{t-1} \equiv \{r_{o,1}, ..., r_{o,t-1}; r_{d,1}, ..., r_{d,t-1}; RV_1, ..., RV_{t-1}\}$, the conditional integrated variance of the *overnight* return cannot be observed even *ex post*. In order to forecast it, we follow Andersen et al. (2011) who note the strong volatility clustering of the overnight return and advocate a GARCH framework to model it. Conditionally on $\Im_{t-1}$, we fit by quasi maximum likelihood (QML)

[14] Barndorff-Nielsen and Shephard (2002) show that realized variance converges to integrated variance at rate $\sqrt{M}$.

[15] The computations are carried out in Oxmetrics 6 using the G@RCH 6.1 package for the overnight modeling and the second step involved in daytime modeling, and the ARFIMA 1.04 package for the daytime modeling.

two equations to characterize the conditional mean and variance, respectively, of the overnight return process: an autoregressive (AR) model and the threshold GARCH (TGARCH) model of Glosten, Jagannathan, and Runkle (1993), expressed as

$$r_{o,t} = a_0 + a_1 r_{d,t-1} + b_1 r_{o,t-1} + \sqrt{h_{o,t}} \cdot \epsilon_{o,t} \tag{14a}$$

$$h_{o,t} = \alpha_0 + \alpha_1 \varepsilon_{o,t-1}^2 + \beta_1 h_{o,t-1} + \gamma_1 I_{t-1}^- \varepsilon_{t-1}^2 \tag{14b}$$

where $\varepsilon_{o,t} \equiv \sqrt{h_{o,t}} \cdot \epsilon_{o,t}$, and the standardized innovation $\epsilon_{o,t}$ is i.i.d. skewed Student-$t(1, 0, \xi_o, \eta_o)$ with ξ_0 and η_o the unknown parameters that capture the degree of asymmetry and fat tailedness, respectively, in the overnight return distribution; ε_t are innovations in a standard AR(1) model for the daily return r_t; the asymmetry indicator function is defined as $I_t^- = 1$ if $\varepsilon_t < 0$ and $I_t^- = 0$ otherwise; $\gamma_1 > 0$ implies that a large negative innovation on day $t-1$ (i.e., bad news) increases the subsequent conditional volatility of overnight returns more than a large positive innovation of the same magnitude. Second, the information set $\Im_{t-1}$ is exploited to obtain forecasts for the conditional mean and variance of the daytime return process as follows. We fit an AR equation to the open-to-close returns and a long memory ARFIMA$(0, d, 1)$ equation to the realized variance, expressed as[16]

$$r_{d,t} = a_0 + a_1 r_{d,t-1} + b_1 r_{o,t-1} + \sqrt{{\sigma_d}^2 \cdot RV_{t|t-1}} \cdot \epsilon_{d,t} \tag{15a}$$

$$(1-L)^d (\ln RV_t - \tau_0 - \tau_1 r_{t-1} - \tau_2 I_{t-1}^- r_{t-1}) = (1 + \theta L) u_t \tag{15b}$$

where the the standardized innovation $\epsilon_{d,t}$ is i.i.d. skewed Student-$t(0, 1, \xi_d, \eta_d)$, the innovation u_t is i.i.d.$(0, \sigma_u^2)$ and the asymmetry indicator function is defined $I_t^- = 1$ if $r_t < 0$ and $I_t^- = 0$ otherwise; parameter values $\tau_2 < 0$ indicate that bad news increase the future daytime volatility more than good news. This conditional variance modeling approach is motivated by the stylized long memory and log-normality properties of the realized variiance. The AR-ARFIMAX parameters can be consistently estimated following Giot and Laurent's (2004) two-step approach. First, equation (15b) is estimated

[16] Other specifications that have been used in the empirical finance literature to approximate the long memory properties of the realized volatility are Corsi's (2004) heterogeneous autoregressive (HAR) model and Ghysels et al.'s (2004) MIxed Data Sampling (MIDAS) model both of which combine information sampled at different frequencies.

by ML under a normality assumption for u_t. Second, equation (15a) is estimated by QML under the assumption that all the dynamics in the conditional variance is captured by the ARFIMA model.

This *bivariate* modeling framework can account for dynamics in the conditional integrated covariance between $r_{o,t}$ and $r_{d,t}$, denoted $IV_{o,d,t|t-1}$. Different modeling strategies ranging from the non-parametric J.P.Morgan RiskMetrics EWMA approach, to fully parametric approaches such as the BEKK model of Engle and Kroner (1995), or further refinements provided, for example, by Kroner and Ng (1998), can be employed to obtain conditional covariance forecasts. Engle (2002) proposes a two-step dynamic conditional correlation (DCC) framework based on individual GARCH processes, although as noted by Engle (2002; p.342): "Nothing would change if this was generalized." We extend the DCC model in this vein, namely, the overnight conditional variance is of GARCH type whereas the daytime conditional variance is of ARFIMA realized volatility type. Formally, we have

$$q_{o,d,t} = \overline{\rho}_{o,d}(1-\alpha-\beta) + \alpha\epsilon_{o,t-1}\epsilon_{d,t-1} + \beta q_{o,d,t-1}, \tag{16}$$

with $\epsilon_{o,t-1}$ and $\epsilon_{d,t-1}$ denoting the standardized overnight and daytime innovations consistently estimated via the AR-TGARCH and AR-ARFIMA model formulations (13) and (14), respectively, $q_{o,d,t}$ is the conditional correlation of those innovations and $\overline{\rho}_{o,d} \equiv E(\epsilon_{o,t} \cdot \epsilon_{d,t})$ the unconditional correlation. Mean reversion of the covariance process follows from $\alpha + \beta < 1$.[17] The conditional covariance of the unstandardized innovations, $\varepsilon_{o,t}$ and $\varepsilon_{d,t}$ is $IV_{o,d,t} = q_{o,d,t}\sqrt{RV_t \times h_{o,t}}$.

An alternative to this novel bivariate modeling proposal is to incorporate the overnight surprise into the realized variance prior to modeling. Effectively, this strategy amounts to *bundling* the squared previous-close to open return with the summation of daytime high-frequency (5 minute) squared returns.[18] Now the overnight-adjusted realized volatilities discussed in Section 3.2, RV_t^{SC}, RV_t^{+ON} and RV_t^{HL}, replace the time-series RV_t in the ARFIMA specification (15b). The daily conditional

[17] J.P.Morgan's RiskMetrics approach based on an EWMA filter amounts to an integrated DCC (IDCC) process, namely, $\alpha + \beta = 1$ in (16), which gives $q_{o,d,t} = (1-\lambda)\sum_{j=1}^{\infty}\lambda^n \epsilon_{o,t-j}\epsilon_{d,t-j}$ with $\lambda = 0.94$ for daily data.

[18] A simpler approach used in a number of papers amounts to 'mismatching' a conditional mean model for the *close-to-close* return with a conditional variance model for realized variance computed from *open-to-close* intraday returns (Brownlees and Gallo, 2010; Wu, 2011; Corsi et al., 2008; Thomakos and Wang, 2003; Andersen et al., 2001).

mean forecast $\mu_{t|t-1}$ comes from a short-memory AR model fitted to the close-to-close return under the assumption, as before, that the dynamics in the conditional variance is long-memory ARFIMA type. In the context of the overnight-adjusted measure RV_t^{HL}, the model is

$$r_t = c_0 + c_1 r_{t-1} + \sqrt{\sigma^2 \cdot RV_{t|t-1}^{HL}} \cdot \epsilon_t \tag{17a}$$

$$(1-L)^d(\ln RV_t^{HL} - \tau_0 - \tau_1 r_{t-1} - \tau_2 I_{t-1}^{-} r_{t-1}) = (1+\theta L)u_t \tag{17b}$$

where the standardized daily innovation ϵ_t is i.i.d. skewed Student-$t(0,1,\xi,\eta)$; likewise for RV_t^{+ON} and RV_t^{SC}. This approach is simpler than the bivariate but more restrictive in that it precludes the modeling of the potentially different dynamics in the overnight and daytime return processes, and the conditional covariance between them. Nevertheless, overnight-adjusted measures are widely employed in the empirical high-frequency finance literature (see e.g. Blair et al., 2001; Martens, 2002; Giot and Laurent, 2004; Koopman et al. 2005; Angelidis and Degiannakis, 2008a; Fuertes and Olmo, 2012).

An alternative framework for a risk manager to set day-ahead VaR-based equity trading limits is to adopt $\Im_{O_t}$ as conditioning information set. Effectively, this amounts to constructing the VaR limits for day t at the market opening of day t when the overnight return $r_{o,t} \equiv p_{O_t} - p_{C_{t-1}}$ is known. Hence, the only uncertainty left pertains to the daytime segment ahead, $IV_{t|O_t} = IV_{d,t|O_t}$, since the overnight segment has elapsed and thus, by definition, $IV_{o,t|O_t} = IV_{o,d,t|O_t} = 0$. In this *ex post* approach, the forecasts $\mu_{d,t|O_t}$ and $IV_{d,t|O_t}$ are obtained through an AR-ARFIMA specification which is a small variation of (14) to exploit the fact that the overnight return for day t is part of the information set at the time the VaR limits for day t are generated. Formally, the model is

$$r_{d,t} = a_0 + a_1 r_{d,t-1} + b_1 r_{o,t-1} + b_0 r_{o,t} + \sqrt{{\sigma_d}^2 \cdot RV_{t|O_t}} \cdot \epsilon_{d,t} \tag{18a}$$

$$(1-L)^d(\ln RV_t - \tau_0 - \tau_1 r_{t-1} - \tau_2 I_{t-1}^{-} r_{t-1} - \tau_3 r_{o,t} - \tau_4 I_{o,t}^{-} r_{o,t}) = (1+\theta_1 L)u_t \tag{18b}$$

and allowing, like before, for a skewed Student-$t(0,1,\xi_d,\eta_d)$ density for the standardized daytime innovations $\epsilon_{d,t}$. A number of studies in the literature adopt a modeling strategy that broadly falls

within this framework in the sense that they work with realized variance measures computed from open-to-close intraday returns matched with daily open-to-close returns (Chevallier et al., 2011; Noureldin et al., 2011; Liu and Maheu, 2009; Fuertes et al., 2009).

3.4 Competing Location-Scale VaR Models

Next we discuss how the above forecasts can be used as building blocks in location-scale VaR models to set daily equity trading limits. This is a relevant risk management application given that banks routinely use VaR to get a sense of the riskiness of the trading behavior of their business lines or 'desks', and to set appropriate trading limits for them to control overall risk. The *bivariate*-based VaR at nominal coverage level α conditional on the information set $\Im_{t-1}$ can be expressed as

$$VaR^{bivariate}_{t|t-1,\alpha} = \mu_{t|t-1} + \sqrt{IV_{t|t-1}}F^{-1}_{\epsilon}(\alpha) \tag{19}$$

where the first term $\mu_{t|t-1}$ is obtained as the sum of two forecasts from the overnight and daytime AR equations, (14a) and (15a), respectively, $\mu_{t|t-1} = \mu_{o,t|t-1} + \mu_{d,t|t-1}$ according to (12); the second term $IV_{t|t-1}$ is obtained by aggregating the three forecasts corresponding, respectively, to the TGARCH, ARFIMA and DCC equations, $IV_{t|t-1} = IV_{o,t|t-1} + IV_{d,t|t-1} + 2IV_{o,d,t|t-1}$, according to (13); and $F^{-1}_{\epsilon}(\alpha)$ is the α-quantile of the skewed Student-t distributed standardized innovation ϵ_t.

Alternatively, in the spirit of what we call above the *bundling* approach, one can construct VaR limits based on overnight-adjusted realized volatility measures. The corresponding VaR forecast is

$$VaR^{bundle}_{t|t-1,\alpha} = \widetilde{\mu}_{t|t-1} + \sqrt{\widetilde{IV}_{t|t-1}}F^{-1}_{\epsilon}(\alpha) \tag{20}$$

with $\widetilde{\mu}_{t|t-1}$ the conditional mean forecast of the close-to-close return, obtained from the AR equation (17a), and $\widetilde{IV}_{t|t-1}$ the integrated daily variance forecast obtained from the ARFIMA equation (17b).

In our third approach, equity VaR limits for day t are set using information up to the opening of the market, denoted $\Im_{O_t}$. The corresponding VaR forecast is

$$VaR^{expost}_{t|O_t,\alpha} = \mu_{t|O_t} + \sqrt{IV_{t|O_t}}F^{-1}_{\epsilon_d}(\alpha), \tag{21}$$

with $\mu_{t|O_t} = r_{o,t} + \mu_{d,t|O_t}$ and $IV_{t|O_t} = IV_{d,t|O_t}$ since the overnight return is now known and becomes an ex post observation. The *daytime* conditional mean forecast $\mu_{d,t|O_t}$ is obtained from the AR(1) equation (18a) above and the *daytime* conditional variance forecasts $\widehat{IV}_{d,t|O_t}$ from the ARFIMA equation (18b); $F^{-1}_{\epsilon_d}(\alpha)$ is the α-quantile of the skewed Student-t distributed standardized innovation $\epsilon_{d,t}$.

4 Backtesting the Performance of VaR Forecasts

Let the in-sample (estimation) period comprise P days and the out-of-sample (evaluation) period comprise n days with $P + n = T$. Let θ_0 denote the true parameter vector of the location-scale model considered. A correctly specified α-th conditional VaR model of the portfolio returns r_t is defined as

$$P(r_t \leq VaR_{t,\alpha}(\theta_0) \mid \Im_{t-1}) = \alpha, \text{ almost surely (a.s.)}, \alpha \in (0,1), \ \forall t \in \mathbb{Z}, \tag{22}$$

a moment restriction often used in theoretical discussions (see e.g., Christoffersen et al., 2001; Engle and Manganelli, 2004; Gourieroux and Jasiak, 2006; Koenker and Xiao, 2006; Fuertes and Olmo, 2012). Let $Hit_{t,\alpha}(\theta_0) \equiv 1(r_t \leq VaR_{t,\alpha}(\theta_0))$ denote the out-of-sample *hits* or exceedances sequence. If the conditional VaR model is correctly specified out-of-sample then it follows that the expected value of the hits conditional on $\Im_{t-1}$ is α. Definition (22) also implies independence or that the demeaned sequence of VaR violations, $Hit_{t,\alpha}(\theta_0) \equiv 1(r_t \leq VaR_{t,\alpha}(\theta_0)) - \alpha$, does not depend on the available information set $\Im_{t-1}$ which includes previous hits, past daily returns, and volatilities. Intuitively, violations of the VaR should be happening in a random, unpredictable manner. In other words, if it can be predicted that the VaR violations will increase in the near future, then this information should be included in the model used to compute the equity trading VaR limits.

In practice, the parameter vector θ_0 is not known and needs to be estimated from the data. Let $\widehat{\theta}_{P,t-1}$ denote a rolling estimator of θ_0 obtained over a moving window of P observations; a natural way to assess the validity of condition (22) is by implementing a dynamic quantile (DQ) test of the null hypothesis $H_0 : E[Hit_{t,\alpha}(\widehat{\theta}_{P,t-1}) \mid x_{t-1}(\widehat{\theta}_{P,t-1})] = 0$, with $x_{t-1}(\widehat{\theta}_{P,t-1})$ a vector of k regressors, including a constant, that subsumes all relevant information, $\Im_{t-1}$[19]. Engle and Manganelli (2004)

[19] In order for this test to be properly defined the estimation sample size P needs to be fixed which rules out the

propose a Wald type statistic for assessing (22) which can be cast as a joint test for the statistical significance of the coefficients in the following linear probability model[20]

$$Hit_{t,\alpha}(\widehat{\theta}_{P,t-1}) = x_{t-1}(\widehat{\theta}_{P,t-1})\gamma + v_t,\ t = P+1,\ldots,T, n = T-P \tag{23}$$

with $Hit(\widehat{\theta}_{P,t-1})$ a $n \times 1$ vector obtained from the sequence of ex post out-of-sample hits or VaR violations, γ a $k \times 1$ parameter vector and v_t a zero-mean iid error sequence. The conditional coverage hypothesis is defined as $H_0 : \gamma_0 = \gamma_1 = ... = \gamma_{k-1} = 0$ which can be tested, using the terminology introduced in Engle and Manganelli (2004), by the following dynamic quantile (DQ) statistic;

$$DQ = n\frac{\widehat{\gamma}_n'[M_n(\widehat{\theta}_P)]\widehat{\gamma}_n}{\alpha(1-\alpha)} \tag{24}$$

where $\widehat{\gamma}_n$ is the OLS estimator of γ in (23), and $M_n(\widehat{\theta}_P) = \frac{1}{n}\sum_{t=P+1}^{T} x_{t-1}'(\widehat{\theta}_{P,t-1})x_{t-1}(\widehat{\theta}_{P,t-1})$ a $k \times k$ matrix. The test statistic converges in distribution to a χ_k^2 density under H_0 as $n \to \infty$.

The binary character of the hits sequence may imply the existence of heteroscedasticity in the error term v_t in the linear probability model (LPM) that underlies the original test of Engle and Manganelli (2004), equation (23). To correct for this heteroscedasticity in the DQ test, Berkowitz et al. (2011) propose a refinement of the method based on the logit version of (23). Probit or logit regressions are generally more convenient than LPMs because they bound the fitted values from the binary model (probabilities) to be between 0 and 1. We adopt the dynamic binary (DB) probit regression

$$E[1(r_t \leq VaR_{t,\alpha}(\widehat{\theta}_{P,t-1})) \mid x_{t-1}(\widehat{\theta}_{P,t-1})] = \Phi(x_{t-1}(\widehat{\theta}_{P,t-1})\beta),\ t = P+1,\ldots,T \tag{25}$$

with $\Phi(\cdot)$ the cumulative standard Normal distribution and $\beta \equiv (\beta_0, \beta_1, ..., \beta_{k-1})'$ the vector of regression parameters. Although the probit model is already deployed in Taylor (2007) in a VaR backtesting context, a formal analysis of its small sample properties (i.e. size and power) has only been provided

adoption of recursive estimation windows for θ_0 for which $P \to \infty$. Thus the uncertainty from using estimated regressors is naturally accounted for. Giacomini and Komunjer (2005) follow the same approach for VaR encompassing inference.

[20] Gaglianone et al. (2011) discuss an alternative testing strategy for (22) based on quantile regressions in order to assess whether the conditional quantile of interest is given by the proposed $VaR_{t,\alpha}(\theta_0)$ model. Other conditional coverage tests have been developed based on the duration between consecutive violations; see e.g. Berkowitz et al. (2011), and Candelon et al. (2011).

recently by Dumitrescu et al. (2012). Their study shows that the probit-based backtesting appears more robust to estimation risk, with less size distortions and better power than the original DQ test, and also that likelihood ratio (LR) tests exhibit better small sample properties (in terms of power) than Wald tests in the probit context. Under stationarity of the regressors, the estimation of probit models can be performed by ML to obtain consistent and asymptotically normal parameter estimators.

In the probit setting, the correct conditional coverage criterion encapsulated in (22) amounts to

$$E[1(r_t \leq VaR_{t,\alpha}(\widehat{\theta}_{P,t-1})) \mid x_{t-1}(\widehat{\theta}_{P,t-1})] = \Phi(\beta_0) = \alpha \tag{26}$$

which can be stated as $\widetilde{H}_0 : \beta_0 = \Phi^{-1}(\alpha), \beta_1 = ... = \beta_{k-1} = 0$. In contrast to the regression model (23), the dependent variable in (25) is not demeaned, implying that under the null hypothesis, the correct specification of the conditional $\alpha-$quantile process is defined by the restrictions $\beta_0 = \Phi^{-1}(\alpha)$ and $\beta_1 = ... = \beta_{k-1} = 0$. This joint hypothesis test is implemented via the LR statistic

$$DB = 2(\mathcal{L} - \mathcal{L}_0) \tag{27}$$

with $\mathcal{L} = \sum_{t=P+1}^{T} [Hit_{t,\alpha}(\widehat{\theta}_{P,t-1}) ln\ \Phi(x_{t-1}(\widehat{\theta}_{P,t-1})\widehat{\beta}_n) + (1 - Hit_{t,\alpha}(\widehat{\theta}_{P,t-1})) ln\ (1 - \Phi(x_{t-1}(\widehat{\theta}_{P,t-1})\widehat{\beta}_n))]$ the log-likelihood where $\widehat{\beta}_n$ is the ML estimator of β. Under the null hypothesis, $\mathcal{L}_0 = \sum_{t=P+1}^{T} [Hit_{t,\alpha}(\widehat{\theta}_{P,t-1}) ln\ \alpha + (1 - Hit_{t,\alpha}(\widehat{\theta}_{P,t-1})) ln\ (1 - \alpha)]$, and DB converges in distribution to a χ^2_k as $n \to \infty$.

We follow Engle and Manganelli (2004) and Dumitrescu et al. (2012) in adopting as regressors

$$x_{t-1}(\widehat{\theta}_{P,t-1}) = (1, r_{t-1}, r^2_{t-1}, VaR_{t-1,\alpha}(\widehat{\theta}_{P,t-1}), Hit_{t-1,\alpha}(\widehat{\theta}_{P,t-1})).$$

This vector of regressors controls for dependence between hits and past returns, past volatility and past VaRs. The inclusion of lagged hits seeks to control for potential autocorrelation of the hits. The specific choice of the vector $x_{t-1}(\widehat{\theta}_{P,t-1})$ is not without importance. One of the main differences between the above framework for testing the correct specification of the VaR model and other standard backtesting procedures lies in the choice of the conditioning information set. Thus the very popular conditional coverage probability likelihood ratio test of Christoffersen (1998) or the spectral density

test and duration based tests discussed in detail in Berkowitz et al. (2011) use an information set only given by the sequence of lagged hits. These tests, in turn, do not exhibit power to reject the existence of dependence of the VaR exceedances with other relevant variables.

5 Empirical Application

This section presents our equity index data, the parameter estimates of the conditional mean, variance, and correlation models, and the VaR backtesting results.

5.1 Data and Descriptive Statistics

Two major stock market indices are chosen as illustrative equity portfolios: S&P 500, by far the most common benchmark for funds (mutual funds, ETFs, and pension funds) that identify themselves as large cap, and Russell 2000, the typical benchmark for small cap funds. The observations are 5-minute quotations from 9:30-16:00 Eastern time (EST) which amounts to $M = 78$ intraday intervals. The observed opening price on day t is the first price recorded after 9:30. The jth intra-day price $p_{t,j}$ with $j = 1, ..., M - 1$ is defined as the last seen tick before the jth 5-minute mark. The day t closing price $p_{C_t} \equiv p_{t,M}$ is defined as the last price observed before 16:00. The two equity portfolios are observed over the most recent 15-year period available from November 12, 1997 to September 31, 2011 ($T = 3491$ trading days). As of September 2011, the S&P 500 index comprises roughly 20% Nasdaq-traded (89/500) and 80% NYSE-traded (411/500) companies, whereas for the Russell 2000, the breakdown is roughly 60% Nasdaq (1250/2000) and 40% NYSE (750/2000). Measured by trading volume, there is a marked difference between the two indices. The average daily volume of trading on S&P 500 futures contracts is about 16 times larger than that on Russell 2000 futures.

Figure 1 plots daytime and overnight returns and clearly reflects the waves of panic sent into the stock market by the disastrous collapse of Lehman Brothers, and the effective nationalization of Fanny Mae and Freddie Mac in September 2008. Volatility remains high in the initial post-Lehman months, and gradually starts subsiding until a second cluster of high volatility is triggered around April 2010,

the month when the Securities and Exchange Commission (SEC) charged Goldman Sachs with fraud in structuring and marketing of a collateralized debt obligation (CDO) tied to subprime mortgages. A final volatility cluster is observed in the final month of the sample (September 2011) when terms of a tranche of bailout funds needed to prevent the default of Greece were discussed. Our out-of-sample period begins on September 2, 2008 ($n = 778$ days), an important landmark of the late 2000s global financial crisis to make the forecasting task more challenging.[21,22]

[Insert Figure 1 around here]

Table 1 summarizes the daily overnight and daytime return distribution over pre- and post-Lehman samples spanning, respectively, the periods 12/11/1998 to 29/08/2008, and 02/09/2008 to 31/09/2011.

[Insert Table 1 around here]

For both indices, and over both sub-periods, the findings confirm previous stylized facts: stock return volatility is substantially higher in daytime (trading) than overnight (non-trading) periods. However, both daytime and overnight return variances increase dramatically in the post-Lehman period. The more negative skewness and higher kurtosis for overnight returns than daytime returns observed during the pre-Lehman period is notably lessened during the post-Lehman crisis period, when the daytime returns also become markedly negatively skewed and kurtosed. The kurtosis of overnight returns is about three times larger than that of daytime returns in the pre-Lehman period, but becomes roughly similar for both segments of the day during the post-Lehman period. Overnight and daytime returns pertaining to the same day ($r_{o,t}, r_{d,t}$) are strongly positively correlated throughout, but there

[21] Fannie Mae and Freddie Mac, two U.S. government sponsored enterprises, owned or guaranteed nearly $5 trillion in mortgage obligations at the time they were placed into conservatorship by the U.S. government on September 7. Lehman Brothers filed for Chapter 11 bankruptcy protection on September 15.

[22] The quotes are from Disk Trading http://www.is99.com/disktrading/. We follow most of the existing literature in adopting the 5-minute sampling frequency. This is the typical sampling interval of choice because, for highly liquid assets, it has been shown to be short enough for the daily volatility dynamics to be picked up with reasonable accuracy (small estimation error) and long enough for the adverse effects of market microstructure noise (e.g., bid-ask bounce, discrete price observations, irregular trading) not to be excessive.

is a substantial increase in correlation from the pre- to the post-Lehman periods. Both returns exhibit (1st order) autocorrelation, typically negative, but in milder form daytime than overnight. Both autocorrelations increase in the post-Lehman period.

The overnight and daytime squared returns are summarized in Panel B of Table 1. To make the comparison more informative, we report the hourly volatility as the mean squared return divided by the number of hours spanned by each segment of the day, 17.5 hr (overnight) and 6.5 hr (daytime). Return volatility is higher daytime than overnight, which is well aligned with market microstructure wisdom that information flow is greater during trading than during non-trading hours. The volatility autocorrelation function shows a much slower decay at daytime (i.e. more persistence) than overnight. Both volatilities increase dramatically from the first to the second sample periods. Overall the statistics in Table 1 reveal important differences in the overnight and daytime return distributions.

Table 2 summarizes the distribution of realized variances, baseline and overnight-adjusted, which proxy the daily integrated volatility. The baseline measure is computed as in Eq. (7), and the overnight-adjusted measures, RV^{SC}, RV^{+ON} and RV^{HL}, correspond with Eqs. (8) to (10).

[Insert Table 2 around here]

All distributions show high positive skewness and kurtosis, corroborating that investors face non-normally distributed risks. Skewness and kurtosis are, by construction, identical for RV_t and RV_t^{SC}. The means of RV_t^{HL} and RV_t^{+ON} are the same by construction. The coefficient of variation, or the relation of standard deviation to the mean, is always the lowest for RV^{HL}, reflecting the fact that the weights for $r_{o,t}^2$ and RV_t are determined in Hansen and Lunde's (2005) approach by minimizing $var(\omega_1 r_{o,t}^2 + \omega_2 RV_t)$.[23] Subsequently, we model the logarithmic realized volatility series, which is standard in the empirical finance literature given that the Gaussianity assumption is more suitable for logs than for levels. In untabulated results, the skewness of the log realized volatilities (baseline and

[23]Note also that the coefficient of variation is equal for RV_t and RV_t^{SC} by construction given that the scaled estimator is a linear transformations of RV_t. The coefficient of variation is larger in the out-of-sample period for all realized variance measures, indicating a decrease in the signal-to-noise ratio during the late 2000s global financial crisis period.

overnight-adjusted) is close to zero, with a range [-0.14, 0.45] for the two indices, and their kurtosis coefficient is close to 3 for both pre- and post-Lehman periods.

Over the entire 15-year period, the percentage of daily volatility that occurs during the inactive period, as proxied by the ratio $r_{o,t}^2/(r_{o,t}^2 + RV_t)$ is 5.95% for the Russell 2000 and 2.83% for S&P 500. For Russell, the ratio more than doubles from 4.70% for the pre-Lehman period (2713 days) to 10.28% in the post-Lehman period (778 days); the increase is milder for S&P 500, from 2.77% to 3.02%.

5.2 Conditional Mean, Variance and Correlation Models

The estimation results and diagnostics of the considered models over the entire sample period with $T = 3491$ days are reported in Table 3. Although these estimates are not used for forecasting, since all models are re-estimated for each rolling window sample (with $P = 2713$ days), the reported estimation results provide insights about the ability of the models to capture various stylized facts.

[Insert Table 3 around here]

The model estimates in Panel A, top exhibit, referring to the overnight period suggest: i) asymmetry in the conditional variance (leverage effect) as borne out by a significantly positive parameter γ_1, ii) fat-tailedness in standardized overnight returns as reflected in a small dof parameter η_0, iii) the sign of $ln(\xi_o)$ indicates that the distribution of innovations to the overnight portfolio returns is negatively skewed. These findings justify the skewed Student-t density assumption for the innovation process. The sufficient condition for stationarity, $\gamma_1 < 2(1 - \alpha_1 - \beta_1)$, is satisfied for both indices.

The model estimates in Panel A, middle exhibit, and Panel C, both of which refer to the daytime period, suggest: i) asymmetry in volatility as borne out by a significantly negative τ_2 parameter in Panel A and significantly negative τ_2 and τ_4 in Panel C, ii) fat-tailedness but less so than in overnight returns as borne out by a larger dof parameter η_d of the residual Student-t density, iii) more marked negative skewness as suggested by the parameter $\ln(\xi_d)$ of the Student-t density, iv) the long-memory parameter d confirms the slow hyperbolic autocorrelation decay of logarithmic realized volatility; in

all cases, it is statistically different from 1/2. For all models, the asymmetry in daytime volatility and the fat-tailedness of daytime innovations appear stronger for the S&P 500 than for the Russell 2000. The long-memory parameter $d < 0.5$ suggests very strong persistence of daytime volatility.

Likewise, the estimation results in Panel B referring to the daily close-to-close period suggest a stronger response of volatility to bad news in the previous period than to good news. The model estimates reported in the bottom exhibit of Panel A indicate that the sufficient stationarity condition, $\alpha_1 + \beta_1 < 1$, is satisfied for the dynamic conditional correlation.

5.3 VaR Predictions Combining Overnight and Intraday Information

We next proceed with the out-of-sample VaR forecasting exercise using a rolling estimation scheme. The first VaR forecast is based on the estimator of the model parameter vector θ_0 formed using information from day 1 to P, denoted $\widehat{\theta}_P$, the second forecast is based on the estimator $\widehat{\theta}_{P+1}$ formed using observations 2 through $P+1$, and so forth. The adjusted RV measures that require parameter estimates (RV^{SC} and RV^{HL}) are calculated in a way that is consistent with the out-of-sample nature of the forecasting exercise. More specifically, the weights needed to obtain, say, the first window $\{RV_t^{HL}\}_{t=1}^{t=P}$ are based on information up to day P, the second window $\{RV_t^{HL}\}_{t=2}^{t=P+1}$ exploits information from day 2 through day $P+1$, and so forth; likewise for RV^{SC}. The fixed-length of the rolling window is $P = 2713$ days and hence, the out-of-sample period contains $n = 778$ days ($T \equiv P + n = 3491$ days).

We present graphically the backtesting results over 279 rolling out-of-sample windows of fixed length $n = 500$ days starting on 02/09/2008. Thus each backtesting outcome corresponds roughly to a 2-year trading period. As these evaluation windows are rolled forward the practical implication is that more observations from the very volatile post-Lehman period are incorporated in the model parameter estimation. Figures 2 and 3 present for the S&P 500 and Russell 2000 indices, respectively, backtesting p-values for the three VaR models, subsumed in equations (19), (20) and (21). The graphs for the bundling approach are based on the overnight-adjusted RV^{HL} measure.[24]

[24]The graphs for RV^{SC} and RV^{+ON} are omitted for space constraints since they are qualitatively similar, as borne out by the summary set out in Table 4 below. Detailed results are available from the authors upon request.

[Insert Figures 2 and 3 around here]

We compare the three parametric VaR models – bivariate, bundling and ex post – from the perspective of correct out-of-sample VaR specification through the dynamic quantile (DQ) test of Engle and Manganelli (2004) and a refinement based on the dynamic binary (DB) probit model. The probit-transformed intercept $\Phi(\beta_0)$ provides the empirical conditional coverage rate or the probability that the actual losses are greater than the VaR limit conditional on the information (regressor) set.

The backtesting p-values reveal that all 3 modeling approaches are unsuccessful in producing VaR forecasts that satisfy (22) in the initial post-Lehman days. In particular, the middle graphs for both Russell 2000 (Figure 2) and S&P 500 (Figure 3) show that it takes longer for the VaRs obtained with the bundling approach to start passing the backtesting (i.e., p-values above the 0.05 significance level threshold). As more and more of the inmediate post-Lehman (very volatile) days are incorporated into the in-sample estimation windows, the VaR backtesting results notably improve.[25] Thus, changes in market conditions can have an important influence on VaR model performance. The 1980-2011 analysis of VaR forecasts in Frey et al. (2012) leads to a similar observation despite important methodological differences with the present study; they forecast the entire return density via Markov Chain Monte Carlo (MCMC) methods based on exponentially (non)affine stochastic volatility models.

Summary information on the DQ and DB probit backtesting outcome obtained dynamically over 279 out-of-sample windows is provided in Table 4. Panel A1 reports the average outcome $\frac{\sum_{j=1}^{J} I(p_j<0.05)}{J}$ with $J = 279$, where $I(p_j < 0.05)$ is the indicator function equal to 1 if the backtesting was failed for window j at the 5% significance level. Thus the reported figures represent percentage of rejections. Next we compute the rejection rates similarly over the first $1/3J$ windows that capture the aftermath

[25] In unreported graphs (available from the authors) we plotted the backtesting results over recursive out-of-sample windows starting from $n = 778$ days which covers the entire out-of-sample period from September 2, 2008 to September 31, 2011 and gradually dropping one day at a time until $n = 500$ days which covers the period from October 8, 2009 to September 31, 2011. These graphs have the advantage that n/P gradually falls over backtesting windows which reduces the impact of estimation noise. On the other hand, the reported graphs have the advantage of making the comparison over time more informative by keeping n fixed so that the results are not influenced by small sample size and power properties of the tests. Both types of graphs give qualitatively the same results.

of post-Lehman debacle (last out-of-sample window begins on January 13, 2009) and the remaining $2/3J$ windows that represents a relatively calmer post-Lehman period. Panels A2 and A3 report the rejection rates corresponding to these two subperiods, respectively. This allows us to compare, on the basis of simple snapshot figures, the overall forecast accuracy of the three tail risk models.

[Insert Table 4 around here]

The summary frequency of backtesting rejections over all 279 out-of-sample windows (Panel A1) suggests that the *bivariate* approach excels for Russell 2000 whereas in contrast the *ex post* approach produces less rejections for S&P 500. Moreover, the bivariate and ex post VaR modeling approaches clearly dominate the horserace for both small cap (Russell 2000) and large cap (S&P 500) portfolios. This is an important finding given that the bundled estimators have become the de facto way of dealing with the overnight period in empirical high-frequency studies. The relative lack of success of the bundling VaR approach to capture tail risk behavior is to some extent unsurprising, given that it implicitly assumes that the overnight squared returns are generated by the same process as the daytime realized volatility. Among the three overnight-adjusted realized volatility measures, the summary backtesting results in Panel A1 of Table 4 indicate that RV^{HL} dominates the other two, RV^{+ON} and RV^{SC}. When considering only the first 1/3 of the windows (Panel A2) which capture the highly volatile aftermath of the post-Lehman debacle, or the last 2/3 of the windows (Panel A3), a relatively calmer period, the previous findings remains unchallenged. Thus the outcome of the horserace of overnight-adjusted VaR modeling approaches appears robust across market conditions.

The empirical conditional coverage (CC) probabilities shown in Figure 4 are also broadly in line with the conclusions drawn above, namely, they are on average closer to the nominal coverage probabilities for S&P 500 with the ex post method, and for the Russell 2000 with the bivariate method. For Russell, the graphs show that the actual CC rates for the 5% VaR obtained through the ex post approach are always well above the nominal coverage probability suggesting that the equity trading VaR limits are always underestimated over the out-of-sample period. In contrast, the actual CC rates

of the 5% VaR obtained through the bivariate modeling approach hover around the 0.05 nominal coverage probability. For the S&P 500 index, the plots suggest that downside tail risk is underestimated for the initial 30 out-of-sample windows reflecting the height of the post-Lehman aftermath (the out-of-sample window crossing the 0.05 nominal coverage line begins on November 17, 2008) and overestimated thereafter. The conditional coverage error or absolute deviation of the empirical CC probabilities from the nominal coverage probability, $\widehat{\alpha}_j - \alpha_j$, is on average larger for the bundling VaR than for the bivariate VaR in the case of Russell, and for the bundling VaR than for the ex post approach in the case of S&P 500, representing underestimated VaRs in both cases.

The summary measures so far discussed give equal weights $w_j \equiv 1/J$ to all backtesting rejections. The loss function of a risk manager may be more sophisticated and assign different importance to different rejections: those associated with large deviations $|\widehat{\alpha}_j - \alpha_j|$ would arguably be more worrisome. Panel B of Table 4 reports the summary outcome of the dynamic backtesting exercise over $J = 279$ rolling windows where the term 'weighted' refers to the fact that each rejection receives a penalty according to the absolute coverage error, i.e. $\sum_{j=1}^{J} \widetilde{w}_j I(p_j < 0.05)$ where $\widetilde{w}_j$ are the standardized weights (so that they add to unity) corresponding to $w_j \equiv |\widehat{\alpha}_j - \alpha_j|$. Finally, the weighting of the rejections in the final two exhibits of the table is asymmetric. Panel C1 presumes that the underprediction of VaR limits is more worrisome (because it entails uncovered losses) for a risk manager than the overprediction of VaR limits (which implies investment opportunity costs) and so the penalty (or weight) is higher at $|\widehat{\alpha}_j - \alpha_j|$ in the former case and lower at $(\widehat{\alpha}_j - \alpha_j)^2$ in the latter case. Panel C2 penalizes more heavily overpredictions than underpredictions instead. These weights are also standardized to sum to unity. Reassuringly, the results are quite robust: the bivariate approach tends to dominate for the Russell 2000 index and the ex post approach for the S&P 500 index.

The contrasting evidence for the two indices could relate to differences in the efficiency of price discovery at the market open for stocks with different trading volumes. The fact that the bivariate approach excels for the Russell 2000 may indicate that the opening quote for this small cap index fails to adequately reflect the information released during non-trading hours. Hence, it is more effective for

the risk manager to step back and set the VaR limits for day t at the closing of day $t-1$ on the basis of conditional mean and (co)variance forecasts for the overnight and daytime return distributions, rather than waiting until the market opens in order to observe the overnight return. For the S&P 500, the opening price seems to convey the information accumulated during non-trading hours more efficiently. This occurs despite the fact that not all stocks open for trading immediately at 9.30 AM, inducing the price staleness and news spillover effects discussed above.

Barclay and Hendershott (2008) show that for Nasdaq stocks, opening prices are more efficient for larger-volume stocks: the more heavily traded the stock, the more price discovery has shifted from the opening price to the pre-open trading period. Importantly, a critical threshold of trading volume is required to shift price discovery from the trading day to the pre-open. Thus, for the most heavily-traded stocks, the opening price conveys more information. Our results are well in line with the findings of Barclay and Hendershott (2008) given the clear difference in results for the two indices. For small-cap stocks, which display smaller trading volumes, the opening price (and hence, the observed overnight return from previous close to open) is less informative, making the bivariate modeling approach more valuable for setting equity trading VaR-based limits.[26]

6 Conclusions

A task faced daily by risk managers is to establish forward-looking mark-to-market loss limits for individual trading desks. Value-at-Risk (VaR) has become a standard risk control tool. In 1996, the Basel II Accord gave regulatory recognition to the widespread use of VaR by banks by introducing a VaR-based capital requirement for positions held for trading intent. Through parametric models, the problem of downside tail risk prediction can be formulated as generating predictions for the conditional

[26] There is evidence that overnight information is disseminated more efficiently on Nasdaq compared to the NYSE due to differences in the opening mechanisms of the exchanges (see Masulis and Shivakumar, 2002 and Greene and Watts, 1996). Both of the indices in our study contain component stocks listed on both exchanges. The marked difference in the results for the two indices suggests that the effect of the trading volume is stronger than the effect of the trading platform. In the latter case, the result could easily be reversed, as more S&P 500 constituents are traded on the NYSE, and more Russell 2000 constituents are traded on Nasdaq.

mean and variance of the return process, which are combined with an α-quantile from a sufficiently general density to approximate departures from Gaussianity.

The present paper makes both theoretical and empirical contributions. We introduce the continuous-time theory to motivate a bivariate approach to obtain VaR limits that involves separate forecasts for overnight volatility, daytime volatility, and the covariance between daytime and overnight returns. A non-zero covariance can be induced by delayed trading or 'price staleness' and 'news spillover' at market open. In our empirical section, we confront this bivariate modeling approach with two other approaches for incorporating overnight information into tail risk prediction. One is the widely-used approach in the empirical high-frequency finance literature of bundling together the overnight and daytime information into single time-series of returns and realized volatility. The other is an ex post approach that generates forecasts at market open once the overnight return is observed and can be incorporated in the information set.

Our empirical illustration reveals that paying separate modeling efforts to the overnight and daytime segments of the day appears beneficial for describing the tail behavior of Russell 2000 returns. In contrast, for the S&P 500 returns, the ex post modeling approach produces the most effective VaR limits. Such a contrast between the two indices may relate to the different levels of trading volume of the constituent stocks. It has been shown in previous research that price discovery at the market open is more efficient for high-volume stocks. Therefore, the large cap stocks in the S&P 500 index may convey enough overnight information in their opening prices to make the index opening levels as a whole quite informative. As a consequence, more accurate VaR-based trading limits are obtained for the S&P 500 portfolio by simply including the observed overnight (squared) returns in the models' conditioning information set as opposed to forecasting the day-ahead overnight return process.

Our VaR backtesting analysis suggests that the bundling approach widely used in academic research to account for the overnight non-trading period can compromise the solvency of trading positions by setting too low trading loss limits. These findings have clear implications for banks' risk managers, other practitioners, and researchers seeking to reliably predict downside tail risk.

References

[1] Ahoniemi, K., & Lanne, M. (2012). Overnight Returns and Realized Volatility. SSRN Working Paper available at: http://ssrn.com/abstract=1945687.

[2] Andersen, T.G., Bollerslev, T., Diebold, F.X. & Ebens, H. (2001). The Distribution of Realized Stock Return Volatility. Journal of Financial Economics, 61, 43–76.

[3] Andersen, T.G., Bollerslev, T. & Huang, X. (2011). A reduced form framework for modeling volatility of speculative prices based on realized variation measures. Journal of Econometrics 160, 176-189

[4] Angelidis, T., & Degiannakis, S. (2008a). Volatility Forecasting: Intra-Day Versus Inter-Day Models. Journal of International Financial Markets, Institutions and Money, 18, 449–465.

[5] Angelidis, T., & Degiannakis, S. (2008b). Econometric Modeling of Value-at-Risk. In New Econometric Modeling Research, ed. William N. Toggins, Nova Science Publishers Inc.

[6] Barclay, M.J. & Hendershott, T. (2008). A Comparison of Trading and Non-Trading Mechanisms for Price Discovery. Journal of Empirical Finance, 15, 839–849.

[7] Barndorff-Nielsen, O.E. and Shephard, N. (2002), Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models, Journal of the Royal Statistical Society, Series B, 64, 253-280.

[8] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., & Sheppard, N. (2008). Designing Realized Kernels to Measure the Ex-Post Variance of Equity Prices in the Presence of Noise. Econometrica, 76, 1481–1536.

[9] Basel Committee on Banking Supervision (1996). Supervisory Framework for the Use of "Backtesting" in Conjunction with the Internal Models Approach to Market Risk Capital Requirements. Bank for International Settlements, January 1996.

[10] Becker, R., Clements, A.E., & White, S.I. (2007). Does Implied Volatility Provide Any Information Beyond That Captured in Model-Based Volatility Forecasts? Journal of Banking & Finance, 31, 2535–2549.

[11] Berkowitz, J., Christoffersen, P., & Pelletier, D. (2011). Evaluating Value-at-Risk Models with Desk-Level Data. Management Science, 57, 2213–2227.

[12] Blair, B.J., Poon, S-H. & Taylor, S.J. (2001). Forecasting S&P 100 Volatility: the Incremental Information Content of Implied Volatilities and High-frequency Index Returns. Journal of Econometrics, 105, 5–26.

[13] Bollerslev, T., Tauchen, G. & Zhou, H. (2009). Expected Stock Returns and Variance Risk Premia. Review of Financial Studies, 22, 4463–4492.

[14] Brownlees, C. T., & Gallo, G. M. (2010). Comparison of Volatility Measures: A Risk Management Perspective. Journal of Financial Econometrics, 8, 29–56.

[15] Candelon, B., Hurlin, C., Colletaz, G. & Tokpavi, S. (2011). Backtesting Value-at-Risk: A GMM Duration-Based Test. Journal of Financial Econometrics, 9, 314-343.

[16] Chan, K., Chan, K.C. & Karolyi, G.A. (1991). Intraday Volatility in the Stock Index and Stock Index Futures Market. Review of Financial Studies, 4, 657–684.

[17] Chevallier, J., Ielpo, F., & Sevi, B. (2011). Do Jumps Help in Forecasting the Density of Returns? SSRN Working Paper available at: http://ssrn.com/abstract=1585822.

[18] Christoffersen, P. (1998). Evaluating Interval Forecasts. International Economic Review 39, 841–108.

[19] Christoffersen, P., Hahn, J., & Inoue, A. (2001). Testing and Comparing Value-at-Risk Measures. Journal of Empirical Finance, 8, 325–352.

[20] Clements, M.P., Galvão, A.B., & Kim, J.H. (2008). Quantile Forecasts of Daily Exchange Returns from Forecasts of Realized Volatility. Journal of Empirical Finance, 15, 729–750.

[21] Cliff, M., Cooper, M.J., & Gulen, H. (2009). Return Differences between Trading and Non-trading Hours: Like Night and Day. SSRN Working Paper 1004081.

[22] Corsi, F. (2004). A Simple Long Memory Model of Realized Volatility, University of Southern Switzerland, mimeo.

[23] Corsi, F., Mittnik, S., Pigorsch, C. & Pigorsch, U. (2008). The Volatility of Realized Volatility. Econometric Reviews, 27, 46–78.

[24] de Gooijer, J.G., Diks, C.G.H. & Gatarek, L.T. (2009). "Information Flows around the Globe: Predicting Opening Gaps from Overnight Foreign Stock Price Patterns," Tinbergen Institute Discussion Papers 09-107/4, Tinbergen Institute.

[25] de Pooter, M., Martens, M. & van Dijk, D. (2008). Predicting the Daily Covariance Matrix for S&P 100 Stocks Using Intraday Data - But Which Frequency to Use? Econometric Reviews, 27, 199–229.

[26] Dumitrescu, E.I., Hurlin, C. & Pham, V. (2012). Backtesting Value-at-Risk: From Dynamic Quantile to Dynamic Binary Tests. University of Orleans, Mimeo.

[27] Engle, R.F. (2002). Dynamic Conditional Correlation - A Simple Class of Multivariate GARCH Models. Journal of Business & Economic Statistics, 20, 339–350.

425–437.

[28] Engle, R.F. & Kroner, K.F. (1995). Multivariate Simultaneous Generalized ARCH. Econometric Theory, 11, 122–150.

[29] Engle, R.F. & Manganelli, S. (2004). CAViaR Conditional Autoregressive Value-at-Risk by Regression Quantiles. Journal of Business & Economic Statistics, 22, 367–381.

[30] Fleming, J., Kirby, C., & Ostdiek, B. (2003). The economic value of volatility timing using realized volatility. Journal of Financial Economics, 67, 473-509.

[31] French, K.R., & Roll, R. (1986). Stock Return Variances: The Arrival of Information and the Reaction of Traders. Journal of Financial Economics, 17, 5–26.

[32] Frey, R., Rodrigues, P., & Seeger, N. (2012). Out-of-Sample Performance of Jump-Diffusion Models for Equity Indices: What the Financial Crisis was Good for. SSRN Working Paper 2022909.

[33] Fuertes, A.-M., Izzeldin, M., & Kalotychou, E., (2009) On Forecasting Daily Stock Volatility: The Role of Intraday Information and Market Conditions. International Journal of Forecasting, 25, 259–281.

[34] Fuertes, A.-M. & Olmo, J. (2012). Optimally Harnessing Inter-Day and Intra-Day Information for Daily Value-at-Risk Prediction. International Journal of Forecasting, forthcoming.

[35] Gaglianone, W.P., Lima, L.R., Linton, O., and Smith, D.R. (2011). Journal of Business and Economic Statistics, 29(1), 150–160.

[36] George, T.J., & Hwang, C-Y. (2001). Information Flow and Pricing Errors: A Unified Approach to Estimation and Testing. Review of Financial Studies, 14, 979–1020.

[37] Ghysels, E., Santa-Clara, P. and Valkanov, R. (2004). The MIDAS Touch: Mixed Data Sampling Regression Models, Chapel Hill, mimeo.

[38] Giacomini, R., & Komunjer, I. (2005). Evaluation and Combination of Conditional Quantile Forecasts. Journal of Business & Economic Statistics, 23, 416–431.

[39] Giot, P., & Laurent, S. (2004). Modeling Daily Value-at-Risk Using Realized Volatility and ARCH Type Models. Journal of Empirical Finance, 11, 379–398.

[40] Glosten, L., Jagannathan, R. & Runkle, D. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. Journal of Finance, 48, 1779–1801.

[41] Gourieroux, C., & Jasiak, J. (2006). Dynamic Quantile Models. Working Paper 2006-4, Department of Economics, York University.

[42] Greene, J.T. & Watts, S.G. (1996). Price Discovery on the NYSE and the Nasdaq: The Case of Overnight and Daytime News Releases. Financial Management, 25, 19–42.

[43] Hansen, P.R., & Lunde, A. (2005). A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data. Journal of Financial Econometrics, 3, 525–554.

[44] Hasbrouck, J. (1991). Measuring the Informatiom Content of Stock Trades. Journal of Finance, 46, 179-207.

[45] Hasbrouck, J. (2012). High Frequency Quoting: Measurement, Detection and Interpretation. Keynote Paper at 5th EMG-ESRC Workshop on the Microstructure of Financial Markets, Cass Business School, May 2012.

[46] Hecq, A., Laurent, S. & Palm, F. (2012). Common Intraday Periodicity. Journal of Financial Econometrics, forthcoming.

[47] Hong, H. & Wang, J. (2000). Trading and Return under Periodic Market Closures. Journal of Finance 55, 297–354.

[48] Jones, C., Kaul, G., & Lipson, M. (1994). Information, Trading and Volatility. Journal of Financial Economics, 36, 127–154.

[49] Koenker, R. & Xiao, Z. (2006). Quantile Autoregression. Journal of the American Statistical Association, 101, 980–990.

[50] Koopman, S.J., Jungbacker, B. & Hol, E. (2005). Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realised, and Implied Volatility Measurements. Journal of Empirical Finance, 12, 445–475.

[51] Kroner, K.F. & Ng, V.K. (1998). Modeling Asymmetric Comovements of Asset Returns. Review of Financial Studies, 11, 817–844.

[52] Kuester, K., Mittnik, S., & Paolella, M.S. (2006) Value-at-Risk Prediction: A Comparison of Alternative Strategies. Journal of Financial Econometrics, 4, 53–89.

[53] Lin, W-L., Engle, R.F. & Ito, T. (1994). Do Bulls and Bears Move Across Borders? International Transmission of Stock Returns and Volatility. Review of Financial Studies, 7, 507–538.

[54] Liu, C., & Maheu, J. (2009). Forecasting Realized Volatility: A Bayesian Model-Averaging Approach. Journal of Applied Econometrics, 24, 709–733.

[55] Lockwood, L.J., & Lin, S. (1990). An Examination of Stock Market Return Volatility During Overenight and Intraday Periods 1964-1989. Journal of Finance, 45, 591-601.

[56] Martens, M. (2002). Measuring and Forecasting S&P 500 Index-Futures Volatility Using High-frequency Data. Journal of Futures Markets, 22, 497–518.

[57] Martens, M., van Dijk, D., & de Pooter, M. (2009). Forecasting S&P 500 Volatility: Long Memory, Level Shifts, Leverage Effects, Day of the Week Seasonality and Macroeconomic Announcements. International Journal of Forecasting, 25, 282–303.

[58] Masulis, R.W. & Ng, V.K. (1995). Overnight and Daytime Stock-Return Dynamics on the London Stock Exchange: The Impacts of "Big Bang" and the 1987 Stock-Market Crash. Journal of Business & Economic Statistics, 13, 365–378.

[59] Masulis, R.W. & Shivakumar, L. (2002). Does Market Structure Affect the Immediacy of Stock Price Responses to News? Journal of Financial and Quantitative Analysis, 37, 617–648.

[60] McAleer, M., & Medeiros, M.C. (2008). Realized Volatility: A Review. Econometric Reviews, 27, 10–45.

[61] Noureldin, D., Shephard N. & Sheppard K. K. (2011). Multivariate high-frequency-based volatility (HEAVY) models. Journal of Applied Econometrics, Forthcoming.

[62] Oldfield, G.S. & Rogalski, R.J. (1980). A Theory of Common Stock Returns Over Trading and Non-Trading Periods. Journal of Finance, 35, 729-751.

[63] Patton, A.J. (2011). Data-Based Ranking of Realised Volatility Estimators. Journal of Econometrics, 161, 284–303.

[64] Ronen, T. (1997). Tests and Properties of Variance Ratios in Microstructure Studies. Journal of Financial and Quantitative Analysis, 32, 183–204.

[65] Shephard N. & Sheppard K. K. (2010). Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. Journal of Applied Econometrics 25: 197231.

[66] Spurlin, W.P., Van Ness, B.F. & Van Ness, R.A. (2008). Open Volume and Time to Open on Option Expiration Days. International Review of Economics and Finance, 17, 245–257.

[67] Stoll, H.R. & Whaley, R.E. (1990a). Stock Market Structure and Volatility. Review of Financial Studies, 3, 37–71.

[68] Stoll, H.R. & Whaley, R.E. (1990b). The Dynamics of Stock Index and Stock Index Futures Returns. Journal of Financial and Quantitative Analysis, 25, 441–467.

[69] Taylor, N. (2007). A Note on the Importance of Overnight Information in Risk Management Models. Journal of Banking & Finance, 31, 161–180.

[70] Thomakos, D.D. & Wang, T. (2003). Realized Volatility in the Futures Markets. Journal of Empirical Finance, 10, 321–353.

[71] Wu, L. (2011). Variance Dynamics: Joint Evidence from Options and High-Frequency Returns. Journal of Econometrics, 160, 280–287.

Appendix A

Proof of Proposition: We assume that the log-price process is driven by the following stochastic differential equation:

$$dp_t = \{\mu_{0,t}dt + \sigma_{o,t}dW_{o,t}\}\,1(C_{t-1} < t \leq O_t) + \{\mu_{d,t}dt + \sigma_{d,t}dW_{i,t}\}\,1(O_t < t \leq C_t) \quad 0 \leq t \leq T.$$

Integrating the expression, we have

$$\Delta p_t = p_{C_t} - p_{C_{t-1}} = \int_{C_{t-1}}^{O_t} \mu_{0,s}ds + \int_{C_{t-1}}^{O_t} \sigma_{o,s}dW_{o,s} + \int_{O_t}^{C_t} \mu_{d,s^*} + \int_{O_t}^{C_t} \sigma_{d,s^*}dW_{d,s^*}.$$

Accordingly, the variance in the log-price process can be expressed as $\Delta p_t = r_{o,t} + r_{d,t}$, with

$$r_{o,t} = \int_{C_{t-1}}^{O_t} \mu_{0,s}ds + \int_{C_{t-1}}^{O_t} \sigma_{o,s}dW_{o,s},$$

and

$$r_{d,t} = \int_{O_t}^{C_t} \mu_{d,s^*} + \int_{O_t}^{C_t} \sigma_{d,s^*}dW_{d,s^*}.$$

This equation can be decomposed as $r_{o,t} = \mu_{0,t} + \varepsilon_{o,t}$, with $\mu_{0,t} \equiv \int_{C_{t-1}}^{O_t} \mu_{o,s}ds$, and $\varepsilon_{o,t} = \int_{C_{t-1}}^{O_t} \sigma_{o,s}dW_{o,s}$, and $r_{d,t} = \mu_{d,t} + \varepsilon_{d,t}$, with $\mu_{d,t} \equiv \int_{O_t}^{C_t} \mu_{d,s^*}ds^*$ and $\varepsilon_{d,t} = \int_{O_t}^{C_t} \sigma_{d,s^*}dW_{d,s^*}$.

Let $IV_{o,t} \equiv V(\int_{C_{t-1}}^{O_t} \sigma_{o,s}dW_{o,s})$, $IV_{d,t} \equiv V(\int_{O_t}^{C_t} \sigma_{d,s^*}dW_{d,s^*})$ and $IV_{o,d,t} = Cov(r_{o,t}, r_{d,t})$. To show the decomposition in the proposition, we compute first $IV_{o,t}$. We use a partition $C_{t-1} < s_1 < s_2 < \ldots < s_n < O_t$, then

$$V(\int_{C_{t-1}}^{O_t} \sigma_{o,s}dW_{o,s}) = V\left(\lim_{n\to\infty}\sum_{i=1}^{n}\sigma_{o,s_i}(W_{o,s_i} - W_{o,s_{i-1}})\right) = \lim_{n\to\infty}\sum_{i=1}^{n}\sigma_{o,s_i}^2 E[\left(W_{o,s_i} - W_{o,s_{i-1}}\right)^2] = \int_{C_{t-1}}^{O_t} \sigma_{o,s}^2 ds,$$

by the properties of the Brownian motion. Similar algebra shows that $IV_{d,t} \equiv \int_{O_t}^{C_t} \sigma_{d,s^*}^2 ds^*$. Finally, to compute $IV_{o,d,t}$ we use that

$$IV_{o,d,t} \equiv Cov(r_{o,t}, r_{d,t}) = E[r_{o,t}r_{d,t}] - E[r_{o,t}]E[r_{d,t}].$$

Note that $E[r_{o,t}] = \mu_{o,t}$, since $E[\varepsilon_{o,t}] = E[\int_{C_{t-1}}^{O_t} \sigma_{o,s}dW_{o,s}] = \int_{C_{t-1}}^{O_t} E[\sigma_{o,s}dW_{o,s}] = 0$; $\sigma_{o,t}$ is a deterministic process and it follows that $E[\sigma_{o,s}dW_{o,s}] = \sigma_{o,s}E[dW_{o,s}] = 0$, by the properties of the Brownian motion. Similarly, we obtain that $E[r_{d,t}] = \mu_{d,t}$ and then

$$E[r_{o,t}r_{d,t}] = \mu_{o,t}\mu_{d,t} + \mu_{o,t}E[\varepsilon_{d,t}] + E[\varepsilon_{o,t}]\mu_{d,t} + E[\varepsilon_{o,t}\varepsilon_{d,t}],$$

which is equal to $\mu_{o,t}\mu_{d,t}+E[\varepsilon_{o,t}\varepsilon_{d,t}]$. Finally, by the Law of Iterated Expectations constructed over the sigma-algebra $\Im_{O_t}$ we have $E[\varepsilon_{o,t}\varepsilon_{d,t}] = E[\varepsilon_{o,t}E[\varepsilon_{d,t}|\Im_{O_t}]] = 0$. Therefore, $IV_{o,d,t} = 0$ by construction, and $IV_t = IV_{o,t} + IV_{d,t}$.

Table 1. Distribution of overnight and daytime returns.

The table summarizes the empirical distribution of daily overnight and daytime returns (Panel A) and the corresponding volatility measure given by the squared return (Panel B) over the post-Lehman period from September 2, 2008 to September 31, 2011 and the preceding period from November 12, 1997 to August 29, 2008. All returns are in percentages. The correlation structure section reports autocorrelations between the overnight return on day *t* and day *t*-1, between the daytime return on day *t* and day *t*-1, and the cross-correlation between the overnight return on day *t* and daytime return on day *t*. Q_{20}(ACF) is the Ljung-Box statistic for the null of no autocorrelation up to 20 lags. *, ** and *** indicate significance at the 10%, 5% or 1% levels.

	In-sample: Nov 1997 - Aug 2008				Out-of-sample: Sep 2008 - Sep 2011			
	Russell		S&P 500		Russell		S&P 500	
Summary statistics	$r_{o,t}$ night	$r_{d,t}$ day	$r_{o,t}$ night	$r_{d,t}$ day	$r_{o,t}$ night	$r_{d,t}$ day	$r_{o,t}$ night	$r_{d,t}$ day
				Panel A: returns				
Mean	0.020	0.000	0.007	0.005	-0.004	-0.014	-0.012	-0.004
Median	0.005	0.040	0.003	0.047	0.007	0.138	0.000	0.095
StDev	0.234	1.271	0.190	1.092	0.572	2.136	0.249	1.780
Skewness	-0.525	-0.074	-0.249	-0.031	-0.283	-0.358	-0.254	-0.362
Kurtosis	19.80	4.026	14.99	5.600	6.930	6.015	7.774	8.586
Correlation structure:								
$r_{o,t-1}$	-0.089 ***	-0.028	-0.092 ***	0.001	-0.149 ***	-0.064 *	-0.118 ***	-0.026
$r_{d,t}$	0.162 ***		0.202 ***		0.273 ***		0.390 ***	
$r_{d,t-1}$	0.041 **	0.015	-0.011	-0.038 **	-0.040	-0.085 **	-0.048	-0.122
Q_{20} (AC)	60.6	38.64	60.78	35.21	36.62	28.07	37.12	52.53
				Panel B: squared returns				
Mean (hourly)	0.003	0.248	0.002	0.183	0.019	0.701	0.004	0.487
Median	0.006	0.597	0.004	0.361	0.096	1.120	0.011	0.457
StDev	0.237	2.810	0.135	2.556	0.797	10.22	0.162	8.724
Skewness	14.16	4.661	11.58	7.089	8.880	5.307	6.353	5.905
Kurtosis	254.6	39.42	196.5	89.15	117.3	43.48	57.41	46.17
Q_{20} (AC)	156.4	1697	247.9	1058	186.8	1151	441.3	1107

Table 2. Distribution of realized variance.

The table summarizes the empirical distribution of *daytime* realized volatility computed as the sum of 5-minute squared returns from market open at 9:30am to close at 4:00pm (RV) and three overnight-adjusted versions that span the entire 24hr period as described in Section 2.2. RV^{+ON} adds the unweighted squared overnight return to the baseline realized variance; RV^{SC} is a rescaled version of the baseline realized variance using the scaling factor in Eq. (7); RV^{HL} is Hansen and Lunde's (2005) weight optimized realized variance. Figures are in percentages.

Summary	In-sample: Nov 1997 - Aug 2008				Out-of-sample: Sep 2008 - Sep 2011			
statistics	RV_t	RV_t^{SC}	RV_t^{+ON}	RV_t^{HL}	RV_t	RV_t^{SC}	RV_t^{+ON}	RV_t^{HL}
				Panel A: Russell 2000				
Mean	1.091	1.128	1.145	1.14484	3.735	4.003	4.062	4.062
Median	0.658	0.679	0.698	0.69699	1.590	1.704	1.867	1.948
St Dev / Mean	1.257	1.257	1.240	1.240	1.663	1.663	1.598	1.540
Skewness	4.368	4.368	4.253	4.24478	4.208	4.208	4.082	4.012
Kurtosis	34.78	34.78	32.73	32.5309	26.17	26.17	25.14	24.46
d	0.476	0.476	0.484	0.476	0.493	0.493	0.490	0.495
				Panel B: S&P 500				
Mean	1.189	1.225	1.225	1.225	3.295	3.359	3.357	3.357
Median	0.736	0.759	0.761	0.761	1.109	1.131	1.132	1.202
St Dev / Mean	1.272	1.272	1.261	1.258	1.939	1.939	1.928	1.865
Skewness	4.584	4.584	4.453	4.353	4.359	4.359	4.371	4.664
Kurtosis	37.35	37.35	35.29	33.56	26.75	26.75	27.00	32.07
d	0.492	0.492	0.493	0.493	0.495	0.495	0.491	0.495

Table 3. Conditional mean, variance and correlation model estimates.

The table reports parameter estimates and standard errors in parenthesis over the entire sample period. Panel A, top panel, pertains to the overnight AR and TGARCH equations (13a) and (13b), respectively; middle panel to daytime AR and ARFIMA equations (14a) and (14b), respectively; bottom panel to correlation equation (15). Panel B corresponds to AR equation (16a) for the close-to-close return and ARFIMA equation (16b) for RV^{HL}. Panel C corresponds to the daytime AR and ARFIMA equations (17a) and (17b), respectively. Intercepts are not reported to preserve space.

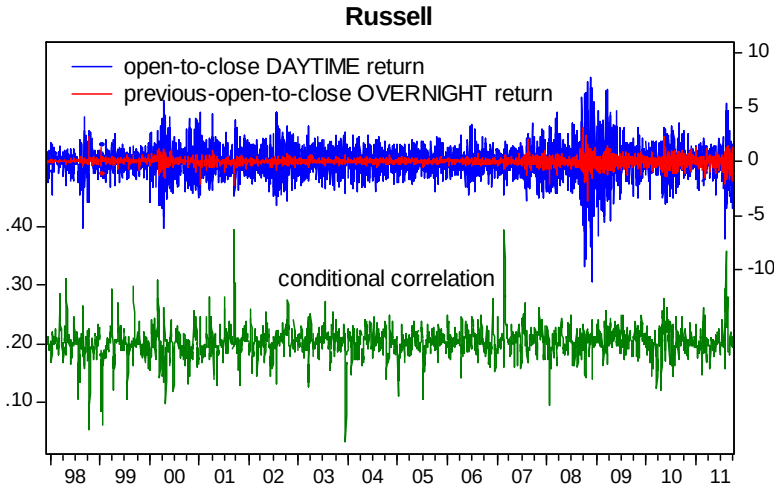
Panel A: Bivariate approach				
Parameters	Russell 2000		S&P 500	
AR-TGARCH model (*overnight*)				
a_1	-0.0031	(0.0021)	-0.0074	(0.0018)
b_1	-0.0659	(0.0144)	-0.0692	(0.0140)
α_1	0.0900	(0.0171)	0.0720	(0.0127)
β_1	0.8620	(0.0169)	0.9080	(0.0124)
γ_1	0.0020	(0.0006)	0.0011	(0.0004)
$\ln\xi_o$	-0.0274	(0.0144)	-0.0340	(0.0156)
η_o	2.9625	(0.1212)	2.9017	(0.1679)
AR-ARFIMA model (daytime)				
a_1	0.0425	(0.0189)	-0.0407	(0.0167)
b_1	-0.0988	(0.0917)	0.0790	(0.1081)
τ_1	0.3059	(0.0194)	0.3576	(0.0207)
τ_2	-0.8021	(0.0308)	-0.9397	(0.0325)
θ	-0.5482	(0.0149)	-0.5386	(0.0127)
d	0.4933	(0.0088)	0.4966	(0.0046)
σ^2	1.1431	(0.0736)	1.0666	(0.0620)
$\ln\xi_d$	-0.1310	(0.0240)	-0.1138	(0.0196)
η_d	14.1627	(3.0623)	6.5079	(0.7316)
DCC model (*overnight* vs. *daytime*)				
α	0.0146	(0.0105)	0.0049	(0.0085)
β	0.6828	(0.2402)	0.7716	(0.1480)

Panel B: Bundling approach				
Parameters	Russell 2000		S&P 500	
AR-ARFIMA model (*close-to-close*)				
c_1	0.0466	(0.0206)	-0.0428	(0.0183)
τ_1	0.3092	(0.0195)	0.3568	(0.0207)
τ_2	-0.8086	(0.0310)	-0.9382	(0.0325)
θ	-0.5497	(0.0148)	-0.5383	(0.0128)
d	0.4934	(0.0087)	0.4966	(0.0046)
σ^2	1.2091	(0.0975)	1.0443	(0.0675)
$\ln\xi$	-0.1042	(0.0268)	-0.0806	(0.0233)
η	21.1876	(7.5066)	8.8438	(1.5029)

Panel C: Ex post approach				
Parameters	Russell 2000		S&P 500	
AR-ARFIMA model (*daytime*)				
a_0	-0.0291	(0.0207)	-0.0040	(0.0154)
a_1	0.0440	(0.0184)	-0.0315	(0.0162)
b_0	1.0553	(0.0893)	1.6801	(0.1113)
τ_0	-0.8061	(0.0247)	-0.8067	(0.0228)
τ_1	0.2313	(0.0186)	0.3115	(0.0196)
τ_2	-0.6453	(0.0300)	-0.8424	(0.0310)
τ_3	1.0250	(0.0797)	1.7658	(0.1212)
τ_4	-2.4760	(0.1214)	-3.9167	(0.1868)
d	0.4829	(0.0186)	0.4949	(0.0070)
σ^2	0.9752	(0.0692)	0.9255	(0.0558)
$\ln\xi_d$	-0.1008	(0.0242)	-0.1273	(0.0211)
η_d	12.1365	(2.2368)	6.6662	(0.7253)

Figure 1. Returns and conditional correlation.

Each figure graphs the conditional TGARCH variance corresponding to the overnight segment of the day, the conditional ARFIMA variance corresponding to the daytime segment of the day, and the dynamic conditional correlation between the standardized overnight and daytime innovations. The sample period runs from November 1997 to September 2011 spanning 3491 days.



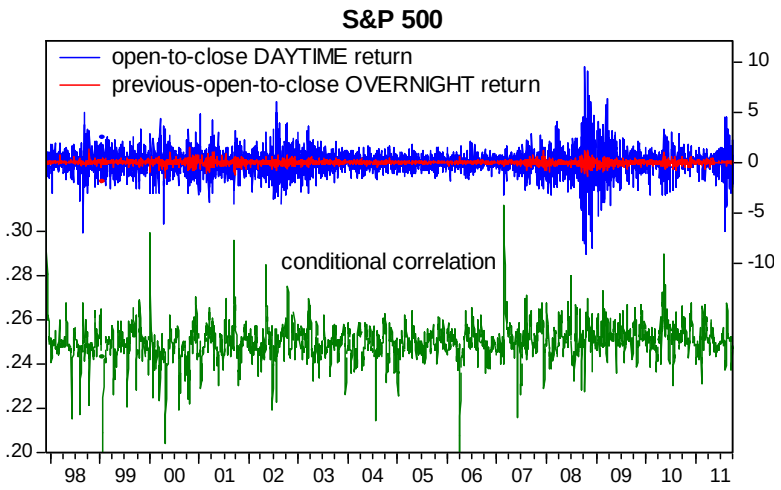
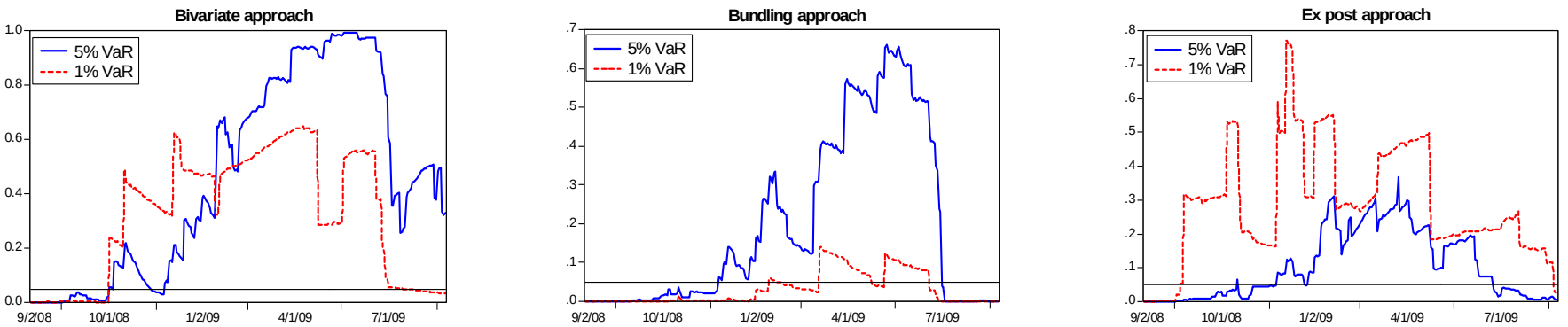


Figure 2. Dynamic backtesting for correct out-of-sample VaR specification: Russell 2000 index

The graphs report *p*-values of the DQ test statistic (24) and the DB probit test statistic (27) for the null hypothesis of correct conditional VaR specification. The first (last) point in each graph represents the *p*-value of the test conducted over a 500-day out-of-sample window starting September 2, 2008 (September 8, 2009). The bottom horizontal line represents the 0.05 significance level for the tests. The bundling approach graphs pertain to the RV^{HL} measure.

Panel A: Dynamic quantile (DQ) test



Panel B: Dynamic binary (DB) probit test

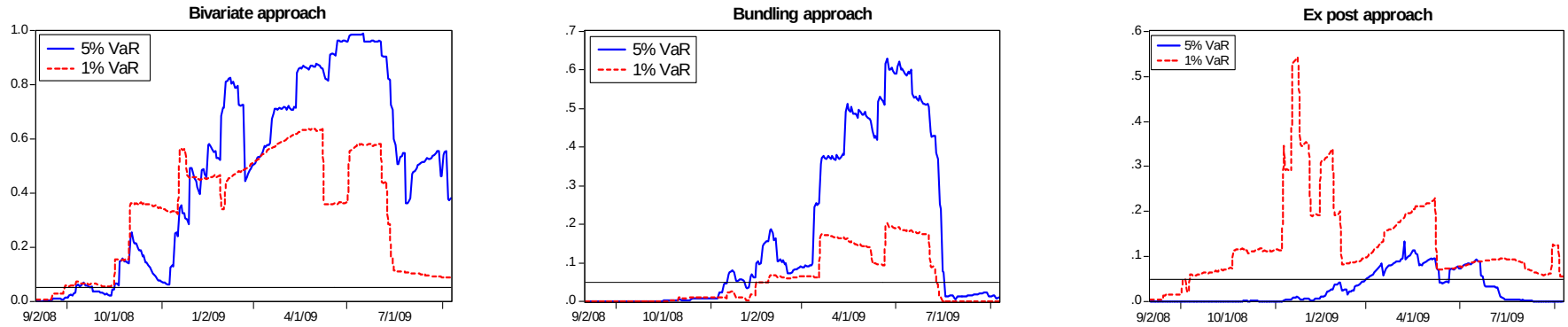
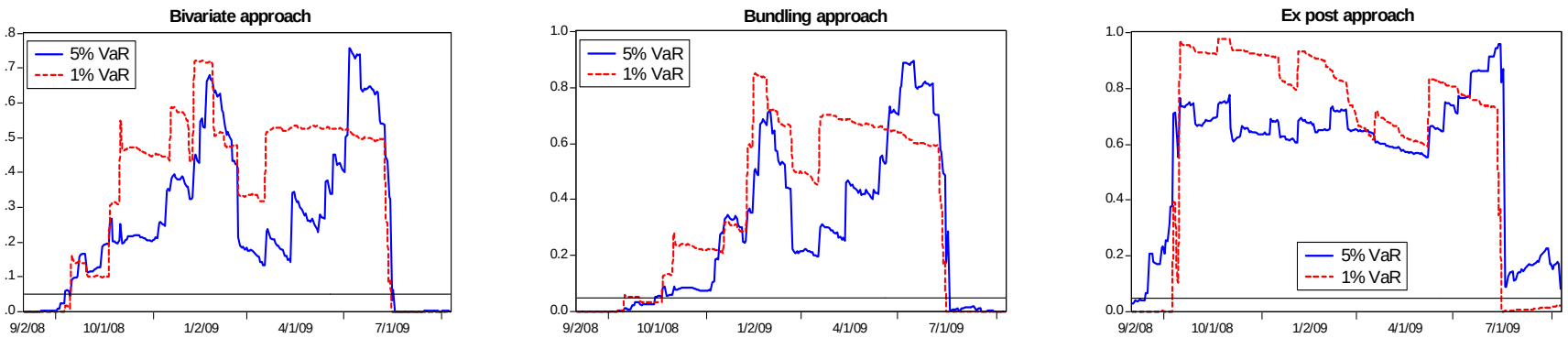


Figure 3. Dynamic backtesting for correct out-of-sample VaR specification: S&P 500 index

The graphs report *p*-values of the DQ test statistic (24) and the DB probit test statistic (27) for the null hypothesis of correct conditional VaR specification. The first (last) point in each graph represents the *p*-value of the test conducted over a 500-day out-of-sample window starting September 2, 2008 (September 8, 2009). The bottom horizontal line represents the 0.05 significance level for the tests. The bundling approach graphs pertain to the RV^{HL} measure.

Panel A: Dynamic quantile (DQ) test



Panel B: Dynamic binary (DB) probit test

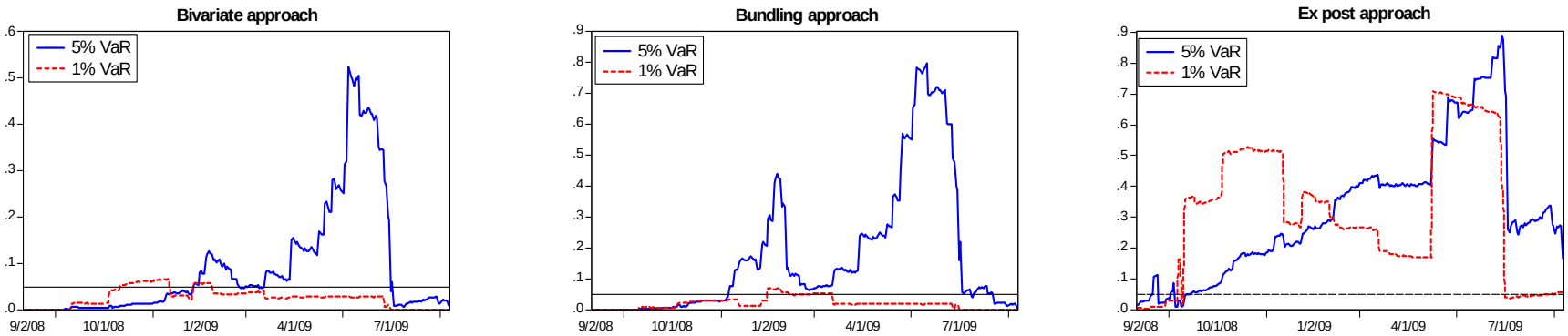
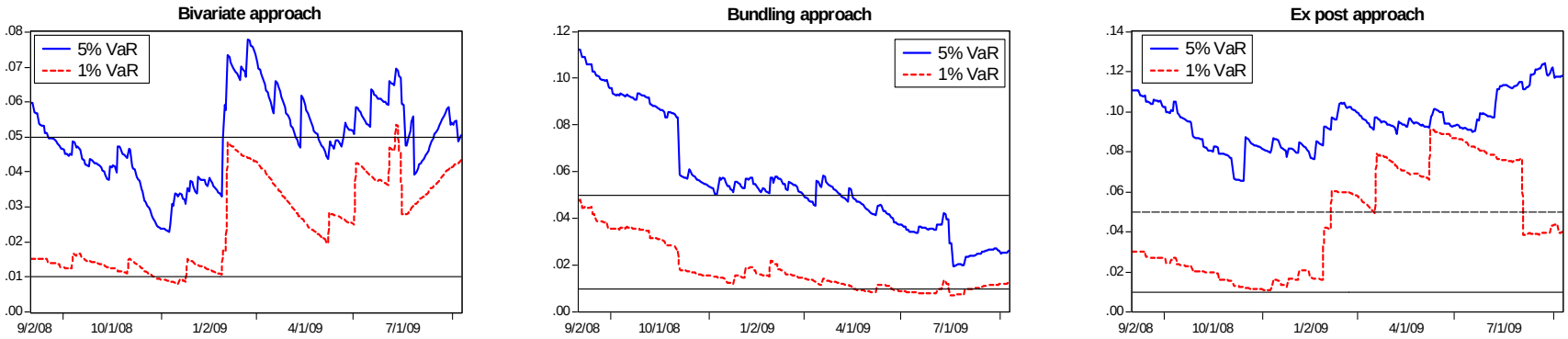


Figure 4. Empirical conditional coverage probabilities.

The graphs plot the empirical conditional coverage probability of the VaR obtained as the probit-transformed intercept from the probit regression (26). The top and bottom horizontal lines plot the nominal coverage probabilities at 0.05 and 0.01. The bundling approach graphs pertain to the RV^{HL} measure.

Panel A: Russell 2000 index



Panel B: S&P 500 index

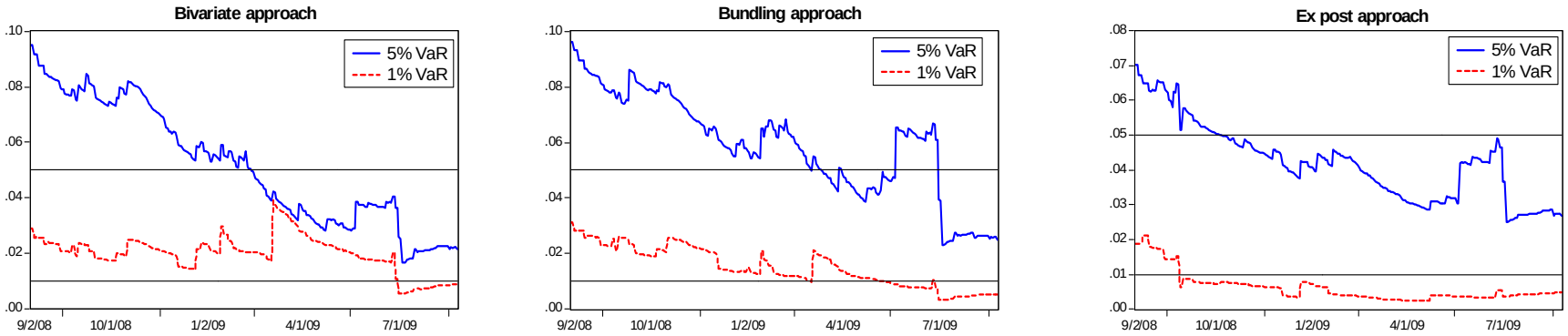


Table 4. Summary average rejection of correct out-of-sample VaR specification.

This table summarises the dynamic quantile (DQ) and dynamic binary (DB) probit test for the null of correct conditional VaR specification, criterion (22), assessed out-of-sample over J=279 rolling windows of 500 days each. Panel A1 reports the rejection frequency over 279 windows; the first window begins September 2, 2008 and the last window begins September 8, 2009. Panel A2 focuses on the first 1/3J windows (highly volatile post-Lehman period; last window begins on January 13, 2009) and Panel A3 focuses on the remaining 2/3J windows. The last three panels report weighted rejection frequencies. Panel B weights each rejection according to the distance of the deviation of the empirical conditional coverage (CC) probability from the nominal coverage. Panel C1 weights more heavily the rejections for which the empirical CC probability is above the nominal coverage (absolute distance) than the rejections for which the empirical CC is below the nominal coverage (squared distance). Panel C2 weights more heavily the rejections for which the empirical CC probability is below the nominal coverage (absolute distance) than the rejections for which the empirical CC is above the nominal coverage (squared distance). Shaded areas report the lowest (weighted) average rejection rate across VaR modeling approaches for each test.

	Bivariate	RV^{SC}	RV^{+ON}	RV^{HL}	Ex post	Bivariate	RV^{SC}	RV^{+ON}	RV^{HL}	Ex post
	Russell 2000					S&P 500				
Panel A1: Equally-weighted J backtesting windows (J=279)										
DQ test (5% VaR)	*0.222*	0.541	0.462	0.462	0.480	0.237	0.620	0.441	0.315	*0.032*
DQ test (1% VaR)	0.291	1.000	0.749	0.742	*0.097*	0.251	0.910	0.910	0.305	*0.237*
Probit test (5% VaR)	*0.151*	0.581	0.487	0.487	0.741	0.566	0.846	0.541	0.387	*0.097*
Probit test (1% VaR)	*0.068*	0.767	0.591	0.577	0.097	0.835	1.000	1.000	0.878	*0.194*
Panel A2: Equally-weighted 1/3J backtesting windows										
DQ test (5% VaR)	*0.667*	1.000	0.968	0.968	0.957	0.301	1.000	0.914	0.548	*0.097*
DQ test (1% VaR)	0.570	1.000	1.000	1.000	*0.269*	0.333	1.000	1.000	0.495	*0.290*
Probit test (5% VaR)	*0.452*	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.968	*0.290*
Probit test (1% VaR)	*0.204*	1.000	1.000	1.000	0.290	0.677	1.000	1.000	1.000	*0.312*
Panel A3: Equally-weighted 2/3J backtesting windows										
DQ test (5% VaR)	*0.000*	0.308	0.205	0.205	0.238	0.200	0.427	0.200	0.195	*0.000*
DQ test (1% VaR)	0.151	1.000	0.622	0.611	*0.011*	*0.205*	0.865	0.865	0.205	*0.205*
Probit test (5% VaR)	*0.000*	0.368	0.227	0.227	0.611	0.346	0.768	0.308	0.092	*0.000*
Probit test (1% VaR)	*0.000*	0.649	0.384	0.362	*0.000*	0.914	1.000	1.000	0.816	*0.135*
Panel B: Weighted by Absolute Coverage Error										
DQ test (5% VaR)	*0.191*	0.807	0.798	0.803	0.514	0.381	0.794	0.725	0.529	*0.047*
DQ test (1% VaR)	0.242	1.000	0.932	0.927	*0.054*	*0.171*	0.883	0.883	0.432	*0.281*
Probit test (5% VaR)	*0.085*	0.832	0.801	0.807	0.750	0.726	0.972	0.787	0.628	*0.113*
Probit test (1% VaR)	*0.021*	0.954	0.863	0.855	0.048	0.813	1.000	1.000	0.941	*0.236*
Panel C1: Weighted Asymmetrically Underprediction > Overprediction										
DQ test (5% VaR)	*0.050*	0.804	0.905	0.902	0.514	0.329	0.785	0.777	0.457	*0.280*
DQ test (1% VaR)	0.243	1.000	0.955	0.948	*0.054*	*0.143*	0.882	0.881	0.375	0.924
Probit test (5% VaR)	*0.045*	0.836	0.921	0.919	0.750	0.939	0.986	0.865	0.705	*0.672*
Probit test (1% VaR)	*0.021*	0.957	0.884	0.873	0.048	*0.807*	1.000	1.000	0.932	0.919
Panel C2: Weighted Asymmetrically Overprediction > Underprediction										
DQ test (5% VaR)	*0.318*	0.839	0.615	0.630	0.567	0.447	0.841	0.627	0.757	*0.001*
DQ test (1% VaR)	0.184	1.000	0.568	0.580	*0.019*	0.749	0.000	0.934	0.765	*0.174*
Probit test (5% VaR)	*0.120*	0.834	0.593	0.607	0.775	0.482	0.916	0.633	0.411	*0.002*
Probit test (1% VaR)	*0.003*	0.897	0.551	0.568	0.016	0.949	1.000	1.000	0.997	*0.122*