Graduate School of Decision Sciences

Universität Konstanz

GSDS
Working Paper
No. 2018-04

How Informative is High-Frequency Data for Tail Risk Estimation and Forecasting? An Intrinsic Time Perspectice

Timo Dimitriadis
Roxana Halbleib

April 2018

Graduate School of Decision Sciences

All processes within our society are based on decisions – whether they are individual or collective decisions. Understanding how these decisions are made will provide the tools with which we can address the root causes of social science issues.

The GSDS offers an open and communicative academic environment for doctoral researchers who deal with issues of decision making and their application to important social science problems. It combines the perspectives of the various social science disciplines for a comprehensive understanding of human decision behavior and its economic and political consequences.

The GSDS primarily focuses on economics, political science and psychology, but also encompasses the complementary disciplines computer science, sociology and statistics. The GSDS is structured around four interdisciplinary research areas: (A) Behavioural Decision Making, (B) Intertemporal Choice and Markets, (C) Political Decisions and Institutions and (D) Information Processing and Statistical Analysis.

GSDS – Graduate School of Decision Sciences
University of Konstanz
Box 146
78457 Konstanz

Phone: +49 (0)7531 88 3633
Fax: +49 (0)7531 88 5193
E-mail: gsds.office@uni-konstanz.de

-gsds.uni-konstanz.de

ISSN: 2365-4120
April 2018

How informative is high-frequency data for tail risk estimation and forecasting? An intrinsic time perspective

Timo Dimitriadis[1] and Roxana Halbleib[2]

April 27, 2018

Abstract

This paper proposes a novel and simple approach to compute daily Value at Risk (VaR) and Expected Shortfall (ES) directly from high-frequency data. It assumes that financial logarithm prices are subordinated unifractal processes in the intrinsic time, which stochastically transforms the clock time in accordance with the markets activity. This is a very general assumption that allows for a simple computation of daily VaR and ES by scaling up their intraday counterparts computed from data sampled in intrinsic time. In the empirical exercise, we discuss the statistical and dynamic properties of the resulting daily VaR and ES estimates and show that our method outperforms standard ones in accurately estimating and forecasting VaR and ES.

Keywords: Value at Risk, Expected Shortfall, Intrinsic Time, Subordinated Process, High-Frequency Data, Scaling Law

[1]Department of Economics, University of Konstanz, Germany; email: timo.dimitriadis@uni-konstanz.de.

[2]Department of Economics, University of Konstanz, Germany; email: roxana.halbleib@uni-konstanz.de.

Corresponding address: Roxana Halbleib, University of Konstanz, P.O. Box 124, Universitaetstrasse 10, 78464, Konstanz, Germany; phone: 0049 (0) 7531 88 5373; fax: 0049 (0) 7531 88 4450.

1 Introduction

The huge losses during and following the previous financial crisis have shown that the existing risk models have failed when they were needed most. Since its implementation as a standard approach by the Basel Committee on Banking Supervision in 1996 (Basel Committee, 1996), Value at Risk (VaR) has become the most known risk measure in today's financial world. In 2016, the Basel Committee recommended a shift from VaR to Expected Shortfall (ES) as a "more prudent" measure of risk given that it accounts for the distribution of financial losses beyond VaR (Basel Committee, 2016). Although very appealing, the literature on estimating and forecasting ES relies heavily on the one of VaR. The standard way of computing e.g., (daily) VaR is by using a location-scale model that involves the estimation and forecasting of the mean, of the standard deviation (or volatility) and of the quantiles of standardized residuals of (daily) financial returns from data sampled in calendar time and, usually, at low (daily) frequencies.[1] The location-scale approach stems originally from the assumption that daily financial returns are normally distributed, which has proven to be an unrealistic assumption during the last decades. To overcome this problem, the common practice is to replace the quantile of standardized residuals with one of a fat-tailed distribution, such as Student's t. However, Halbleib and Pohlmeier (2012), among others, provide empirical evidence on the general pitfalls of the location-scale model for VaR, especially during turbulent financial times, and the necessity for more flexible and precise approaches.

This paper proposes a novel and simple method of estimating and forecasting daily VaR and ES directly from high-frequency (HF) data. It consists in scaling-up quantiles and ES of intraday returns sampled not in calendar (clock or physical) time, but in a time dimension that aims at capturing the real "heartbeat" of the market. This new time dimension transforms the time in accordance with market's activity and allows to sample data more often during active/turbulent financial periods and less often during calm ones. Thus, the resulting intraday returns become very informative about the market's activity and riskiness, differently from the ones sampled in the standard clock time, which ignores such information. Our methods, which we denote the Scaled-up VaR (ScaVaR) and the Scaled-up ES (ScaES), are the first in the literature that directly exploit the rich information content of HF data from the intrinsic time perspective for tail risk estimation and forecasting.

The theoretical background of our approach is built on the assumption that the financial logarithm price process is a subordinated process. In particular, we assume that it is a unifractal process subordinated (indexed) by the stochastic intrinsic time; i.e. $p(t) = U(\theta(t))$, where $U(t)$ is the subordinated unifractal process and $\theta(t)$ is the stochastic intrinsic time process (the subordinator). The unifractality assumption is particularly attractive, as it allows for describing the relation between the distributions of financial returns computed at different aggregation (frequency) levels through one single scaling law driven by one scaling index. This enables us to compute daily estimates of VaR and ES by scaling up their intraday counterparts by the single scaling index estimated from the

[1] Intraday data sampled in calendar time has found application in estimating and forecasting daily VaR by using realized volatility estimates for the standard deviation in the location-scale model (see Giot and Laurent (2004) among others).

data, which makes our procedure very appealing compared to the location-scale models, as it avoids heavy parametric specifications for the mean, volatility and the distribution of the standardized returns.

The idea of subordination goes back to Mandelbrot and Taylor (1967), which generalize the classical log-normal asset price model by substituting the physical time with the (non-negative stochastic) intrinsic one. This approach is able to capture the fat-tailedness of financial returns in a more tractable way than by assuming a heavy-tailed distribution, such as the α-stable. Most of the existing subordinting processes use the Brownian motion as the subordinated one, but differ in what regards the (parametric) assumption on the stochastic time process (subordinator): the skewed $\alpha/2$-stable as in Mandelbrot and Taylor (1967), the inverse Gaussian as in Eberlein and Keller (1995), the log-normal as in Clark (1973) or the multifractal as in Mandelbrot et al. (1997).

The main two limitations of these subordinated models are that (1) the financial returns are imposed to be Gaussian and (2) they use daily data for the estimation. Our model overcomes these limitations as the unifractality assumption imposes no Gaussian constraints for the price increments and it uses HF data for the estimation i.e., the intrinsic time is derived directly from the intraday HF information based on a choice of an intensity measure of the market's activity (e.g., intraday trading pattern). Moreover, by assuming that the underlying intensity is a multifractal measure (Mandelbrot et al., 1997; Calvet and Fisher, 2002), we show theoretically that our assumption on the logarithm price process is very general as the the logarithm price process becomes a multifractal one characterized by different scaling laws describing the relationships among the distributions of aggregated returns, in accordance with the aggregation (frequency) level. As shown by Mandelbrot (1999); Calvet and Fisher (2008); Lux and Segnon (2018), the multifractality assumption is particularly attractive for financial price processes as it accommodates various empirical features such as the fat-tailedness of the distribution of returns as well as volatility clustering and persistence in the return dynamics. As a consequence, our basic assumption that the financial logarithm price process is a unifractal process subordinated (indexed) by the intrinsic time, is theoretically a very general one and, in the same time, practically, very usable due to the simple relationships typical to unifractality.

Despite its practical relevance, the concept of intrinsic time has got so far only limited attention from the economics and finance literature and, so far, only for other purposes than the ones of this paper (Stock, 1988; Müller et al., 1995; Ghysels et al., 1997; Oomen, 2006; Guillaume et al., 1997). It has circulated under different names, such as "time deformation" (Mandelbrot et al., 1997), "deformed time" (Gouriéroux and Jasiak, 2001) or "operational time" (Clark, 1973) and, as mentioned above, it has been mainly treated as a latent process with increments following certain parametric distributions estimated from low frequency data (Hurst et al., 1997). However, the availability of HF data in the last one and a half decades opens new possibilities of measuring and estimating the intrinsic time (Marinelli et al., 2001, 2000) based on market intensity measures, such the trading volume (Mandelbrot and Taylor, 1967; Clark, 1973), the intraday volatility pattern (Boudt et al., 2011; Dong and Tse, 2014), the intraday trading pattern (Oomen, 2005; Griffin and Oomen, 2008; Wu, 2012) or in the context of duration models (Engle and Russell, 1998; Gerhard and Hautsch, 2002).

The application of the fractal theory in finance has mainly focused so far on providing models for asset returns and their volatilities as in Mandelbrot et al. (1997), Calvet et al. (1997) and Calvet and Fisher (2001) (see Calvet and Fisher (2008) and Lux and Segnon (2018) for an overview). Hallam and Olmo (2014*a*) and Hallam and Olmo (2014*b*) apply scaling relationships typical to fractality to estimate and forecast the whole probability density function of daily financial returns from their intraday counterparts computed from data sampled in calendar time. The outcome of these two approaches is mainly driven by the information from the central region of the distribution instead of the tails, which makes them inappropriate for risk estimation and forecasting.

Differently from these fractal-based models, our ScaVar and ScaES focus on the estimation and forecast of the tails of return distributions by accounting for the intraday information on market's activity within the minimal parametric framework imposed by the fractal theory. The practical implementation of ScaVaR and ScaES consists in several simple steps, as it follows: (1) sample HF prices in the intrinsic time based on a choice of intensity measure and at a certain frequency[2]; (2) estimate quantiles and ES of the resulting intraday logarithm returns; (3) estimate the scaling index that describes the unique scaling law of the unifractal price process and (4) compute daily estimates of VaR and ES by scaling up the intraday counterparts from point (2) with the scaling index obtained at point (3).

In our empirical application, we apply ScaVaR and ScaES to one exchange rate and two stocks. For our purposes, we use the intraday trading pattern as the intensity measure driving the intrinsic time sampling scheme: i.e., the sampled prices are "equidistant" in terms of the number of trades. Further alternative intensity measures such as price changes, volumes or intraday volatility patterns may also be considered. For the sake of comparison, we also implement the standard clock-time sampling scheme that samples prices equidistantly in time.

We provide a comprehensive description of the empirical properties of scaled-up daily VaR and ES estimates and find that the estimates of foreign exchanges are less extreme, less volatile and less persistent than the ones of the stocks. Among the stocks, these effects are mostly pronounced for the asset from the financial sector (BAC). All VaR and ES estimates exhibit long memory and some skewness and overkurtosis.

Furthermore, we evaluate and compare the accuracy of ScaVaR and ScaES to estimate and forecast daily VaR and ES against a standard location-scale model and the fractal-based approach of Hallam and Olmo (2014*a*). We show that ScaVaR and ScaES significantly outperform the alternative methods, especially the method of Hallam and Olmo (2014*a*), in accurately estimating and forecasting daily VaR and ES. Comparing the two sampling schemes, we show that sampling the HF data in accordance with the trading activity is much more valuable for risk estimation and forecasting than the calendar one.

This paper is organized as follows: Section 2 presents the theoretical background of

[2]Please note that in this paper we do not deal with the market microstructure noise, which affects HF prices. As it is a very complex topic by itself (as depicted in the realized volatility literature: Hansen and Lunde (2004), Bandi and Russell (2005), Zhang et al. (2005), Barndorff-Nielsen et al. (2008), Dahlhaus and Tunyavetchakit (2018) among others), we leave it for further research.

our approach and its implementation. Section 3 provides empirical results from applying our method and the alternatives to real data. Section 4 concludes.

2 Theory

Section 2.1 briefly introduces the key mathematical concepts for our methodology, Section 2.2 provides its formal description and properties and Section 2.3 presents the practical steps in implementing it.

2.1 Key Theoretical Concepts

The most basic concept at the basis of our approach is the one of *subordination* (Mandelbrot and Taylor, 1967) and consists in replacing the calendar time with an *intrinsic* one in defining a stochastic process that evolves in time. Let $X = \{X(t), t \geq 0\}$ be a stochastic process and $\theta = \{\theta(t), t \geq 0\}$ be a non-negative stochastic process. Then the new process $\{X(\theta(t)), t \geq 0\}$ is said to be subordinated to X by the intrinsic time $\theta(t)$. For a detailed discussion of subordinated processes and their application in modeling financial logarithm price processes see Chapter 5 of Rachev and Mittnik (2000), among others. In what follows, we introduce in more detail the concept of intrinsic time as applied in our methodology.

Assume $t \in [0, T]$ where $T > 0$. Let λ denote a real-valued intensity measure on $[0, T]$, which measures the market's activity at a certain point of time. From this intensity measure, we generate the time transformation $\theta : [0, T] \to [0, T]$, which transforms the calendar time t into the *intrinsic time* $\theta(t)$:

$$\theta(t) = \lambda\left([0, t]\right). \tag{1}$$

$\theta(t)$ is an increasing function of the clock time t in accordance with the intensity measure λ, i.e., given the market's intensity λ between two time points $s, t \in [0, T], s < t$, which are close to each other, the transformed time $\theta(t)$ elapses such that $\theta(t) - \theta(s) = \lambda\left([0, t]\right) - \lambda\left([0, s]\right) = \lambda\left([s, t]\right)$. Thus, the intrinsic time aims at capturing the real "heartbeat" of the market in accordance with its market intensity: it accelerates the time when the market's activity is intense and it slows it down when it is calm. Choices of intensity markers could be: the number of trades, the number of price changes, trading volumes and intraday volatility patterns (see Clark (1973), Gouriéroux and Jasiak (2001), Oomen (2005), Dong and Tse (2014), Boudt et al. (2011), Mandelbrot et al. (1997), among others).

The next concept used in our methodology is the one of *unifractality* (Mandelbrot, 1963, 1982), which is defined in the following manner:

Definition 2.1. *A real-valued stochastic process* $\{U(t), t \geq 0\}$ *is said to be* unifractal *or* self-affine *if for some* $H > 0$ *and all* $c \geq 0$*,*

$$U(ct) \stackrel{d}{=} c^H U(t), \tag{2}$$

where $\stackrel{d}{=}$ *denotes equality in distribution.*

Hereby, the parameter H is known as the *self-affinity* or *scaling index* and it gives the (unique) scaling law describing the unifracal process. c is the scaling or aggregation factor that defines the time scale change from t to ct. It holds that, if $E[|U(1)|] < \infty$, then $H \leq 1$ and if H =1, then $U(t) = tU(1)$ almost surely (Embrechts and Maejima, 2002). For this reason, we focus in our paper on the cases where $H \in (0, 1)$.

Assuming that the increments are Gaussian and if $H = 0.5$, then $U(t)$ is a standard Brownian Motion (BM), while if $H \neq 0.5$, then $U(t)$ is a fractional Brownian Motion (fBM). While both BM and fBM have stationary (Gaussian) increments, they are independent in the case of BM and dependent in the case of fBM.[3] fBM with $H > 0.5$ exhibits long memory. As a consequence, the self-affinity index H is in the long memory literature also known as the Hurst coefficient (see Beran et al. (2013), among others).

Proposition 2.2. *Let* $\{U(t), t \geq 0\}$ *be a unifractal process as given in Equation (2) with* $H \in (0, 1)$ *and with stationary increments. Then, for all* $c \geq 0$ *and for all* $\Delta > 0$*, it holds that*

$$U(t + c\Delta) - U(t) \stackrel{d}{=} c^H[U(t + \Delta) - U(t)]. \tag{3}$$

The proof of the Proposition 2.2 is given in Appendix A. The proposition states that, if a process is unifractal and has stationary increments, then the increments are also unifractal following the same scaling law as the process itself.

The scaling relationship describing a unifractal process (given in Equation (2)) remains unchanged w.r.t. c: i.e., H does not depend on the choice of c. This might be very restrictive in many applications, including in the financial ones, as the relationships among variables scaled (aggregated) at different time scales (frequencies) may not be simple (unifractal), but complex, as they might change with the scale or aggregation factor. Therefore, Mandelbrot (1974) introduces a more general concept, namely the *multifractality* that, differently from the unifractality, allows for the scaling relationships to change with the scaling factor c. The multifractality can be defined in two equivalent ways as it follows (Mandelbrot et al., 1997; Calvet and Fisher, 2002):

Definition 2.3. *A real-valued process* $\{M(t), t \geq 0\}$ *is said to be* multifractal

1. *If there exists a real-valued random variable* $H(c)$ *such that for all* $c \geq 0$*,*

$$M(ct) \stackrel{d}{=} c^{H(c)}M(t), \tag{4}$$

where $M(t)$ *and* $H(c)$ *are independent random functions and* $H(c)$ *is time independent* or

[3] A self-affine process $\{U(t), t \geq 0\}$ with stationary and independent increments is an α-stable Lévi process if the distribution of $U(1)$ is the α-stable described by the tail index $0 < \alpha \leq 2$. Thus, $H = 1/\alpha$ and for $\alpha = 2$, the α-stable distribution becomes the Gaussian distribution (Embrechts and Maejima, 2002).

2. *If $M(t)$ has stationary increments and satisfies*

$$E[|M(t)|^q] = c(q)t^{\tau(q)+1}, \quad \forall q \in \mathcal{Q} \quad \textit{and} \quad \forall t \in \mathcal{T} \tag{5}$$

where $c(q)$ and $\tau(q)$ are deterministic functions of q, $\tau(q)$ is denoted the scaling function, $\mathcal{Q}$ and $\mathcal{T}$ are intervals on the real line such that $[0,1] \subseteq \mathcal{Q}$, $\mathcal{T} \subseteq [0,\infty)$ and $\{0\} \in \mathcal{T}$.

While the first definition refers to the variation of the scaling relationships with respect to the choice of the scaling factor c, the second one states that a multifractal process has stationary increments and its moments exhibit multi-scaling laws, denoted here as *moment-scaling relationships*. $H(c)$ is known as the generalized Hurst component and it is a function of c: it allows for different scaling relationships among the distribution of $M(t)$ sampled at different timescales (frequencies). From Definition 2.3, one can see that the unifractality with stationary increments is a special case of the multifractality: if $M(t)$ is unifractal then, $H(c) = H, \forall c \geq 0$ and $M(t) \stackrel{d}{=} t^H M(1)$. Similar to Proposition 2.2, one can easily show that the increments of a multifractal process are also multifractal and have the same moment-scaling relationships as the process itself (Mandelbrot et al., 1997).

While the concept of multifractality may seem at a first glance simple to understand, it is difficult to implement in practice, especially in the form defined in Equation (4), due to the difficulty of computing $H(c)$ for all choices of c. Most of the applications of multifractality in the finance literature (e.g., Hallam and Olmo (2014*a*), Calvet and Fisher (2004)) make use of the second definition that involves concepts (stationarity and moment-scaling relationships), which are easier to deal with in practice.

Multifractal theory has found so far a wide application in many disciplines as it is able to describe complex systems characterized by long dynamic dependencies, occurrence of extreme events and multi-scaling relationships (Kantelhardt, 2009). The multifractal theory is, thus, able to address all these features within a unified mathematical framework (Lux and Segnon, 2018; Kobeissi, 2013). The fact that financial returns exhibit such empirical properties (volatility clustering and fat-tails), makes the multifractal framework an attractive one in describing the price process of the financial assets as shown by Mandelbrot et al. (1997) and Calvet and Fisher (2004).

2.2 ScaVaR and ScaES

The main assumption at the fundament of our methodology is that the financial logarithm price process is a subordinated unifractal process indexed by in the intrinsic time $\theta(t)$. More formally:

Assumption 2.4. *Let $P(t)$ be the financial price process and $p(t) = \ln P(t)$.*[4] *We assume that:*

$$p(t) = U(\theta(t)), \tag{6}$$

[4]Notice that we normalize the logarithm price process by subtracting $p(0)$ from $p(t)$ as Definition 2.1 implies that for a unifractal process $U(t)$, it must hold that $U(0) = 0$ almost surely.

where $U(t)$ is a unifractal process with the Hurst coefficient H and stationary increments and $\theta(t)$ is independent of $U(t)$.

Assumption 2.4 is a very general one for financial logarithm price processes and includes as special cases: (1) the BM assumption when $H = 0.5$, $\theta(t) = t$ and the increments are independent and Gaussian, which is the mostly widely-spread assumption in the finance theory (Black and Scholes, 1973; Markowitz, 1952); (2) the fBM assumption when $H \neq 0.5$, $\theta(t) = t$ and the increments are Gaussian, but dependent, which is promoted by Comte and Renault (1998), Comte and Renault (1996) and Mandelbrot (1997), among others and (3) the assumption that the logarithm price process is a subordinated BM or fBM in $\theta(t)$ as in the multifractal model of asset returns (MMAR) of Mandelbrot et al. (1997) and Calvet and Fisher (2002). A major concern regarding all these previous assumptions is about the Gaussianity of the financial returns (sampled in intrinsic time), which is empirically unrealistic for the current financial returns as they exhibit fat-tailed distributions. In contrast, Assumption 2.4 only assumes that $p(t)$ follows a unifractal law not in the clock time t, but in its transformation $\theta(t)$: i.e., $U(c\theta(t)) \stackrel{d}{=} c^H U(\theta(t))$. This allows for a simple derivation of the distribution and, for our purposes, of the quantiles of aggregated (daily) returns from their intraday counterparts sampled in intrinsic time by simply scaling them up by c^H.

In what follows, we provide further theoretical evidence on the very general character of Assumption 2.4 compared to the existing ones (e.g., BM, fBM) by showing that, assuming that $\lambda(t)$ is a multifractal measure, the logarithm price process in Equation (6) becomes a multifractal one that, as mentioned above, accommodates many important empirical stylized facts of financial returns, such as volatility clustering and fat-tailedness (Lux and Segnon, 2018).

In order to show this, we first define the concept of multifractal measure as in Calvet and Fisher (2002).

Definition 2.5. *Let μ be a random measure on $[0, T]$. μ is a multifractal measure if*

$$\mathbb{E}\left[\mu\left([t, t+\Delta]^q\right)\right] = c_\mu(q)(\Delta)^{\tau_\mu(q)+1}, \tag{7}$$

for all $(t, \Delta) \in \mathfrak{D}$, where $\mathfrak{D}$ is a subset of $[0, T]\times[0, T]$, such that the closure of $\mathfrak{D}$ contains the set $[0, T] \times \{0\}$, i.e., the scaling relation in Equation (7) holds "in the neighborhood of any instant" and all $q \in \mathcal{Q}$, where $c_\mu(q)$ and $\tau_\mu(q)$ are deterministic functions of q and $\mathcal{Q}$ is specified in Definition 2.3.

Thus, a multifractal measure is characterized by moment-scaling relationships similar to the ones given in Equation (5). In what follows, we assume that the intensity measure λ is a multifractal measure:

Assumption 2.6. *The intensity measure λ defined on $[0, T]$ is a multifractal measure i.e., $\mathbb{E}\left[\lambda\left([t, t+\Delta]^q\right)\right] = c_\lambda(q)(\Delta)^{\tau_\lambda(q)+1}$, where $c_\lambda(q)$ and $\tau_\lambda(q)$ are deterministic functions of q.*

Based on Assumption 2.6, we now show that $\theta(t)$ is a multifractal process as defined in Section 2.1:

Theorem 2.7. *Let $\lambda(t)$ be a multifractal measure on $[0,T]$. If there exists $\varepsilon \in (0,1)$ such that $(-\varepsilon, 1] \subseteq \mathcal{Q}$, then, $\theta(t) = \lambda([0,t])$ is a multifractal process with continuous and non-decreasing paths.*

The proof of the theorem is in Appendix A. Based on Theorem 2.7, one can now show that the price process specified in Assumption 2.4 is a multifractal one:

Theorem 2.8. *Assume that the price process $p(t)$ given in Equation (6) satisfies the condition: $\mathbb{E}[|p(t)|^q] < \infty \; \forall q \in \mathcal{Q}$, where is $\mathcal{Q}$ is specified in Definition 2.3. Then, $p(t)$ is multifractal with the scaling function $\tau_p(q) = \tau_\lambda(Hq)$ and the scaling factor $c_p(q) = c_\lambda(Hq)\mathbb{E}[|U(1)|^q]$.*

See Appendix A for the proof. From Proposition 2.2, we further get that the increments of the logarithm price process assumed in Equation (6) are also unifractal in the intrinsic time $\theta(t)$, and consequently, based on Theorem 2.8 also multifractal. While Assumption 2.4 may seem at a first glance ad hoc, Theorem 2.8 proves its very general character, by including the multifractal processes as special cases, along with further assumptions that are less realistic for financial returns such as the unifractality, BM or fractional BM. While the theoretical results so far provide evidence on the generality of Assumption 2.4, below and in Section 3 we demonstrate its usefulness in practical applications, due to the simplicity in describing the relationships among the distributions of returns aggregated at various levels, typical to unifractal processes.

Thus, by replacing t with $\theta(t)$ in Proposition 2.2, one can show that the distribution of the returns of the logarithm price process assumed in Equation (6) follows the same unifractal scaling law as the logarithm price itself:

Corollary 2.9. *For all $c \geq 0$ and for all $\Delta > 0$, it holds that*

$$U(\theta(t) + c\Delta) - U(\theta(t)) \stackrel{d}{=} c^H [U(\theta(t) + \Delta) - U(\theta(t))]. \tag{8}$$

From Corollary 2.9, one gets that the quantiles (VaR) and ES at probability level p of the increments of the logarithm price sampled in intrinsic time have the following scaling relationships:

$$Q_p\big(U(\theta(t) + c\Delta) - U(\theta(t))\big) = c^H Q_p\big(U(\theta(t) + \Delta) - U(\theta(t))\big) \tag{9}$$

$$ES_p\big(U(\theta(t) + c\Delta) - U(\theta(t))\big) = c^H ES_p\big(U(\theta(t) + \Delta) - U(\theta(t))\big), \tag{10}$$

for all $p \in (0,1)$, all $c \geq 0$ and for all $\Delta > 0$. Thus, equations (9) and (10) allow us to compute quantiles (VaR) and ES of increments sampled at the daily frequency by scaling up their counterparts computed from increments sampled at a higher frequency (e.g., 5 minute) and in intrinsic time. We denote these approaches by **Scaling-up VaR (ScaVaR)** and **Scaling-up ES (ScaES)**.

2.3 Implementation of ScaVaR and ScaES

Based on the theoretical assumptions and results presented in Section 2.2, we show here how to obtain daily estimates of quantiles (VaR) and ES from HF data by implementing the following steps:

Step 1. Based on choices of the intensity measure λ, generate the time transformation function $\theta(t) = \lambda([0,t])$ for all $t \in [0,T]$.[5]

In order to sample $c+1$ HF logarithm prices (logarithm ticks) according to the transformed time $\theta(t)$, fix $c \in \mathbb{N}$.[6]

Define the equally spaced calendar time grid $\boldsymbol{t} = \{t_0, t_1, t_2, \ldots, t_c\}$, where $t_0 = 0$, $t_c = T$ and $t_i - t_{i-1}$ is constant for all $i = 1, \ldots, c$.

The intrinsic time grid that samples $c+1$ prices according to the intrinsic time $\theta(t)$ is computed as:

$$\begin{aligned} \tilde{\boldsymbol{t}} &:= \{\theta^{-1}(t), t \in \boldsymbol{t}\} \\ &= \{\theta^{-1}(t_0), \theta_d^{-1}(t_1), \ldots, \theta^{-1}(t_c)\} \\ &:= \{\tilde{t}_0, \tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_c\}, \end{aligned} \tag{11}$$

where we fix the first and the last observation, i.e., $\theta^{-1}(t_0) = \tilde{t}_0 = 0$ and $\theta^{-1}(t_c) = \tilde{t}_c = T$.[7]

From the ticks sampled on the time grid $\tilde{\boldsymbol{t}}$, compute the c intraday log-returns for each day d, denoted here by $\boldsymbol{r}_{f,d} = \{r_{1,d}, \ldots, r_{c,d}\}$ with $d = 1, \ldots \mathcal{D}$.

Step 2. From the intraday log-return series $\boldsymbol{r}_{f,d}$ of day d and for a given probability $p \in (0,1)$, compute the p-th quantiles and corresponding ES, denoted here by $\widehat{Q}_p(\boldsymbol{r}_{f,d})$ and $\widehat{ES}_p(\boldsymbol{r}_{f,d})$, respectively.

Step 3. Estimate the Hurst coefficient H from the intraday log-return series (see Kantelhardt (2009) for a survey of estimation methods of H).

Step 4. Based on equations (9) and (10), compute the p-th quantiles and p-th ES of the daily returns on day d with $p \in (0,1)$, denoted by $\widehat{Q}_{p,d}$ and $\widehat{ES}_{p,d}$, by scaling up their counterparts computed in Step 2 with the scale index H estimated in Step 3, as it follows:

$$\widehat{Q}_{p,d} = c^{\hat{H}} \widehat{Q}_p(\boldsymbol{r}_{f,d}), \tag{12}$$

$$\widehat{ES}_{p,d} = c^{\hat{H}} \widehat{ES}_p(\boldsymbol{r}_{f,d}). \tag{13}$$

[5] We provide concrete examples on how to choose λ in the empirical exercise in Section 3.

[6] Alternatively, one can fix the frequency f of sampling the HF data, which is related to c in the following manner: $f = T/c$, where T is the total amount (measured in units of time: e.g., minutes) of calendar time per day (e.g., on NYSE $T = 390$ minutes). For example, on NYSE, if $c = 78$, then $f = 5$ minute. Note that the term "frequency" here does not necessarily have the classical calendar interpretation. E.g., a frequency of 5 minute does not necessarily mean that we sample every 5 minute, but the sampling points varies with the choice of the intensity measure such that, for instance on NYSE, one samples a total 78 observations per day sampled at 5 minute frequency.

[7] A special case of this procedure is the Calendar Time Sampling (CTS) scheme that samples price observations that are equidistant in calendar time. Thus, the intensity measure λ is fixed and equal to 1 and $\tilde{\boldsymbol{t}} = \boldsymbol{t}$.

3 Empirical Application

3.1 Data Description

We apply ScaVaR and ScaES to stock and foreign exchange data. In particular, we use the tick data of International Business Machines Corporation (IBM) and Bank of America (BAC) collected from TAQ database of NYSE from January 2, 2001 to July 24, 2017. We have a total of $\mathcal{D} = 4124$ trading days for each stock. The raw data consists of the best bid and best ask quotes during the trading period, which is from Monday to Friday from 9:30:00am to 16:00:00pm. From the best bid and best ask, we compute the midquote as their average. The choice of the stocks aims at providing some diversification in what regards the liquidity of trading as well as the sector and the profile of the companies considered: IBM is a common choice in the literature due to its high-liquidity, while BAC is from the financial sector and has experienced some serious losses, especially during the previous financial crisis.

The foreign exchange rate data we use is for the Euro against the US Dollar (EURUSD) from July 5, 2008 until July 20, 2016 ($\mathcal{D} = 2125$). The raw tick data is obtained from TICKDATA[8] that collects foreign exchange rate transactions from multiple data sources covering a variety of market places around the world. The foreign exchange rates are traded around the clock from Sunday 6pm ET until Friday 5:59:59.999pm ET[9], adding up to a total of five days a week with 24 hours of trading activity.

According to Step 1 presented in Section 2.3, the tick data is sampled at the frequency $f = 5$ minute leading to $c = 288$ for the foreign exchange rates and $c = 78$ for the stocks. The choice of this frequency attempts to accommodate the fact that intraday returns are contaminated with market microstructure noise (see Andersen, Bollerslev, Diebold and Ebens (2001) for a similar procedure when estimating daily volatilities by means of realized volatility measures).[10] However, a theoretical and empirically deeper analysis of the impact of the market microstructure noise on our estimates is left for further research.

For our purposes, we implement two sampling schemes, namely (1) the Calendar Time Sampling (CTS) scheme that samples price observations equidistantly in clock time and (2) the Time Transformation Sampling (TTS) scheme that is an intrinsic time scheme that samples price observations that are equidistant in terms of the number of ticks averaged over all past trading days. In this case, the intensity measure $\lambda(t)$ is given by the number of ticks at time t averaged over all days prior to and including the present day. This sampling scheme accounts for the intraday periodicity in the trading activity as a common pattern across all trading days (Wu, 2012).[11]

[8] https://www.tickdata.com/

[9] Before August 2, 2014, the trading data provided by TICKDATA covers trades from Sunday 4pm ET until Friday 3:59:59.999pm ET every week.

[10] Results from implementing other frequencies, such as 1 minute, 3 minute, 10 minute and 30 minute can be obtained from the authors upon request.

[11] Implementing further intrinsic sampling schemes based on other types of intensity measures, such as non-zero changes in the prices, volumes or volatility patterns (Clark, 1973; Engle and Russell, 1998; Griffin and Oomen, 2008; Oomen, 2005, 2006; Boudt et al., 2011; Dong and Tse, 2014) provides, in general, similar results to TTS when compared to CTS and the alternatives described below. They can be obtained

Descriptive statistics of HF returns sampled by the two sampling schemes are given in Table B.1 in Appendix B. As one may observe from the table, the magnitude of the returns is very small, but their kurtosis is very large and much above the value of 3. However, the value of the kurtosis reduces from the calendar time to the intrinsic one.[12] This indicates that the intrinsic time sampling scheme leads to a reduction in the sampled extreme values compared to CTS as it picks more price observations during highly active periods (with high fluctuations) and less during calm one (with low fluctuations). In contrast, CTS samples the data independently of the market's activity.

In what follows, we compute VaR and ES at $p = 1\%, 2.5\%$ and 5%: $p = 1\%$ is the probability specified by the Basel Committee to compute VaR (Basel Committee, 1996), $p = 2.5\%$ is the probability to compute ES according to Basel Committee (2016) and $p = 5\%$ is the probability that is most popular among researches and practitioners (e.g., Kuester et al. (2006), among others).

In order to implement Step 2 presented in Section 2.3, we estimate the quantiles of the intraday returns by means of empirical quantiles by employing the MATLAB function $'quantile(x)'$. The ES estimator is computed as the average of the observations smaller than the empirical quantile estimator. The empirical quantile and ES estimators are little precise in small samples (see Harrell and Davis (1982), among others), which is also our case, as the number of intraday observations is limited. One way of improving the quality of these estimators is to increase the number of intraday returns, namely c. This leads to increasing the frequency of sampling the HF information, which may induce a bias in the estimates due to the market microstructure noise present in the intraday prices (due to bid-ask spreads, discrete price changes and asymmetric market information, among others). This issue is treated at length in the realized volatility literature that finds that sampling returns at frequencies between 5 and 30 minute provide an optimal trade-off between accuracy and efficiency of the realized volatility estimators (for an overview of the literature see Hansen and Lunde (2006)).

To implement Step 3 in Section 2.3, we estimate the Hurst coefficient by means of the detrended moving average (DMA) estimator of Alessio et al. (2002). We also implement further estimators, such as the detrended fluctuation analysis of Peng et al. (1994) and the estimators of Sánchez-Granero et al. (2012). However, given that the in-sample and out-of-sample results do not change significantly by changing the estimation method of H, we stick here to the results computed from implementing DMA.

Based on the estimate of H and of the intraday empirical quantiles and ES, we are now able to implement Step 4 presented in Section 2.3 in order to compute the daily scaled-up estimates of VaR and ES as given in equations (12) and (13). Tables B.2 and B.3 provide descriptive statistics of the logarithm transformations of their negative values namely of $\hat{q}_{p,d} := \ln(-\widehat{Q}_{p,d})$ and $\widehat{es}_{p,d} := \ln(-\widehat{ES}_{p,d})$. From both tables, we observe that the $\widehat{q}_{p,d}$ and $\widehat{es}_{p,d}$ exhibit overkurtosis and skewness and one rejects the normal distribution assumption in all cases. In general, there is a tendency in reducing the kurtosis and JB-test values from CTS to TTS, which may indicate a reduction in the extreme values in the estimated VaR and ES series. However, the most evident result is that exchange rates

from the authors upon request.

[12] Similar empirical findings are documented by Clark (1973) and Mandelbrot and Taylor (1967).

exhibit smaller kurtosis and JB-test values than the stocks, with BAC exhibiting the largest ones. This can be explained by the fact that the stock markets are more risky and that, financial crises, included in the sample, affect more the stock market than the currency one, especially the stocks from the financial sector.

A graphical inspection of the VaR series in Figure C.1 confirms these findings.[13] The estimates of the stocks are much more volatile, with higher volatility clustering and higher peaks than the ones of the exchange rates, with the mostly pronounced ones for BAC. However, the clustering effects reduce by increasing p. The ACF of the series, depicted in Figure C.2 , indicate that the transformed VaR estimates exhibit slowly decaying ACF, which is typical for long-memory, however with different decaying patterns: the fastest for the exchange rate, followed by the ones of the stocks, with BAC exhibiting the slowest decaying ACF pattern.

The histograms and the QQ-plots given in figures C.3-C.5 and C.6-C.8, respectively, show that the estimates for the exchange rates are closer to the normal distribution than the ones of the stocks as they exhibit less extreme values, which is in line with the results presented in tables B.2 and B.3. Among the stocks, BAC seems to have the most pronounced tails in VaR estimates, which indicates that this asset is particularly risky compared to the other two.

3.2 In-Sample Results

For the assessment of the empirical performance of our new methods to accurately estimate daily VaR and ES, we compare them against (1) a standard location-scale approach with constant conditional mean, the conditional variance stemming from a GARCH(1,1) specification and the Student's t with estimated degrees of freedom for the distribution of the standardized residuals of daily returns (we denote this specification the GARCH-t) and (2) the multiscaling approach of Hallam and Olmo (2014*a*) (denoted here HOMF) that estimates the whole density function of daily returns from matching the scaled up moments of intraday returns sampled in CTS to the moments derived from a parametric distributional assumption for daily returns. From these estimated densities, we compute the quantiles and ES that we need for our comparison. While the GARCH-t approach is a standard choice within the the location-scale framework to estimate and forecast VaR (see Halbleib and Pohlmeier (2012), among others), HOMF may be regarded as being the closest to our methods, as it develops from the assumption that the logarithm price process is multifractal. However, different from our approach, it uses the moment-multiscaling relationships typical to multifractality applied to data sampled in the clock time, ignoring, thus,any type of market's activity information.[14]

[13] Given that the graphs for the daily estimated ES are similar to the ones of VaR, they are not presented here. However, they can be obtained from the authors upon request.

[14] For comparison reasons, we also implement a location-scale model where the conditional variance stems from RV specifications instead of GARCH, as well as the unifractal-based method of Hallam and Olmo (2014*b*) that estimates the density of daily returns by scaling up the one of intraday returns sampled in CTS. However, given that the results for these methods are not significantly different from the ones of the comparative models mentioned above, we choose to disregard them for the present analysis.

To evaluate the performance of the models in accurately estimating VaR and ES, we use strictly consistent scoring (loss) functions as described by Gneiting (2011). These scoring functions allow for a reliable ranking of the forecasts, assuring that the forecast with the smallest loss value is the best. More precisely to make the comparison for quantiles, we employ the strictly consistent *asymmetric piece-wise linear scoring function* (sometimes also called *check loss* function or the *tick loss* function), which is the most popular scoring function for quantiles according to Gneiting (2011) and which is given by

$$S_p^{VaR}\big(\widehat{Q}_{p,d}, r_d\big) = (r_d - \widehat{Q}_{p,d})(p - \mathbb{1}_{\{\widehat{Q}_{p,d} \geq r_d\}}). \tag{14}$$

As there exists no strictly consistent scoring function for ES alone (Gneiting, 2011), we choose to use a joint scoring function for the pair VaR and ES as proposed by Fissler and Ziegel (2016), namely:

$$S_p^{VaRES}\big(\widehat{Q}_{p,d}, \widehat{ES}_{p,d}, r_d\big) = \frac{1}{\widehat{ES}_{p,d}}\left(\widehat{ES}_{p,d} - \widehat{Q}_{p,d} + \frac{(\widehat{Q}_{p,d} - r_d)\mathbb{1}_{\{r_d \leq \widehat{Q}_{p,d}\}}}{p}\right) + \ln(-\widehat{ES}_{p,d}). \tag{15}$$

Thus, over the whole evaluation window composed of $\mathcal{D}$ days, we compute the final scoring functions as the average over all days:

$$\bar{S}_p^{VaR} = \frac{1}{\mathcal{D}}\sum_{d=1}^{\mathcal{D}} S_p^{VaR}\big(\widehat{Q}_{p,d}, r_d\big), \tag{16}$$

$$\bar{S}_p^{VaRES} = \frac{1}{\mathcal{D}}\sum_{d=1}^{\mathcal{D}} S_p^{VaRES}\big(\widehat{Q}_{p,d}, \widehat{ES}_{p,d}, r_d\big), \tag{17}$$

whereas the model with the smallest average score exhibits the best performance. In order to test for differences in the scoring functions among the models, we implement the Diebold-Mariano (DM) test originally developed by Diebold and Mariano (1995).

Table 1 provides the VaR and VaRES scores and the results from the DM test with the ScaVaR/ScaES computed from data sampled in TTS as the benchmark. As one may see from the table, the ScaVaR/ScaES provides the smallest score values for all p's and assets. The TTS provides in general the most improvements compared to CTS.

For stocks, the DM-test cannot reject that the quality of the estimates between the two sampling schemes is the same. However, for the exchange rate compared to the stocks, it seems that TTS provides more improvements in terms of score values compared to CTS, especially when estimating extreme risks ($p = 1\%$). This may be due to the fact that the intraday trading pattern may not be as informative for stocks as it is for the exchange rates. To improve the results for stocks, one should consider other types of intrinsic time sampling schemes, built on intensity measures, which may be more informative for stocks, such as, for instance, the intraday volatility pattern.[15]

[15]It is a well-known fact that, during the trading period, stocks exhibit a pronounced "U"-shape in the

Table 1: In-sample results: VaR and VaRES scores. The entries in bold correspond to the smallest score values. The entries in parentheses give the p-values of the DM test with the ScaVaR/ScaES computed from data sampled in TTS as the benchmark. The VaR score values are scaled by 10^4.

Model	1%		2.5%		5%	
	VaR	VaRES	VaR	VaRES	VaR	VaRES
IBM						
ScaVaR/ScaES-TTS	**3.4056** (-)	-3.5346 (-)	**6.4961** (-)	**-3.7759** (-)	**10.4120** (-)	**-3.9944** (-)
ScaVaR/ScaES-CTS	3.4059 (0.9841)	**-3.5356** (0.8291)	6.5118 (0.5618)	-3.7754 (0.8886)	10.4510 (0.4599)	-3.9912 (0.3753)
HOMF	3.7214 (0.0624)	-3.2112 (0.0004)	7.3741 (0.0001)	-3.5301 (0.0000)	12.2920 (0.0000)	-3.7559 (0.0000)
GARCH-t	3.7186 (0.0990)	-3.3284 (0.0024)	7.4421 (0.0006)	-3.5654 (0.0000)	12.3910 (0.0000)	-3.7676 (0.0000)
BAC						
ScaVaR/ScaES-TTS	**5.1814** (-)	**-3.2071** (-)	**9.7204** (-)	**-3.4406** (-)	15.417 (-)	-3.6576 (-)
ScaVaR/ScaES-CTS	5.2077 (0.2714)	-3.2021 (0.1493)	9.7367 (0.8397)	-3.4438 (0.5134)	**15.451** (0.7922)	**-3.66** (0.6010)
HOMF	6.8742 (0.0003)	-2.6057 (0.0000)	13.1520 (0.0000)	-3.0221 (0.0000)	21.2900 (0.0000)	-3.2861 (0.0000)
GARCH-t	6.4913 (0.0025)	-2.8521 (0.0002)	13.183 (0.0000)	-3.1053 (0.0000)	21.592 (0.0000)	-3.3196 (0.0000)
EURUSD						
ScaVaR/ScaES-TTS	**1.8739** (-)	**-4.0261** (-)	**3.5462** (-)	**-4.2857** (-)	5.8535 (-)	**-4.4889** (-)
ScaVaR/ScaES-CTS	1.9234 (0.0000)	-3.9945 (0.0000)	3.5713 (0.3474)	-4.2712 (0.0003)	**5.8474** (0.8938)	-4.4865 (0.5620)
HOMF	2.0127 (0.2812)	-3.8267 (0.1089)	4.0978 (0.0011)	-4.0852 (0.0035)	6.9601 (0.0000)	-4.2643 (0.0000)
GARCH-t	1.9607 (0.4922)	-3.9002 (0.2077)	4.0799 (0.0029)	-4.1044 (0.0037)	6.991 (0.0000)	-4.2682 (0.0000)

Among the alternatives, the HOMF performs worst in terms of the score values, which indicates that estimating the whole density by scaling up intraday moments and using data sampled in CTS seems not to be appropriate for the estimation of tail risks.

3.3 Out-of-Sample Results

We compute one-step ahead forecasts of daily VaR and ES by using the rolling-window technique. The in-sample and out-of-sample windows are chosen such that they have the same lengths: i.e., for the foreign exchange rates, the out-of-sample window contains the last $\mathcal{D}^* = 1062$ days (for the period from May 7, 2008 until July 20, 2016) and for the stock data, the out-of-sample window contains the last $\mathcal{D}^* = 2062$ days (for the period from January 2, 2001 to July 24, 2017).

In order to make forecasts based on the daily estimates of VaR and ES stemming from ScaVaR and ScaES, we apply the HAR model of Corsi (2009) on the series $\hat{q}_{p,d}$ and $\widehat{es}_{p,d}$ with $d = 1, \ldots, \mathcal{D} - \mathcal{D}^* + 1$ in order to account for their long persistence depicted in Figure C.2.

Thus, the one-step ahead forecasts $\widehat{q}_{p,d+1|d}$ and $\widehat{es}_{p,d+1|d}$ are obtained in the following manner[16]:

$$\hat{q}_{p,d+1|d} = \hat{\beta}^{(0)} + \beta^{(1)}\hat{q}_{p,d} + \hat{\beta}^{(5)}\hat{q}^{(5)}_{p,d} + \hat{\beta}^{(22)}\hat{q}^{(22)}_{p,d}, \tag{18}$$

$$\widehat{es}_{p,d+1|d} = \hat{\gamma}^{(0)} + \hat{\gamma}^{(1)}\widehat{es}_{p,d} + \hat{\gamma}^{(5)}\widehat{es}^{(5)}_{p,d} + \hat{\gamma}^{(22)}\widehat{es}^{(22)}_{p,d}, \tag{19}$$

where $\hat{q}^{(m)}_{p,d} = \frac{1}{m}\sum_{j=1}^{m}\hat{q}_{p,d-j+1}$ and $\hat{es}^{(m)}_{p,d} = \frac{1}{m}\sum_{j=1}^{m}\hat{es}_{p,d-j+1}$ are the averages over the past $m = 1, 5, 22$ days. The one-step ahead forecasts of $\widehat{Q}_{p,d+1|d}$ and $\widehat{ES}_{p,d+1|d}$ are obtained by applying the transformation[17]:

$$\widehat{Q}_{p,d+1|d} = -\exp\left(\hat{q}_{p,d+1|d}\right), \tag{20}$$

$$\widehat{ES}_{p,d+1|d} = -\exp\left(\widehat{es}_{p,d+1|d}\right). \tag{21}$$

intraday volatility pattern (see Harris (1986) and Andersen and Bollerslev (1997), among others). For this reason, in a preliminary attempt, besides further trading-based sampling schemes following Engle and Russell (1998), Griffin and Oomen (2008), Oomen (2005) Oomen (2006), we also implement some sampling schemes based on the intraday volatility pattern as described by Boudt et al. (2011) and Dong and Tse (2014). Our preliminary results show the volatility-based sampling scheme provides better score values for estimating and forecasting the VaR and ES for stocks compared to the exchange rate. A deeper empirical analysis of which type of sampling scheme is more valuable for which type of asset is left, however, for further research.

[16] Alternatively, one could consider other specifications, such as Autoregressive Fractional Integrated Moving Average (ARFIMA) model as applied by Andersen, Bollerslev, Diebold and Labys (2001) or MIDAS approach of Ghysels et al. (2006) to capture the long persistence dynamics similarly to realized volatilities.

[17] Note that due to the transformation, the forecasts are biased. A bias correction as the one proposed by Bianchi and Calzolari (1980) and Oomen (2001) adds, however, further estimation noise to the forecasts (see e.g., Halbleib and Voev (2011)) and, therefore, for the moment, we decide to ignore it in our empirical exercise.

Table 2: Forecasting results: VaR and VaRES scores. The entries in bold correspond to the smallest score values. The entries in parentheses give the p-values of the DM test with the ScaVaR/ScaES computed from data sampled in TTS as the benchmark. The VaR score values are scaled by 10^4.

Model	1%		2.5%		5%	
	VaR	VaRES	VaR	VaRES	VaR	VaRES
IBM						
ScaVaR/ScaES-TTS	**2.8612** (-)	**-3.5967** (-)	**5.6670** (-)	**-3.8291** (-)	**9.4745** (-)	**-4.0152** (-)
ScaVaR/ScaES-CTS	2.8819 (0.0216)	-3.5890 (0.0001)	5.6899 (0.1461)	-3.8235 (0.0212)	9.4769 (0.8821)	-4.0128 (0.1882)
HOMF	3.1472 (0.1825)	-3.2739 (0.0280)	6.0910 (0.0677)	-3.6506 (0.0162)	9.9882 (0.0270)	-3.9038 (0.0112)
GARCH-t	2.9129 (0.7721)	-3.5381 (0.4937)	5.9644 (0.2068)	-3.7435 (0.1276)	9.8456 (0.1576)	-3.9478 (0.0871)
BAC						
ScaVaR/ScaES-TTS	5.8081 (-)	-2.9161 (-)	**11.8621** (-)	**-3.1344** (-)	19.4769 (-)	**-3.3272** (-)
ScaVaR/ScaES-CTS	**5.7742** (0.2834)	**-2.9201** (0.2633)	11.9429 (0.9154)	-3.1298 (0.9082)	**19.4757** (0.9014)	-3.3252 (0.4988)
HOMF	6.9563 (0.0251)	-2.4055 (0.0026)	12.9338 (0.0196)	-2.8758 (0.0010)	20.7714 (0.0009)	-3.1553 (0.0001)
GARCH-t	6.3756 (0.1701)	-2.7287 (0.0888)	12.5178 (0.0973)	-3.0156 (0.0352)	20.5219 (0.0046)	-3.2265 (0.0045)
EURUSD						
ScaVaR/ScaES-TTS	1.7749 (-)	**-4.0216** (-)	**3.3912** (-)	**-4.3099** (-)	5.6807 (-)	-4.5006 (-)
ScaVaR/ScaES-CTS	1.8141 (0.0134)	-3.9928 (0.0746)	3.4053 (0.1146)	-4.2956 (0.0326)	**5.6523** (0.5552)	**-4.5015** (0.1819)
HOMF	1.8317 (0.5706)	-3.8836 (0.3541)	3.5365 (0.1911)	-4.2225 (0.2282)	5.9195 (0.0193)	-4.4249 (0.0525)
GARCH-t	**1.7657** (0.9655)	-3.9625 (0.6083)	3.4905 (0.3270)	-4.2533 (0.3666)	5.8475 (0.0737)	-4.4479 (0.1167)

The forecast evaluation is done similarly to the one from the in-sample analysis, i.e., by implementing the two score functions defined in equations (16) and (17) along with the DM test with the ScaVaR and ScaES with TTS as a benchmark. Table 2 provides the out-of-sample results. As one may see from the table, our models perform very good compared to the two alternatives: it provides in almost all cases the smallest score values. Between the two sampling schemes, it seems that the intrinsic one (TTS) provides the best results, especially for IBM. Whenever CTS delivers the smallest score values, the DM test cannot reject that they are equal to the ones stemming from TTS. The very good performance of our approaches is regardless of the choice of p. Among the alternatives, GARCH-t provides the smallest scores for all p's. In one case, GARCH-t has a score value smaller than of our ScaVaR/ScaES computed based on TTS, however the DM test does not reject that they are equal.

4 Conclusions

We provide a novel approach of estimating and forecasting daily VaR and ES directly from HF data. Our method assumes that financial logarithm price processes are subordinated unifractal ones (which are driven by a single scaling law) in an alternative time dimension, denoted as the intrinsic time. This time dimension transforms the clock time according to an intensity measure that captures the real "heartbeat" of the market's activity. We show theoretically that this assumption is very general by proving that, under certain conditions, the unifractal process in intrinsic time is a fact multifractal one. The multifractality is very appealing for financial returns as it accommodates important empirical feature, such as volatility clustering and fat-tailedness, being, thus, more general and more realistic than existing ones, such as the Brownian Motion or the fractional Brownian Motion.

In practice, our method is very simple to implement and consists in computing daily estimates of VaR and ES by scaling up their intraday empirical counterparts computed from returns sampled in intrinsic time. In our empirical exercise, we describe the statistical and dynamic properties of these estimates and show that our method outperforms the standard ones in accurately estimating and forecasting daily VaR and ES. In particular, the empirical results show that data sampled in intrinsic time in accordance with the intraday trading pattern is very valuable compared to the calendar one when estimating and forecasting extreme risks.

These results are very promising and of great importance in the light of the current challenges on forecasting financial risks, but also in the light of the huge amount of high frequency data (recorded in nanoseconds) that has become available lately. They open a new agenda of research on how to exploit at best the richness of the information content of high-frequency data from another time perspective for risk estimation and forecasting, but also for other financial purposes. Appropriate mathematical frameworks for this type of analysis are provided by subordination models that, however, have found so far only little application in the finance and economics literature, as they are in general built on heavily parametric assumptions and estimated from less informative low frequency data. Thus, the current paper attempts to "revive" the concept of subordination in the new context of

abundant HF information available for academic research (Marinelli et al., 2000, 2001; Oomen, 2005, 2006; Dahlhaus and Tunyavetchakit, 2018).

From our current analysis, various immediate questions arise that are left to be answered by the future research, such as: How one should deal with the market microstructure noise in the subordinating framework? Some own empirical attempts, e.g., by implementing a subsampling type of estimator in the spirit of Zhang et al. (2005) in order to compute the quantiles of intraday returns, provide promising empirical results that need to be further exploited. Which intensity measure (such as based on volatility patterns, volumes, price changes, transaction patterns) and resulting intrinsic time-based sampling scheme are mostly useful to forecast financial risks? Which are the dynamic and statistical properties of the data sampled in intrinsic time and of the measures (volatilities, quantiles, etc) computed from it? The answers to these questions help at developing a new understanding on how financial markets function and how their risks can be measured and predicted by changing the (time) perspective.

Acknowledgements

For helpful comments we would like to thank Sebastian Bayer, Mark Hallam, Jens Jackwerth, Winfried Pohlmeier and members of the Chair of Econometrics at the Department of Economics at University of Konstanz, Germany. All remaining errors are ours. Timo Dimitriadis gratefully acknowledges financial support from the WIN-Kolleg of the Heidelberg Academy of Sciences and Humanities, Germany and from the Graduate School of Decision Sciences (GSDS) at University of Konstanz, Germany. Roxana Halbleib gratefully acknowledges financial support from the European Social Fund, Ministry of Science, Research and Arts of Baden–Württemberg, Germany, from the Deutsche Forschungsgemeinschaft through the project PO375/11-1 titled ”Robust Risk Measures in Real Time Settings” and from the Zukunftskolleg at University of Konstanz.

References

Alessio, E., Carbone, A., Castelli, G. and Frappietro, V. (2002), 'Second-order moving average and scaling of stochastic time series', *The European Physical Journal B - Condensed Matter and Complex Systems* **27**(2), 197–200.

Andersen, T. G. and Bollerslev, T. (1997), 'Intraday periodicity and volatility persistence in financial markets', *Journal of Empirical Finance* **4**, 115–158.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001), 'The distribution of realized exchange rate volatility', *Journal of the American Statistical Association* **96**, 42–55.

Bandi, F. M. and Russell, J. R. (2005), Microstructure noise, realized volatility and optimal sampling. Working Paper, Graduate School of Business, The University of Chicago.

Barndorff-Nielsen, O. E., Hansen, P., Lunde, A. and Shephard, N. (2008), 'Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise', *Econometrica* **76**(6), 14811536.

Basel Committee (1996), Overview over the amendment to the capital accord to incorporate market risks, Technical report, Basel Committee on Banking Supervision.

Basel Committee (2016), *Minimum capital requirements for Market Risk*.

Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013), *Long-Memory Processes: Probabilistic Properties and Statistical Methods*, Springer.

Bianchi, C. and Calzolari, G. (1980), 'The one-period forecast errors in nonlinear econometric models', **21**, 201–208.

Black, F. and Scholes, M. S. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**(3), 637–54.

Boudt, K., Croux, C. and Laurent, S. (2011), 'Robust estimation of intraweek periodicity in volatility and jump detection', *Journal of Empirical Finance* **18**, 353–367.

Calvet, L. E. and Fisher, A. J. (2001), 'Forecasting multifractal volatility', *Journal of Econometrics* **105**(1), 27–58.

Calvet, L. and Fisher, A. (2002), 'Multifractality in asset returns: Theory and evidence', *Review of Economics and Statistics* **84**(3), 381–406.

Calvet, L. and Fisher, A. (2004), 'How to forecast long-run volatility: Regime switching and the estimation of multifractal processes', *Journal of Financial Econometrics* **2**(1), 49–83.

Calvet, L. and Fisher, A. (2008), *Multifractal Volatility: Theory, Forecasting, and Pricing*, Academic Press.

Calvet, L., Fisher, A. and Mandelbrot, B. (1997), Large deviation and the distribution of price changes. Cowles Foundation Discussion Paper.

Clark, P. K. (1973), 'A subordinated stochastic process model with finite variance for speculative prices', *Econometrica* **41**, 135–155.

Comte, F. and Renault, E. (1996), 'Long memory continuous time models', *Journal of Econometrics* **73**, 101–149.

Comte, F. and Renault, E. (1998), 'Long memory in continuous-time stochastic volatility models', *Mathematical Finance* **8**, 291–323.

Corsi, F. (2009), 'A simple approximate long-memory model of realized volatility', *Journal of Financial Econometrics* **7**, 174–196.

Dahlhaus, R. and Tunyavetchakit, S. (2018), Volatility decomposition and estimation in time-changed price models. arXiv:1605.02205.

Diebold, F. and Mariano, R. (1995), 'Comparing predictive accuracy', *Journal of Business and Economic Statistics* **13**(3), 253–63.

Dong, Y. and Tse, Y.-K. (2014), Business time sampling scheme and its application. Working Paper.

Eberlein, E. and Keller, K. (1995), 'Hyperbolic distributions in finance', *Bernoulli* **1**, 281–299.

Embrechts, P. and Maejima, M. (2002), *Selfsimilar Processes*, Princeton University Press.

Engle, R. F. and Russell, J. R. (1998), 'Autoregressive conditional duration: A new model for irregularly spaced transaction data', *Econometrica* **66**, 1127–1162.

Fissler, T. and Ziegel, J. F. (2016), 'Higher order elicitability and Osband's principle', *The Annals of Statistics* **44**(4), 1680–1707.

Gerhard, F. and Hautsch, N. (2002), 'Volatility estimation on the basis of price intensities', *Journal of Empirical Finance* **9**, 57–89.

Ghysels, E., Gouriéroux, C. and Jasiak, J. (1997), Trading patterns, time deformation and stochastic volatility in foreign exchange markets, *in* C. Dunis and B. Zhou, eds, 'Nonlinear Modelling in Foreign Exchange Markets', Wiley, New York.

Ghysels, E., Santa-Clara, P. and Valkanov, R. (2006), 'Predicting volatility: Getting the most out of return data sampled at different frequencies', *Journal of Econometrics* **131**, 59–95.

Giot, P. and Laurent, S. (2004), 'Modelling daily value-at-risk using realized volatility and arch type models', *Journal of Empirical Finance* **11**, 379 – 398.

Gneiting, T. (2011), 'Quantiles as optimal point forecasts', *International Journal of Forecasting* **27**, 197–207.

Gouriéroux, C. and Jasiak, J. (2001), *Financial Econometrics*, Princeton University Press, Princeton.

Griffin, J. and Oomen, R. (2008), 'Sampling Returns for Realized Variance Calculations: Tick Time or Transaction Time?', *Econometric Reviews* **27**(1-3), 230–253.

Guillaume, D. M., Dacorogna, M. M., Davé, R., Müller, U. A., Olsen, R. B. and Pictet, O. V. (1997), 'From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets', *Finance and Stochastics* **1**, 95–129.

Halbleib, R. and Pohlmeier, W. (2012), 'Improving the value at risk forecasts: Theory and evidence from the financial crisis', *Journal of Economic Dynamics and Control* **36**, 1212–1228.

Halbleib, R. and Voev, V. (2011), 'Forecasting multivariate volatility using the varfima model on realized covariance cholesky factors', *Journal of Economics and Statistics* **231**, 134–152.

Hallam, M. and Olmo, J. (2014*a*), 'Forecasting daily return densities from intraday data: A multifractal approach', *International Journal of Forecasting* **30**, 863–881.

Hallam, M. and Olmo, J. (2014*b*), 'Semiparametric density forecasts of daily financial returns from intraday data', *Journal of Financial Econometrics* **12**(2), 408–432.

Hansen, P. R. and Lunde, A. (2004), Realized variance and iid market microstructure noise. Brown University Working Paper.

Hansen, P. R. and Lunde, A. (2006), 'Realized variance and market microstructure noise', *Journal of Business and Economic Statistics* **24**, 127–218.

Harrell, F. and Davis, E. (1982), 'A new distribution-free quantile estimator', *Biometrika* **69**, 635–640.

Harris, L. (1986), 'A transaction data study of weekly and intradaily patterns in stock returns', *Journal of Financial Economics* pp. 99–117.

Hurst, S. R., Platen, E. and Rachev, S. T. (1997), 'Subordinated market index models: A comparison', *Financial Engineering and the Japanese Markets* **4**, 97–124.

Kantelhardt, J. (2009), 'Fractal and Multifractal Time Series', *Encyclopedia of Complexity and Systems Science* pp. 463–487.

Kobeissi, Y. H. (2013), *Multifractal Financial Markets*, Springer.

Kuester, K., Mitinik, S. and Paolella, M. S. (2006), 'Value-at-risk prediction: A comparison of alternative strategies', *Journal of Financial Econometrics* **4**(1), 5389.

Lux, T. and Segnon, M. (2018), Multifractal models in finance: their origin, properties, and applications, *in* . M. K. Shu-Heng Chen and Y.-R. Du, eds, 'The Oxford Handbook of Computational Economics and Finance', Oxford University Press, chapter 6.

Mandelbrot, B. (1974), 'Intermittent turbulence in self similar cascades: Divergence of high moments and dimension of the carrier', *Journal of Fluid Mechanics* **62**, 331–358.

Mandelbrot, B. B. (1963), 'The variation of certain speculative prices', *Journal of Business* **36**, 394–419.

Mandelbrot, B. B. (1982), *The Fractal Geometry of Nature*, W. H. Freeman and Company, New York.

Mandelbrot, B. B. (1997), *Fractals and Scaling in Finance: Discontinuity, Concentration, Risk*, Springer, New York.

Mandelbrot, B. B. (1999), *Multifractals and 1/f Noise*, Springer Verlag, New York.

Mandelbrot, B. B., Fisher, A. J. and Calvet, L. E. (1997), A multifractal model of asset returns. Cowles Foundation Discussion Paper No. 1164, Yale University.

Mandelbrot, B. B. and Taylor, H. M. (1967), 'On the distribution of stock price differences', *Operations Research* **15**, 1057–1062.

Marinelli, C., Rachev, S. and Roll, R. (2001), 'Subordinated exchange rate models: Evidence for heavy tailed distributions and long-range dependence', *Mathematical and Computer Modelling* **34**, 955–1001.

Marinelli, C., Rachev, S., Roll, R. and Göppl, H. (2000), Subordinated stock price models: Heavy tails and long-range dependence in the high-frequency deutsche bank price record, *in* G. Bol, G. Nakhaeizadeh and K. Vollmer, eds, 'Datamining und Computational Finance. Wirtschaftswissenschaftliche Beiträge. vol 174', Physica, Heidelberg.

Markowitz, H. (1952), 'Portfolio selection', *Journal of Finance* **7**, 77–91.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Pictet, O. V., Olsen, R. B. and Ward, J. R. (1995), Fractals and intrinsic time - a challenge to econometricians. Working Paper Series, Olsen&Associates.

Oomen, R. C. A. (2001), Using high frequency stock market index data to calculate, model and forecast realized return variance. European University, Economics Discussion Paper No. 2001/6.

Oomen, R. C. A. (2005), 'Properties of bias-corrected realized variance under alternative sampling schemes', *Journal of Financial Econometrics* **3**(4), 555–577.

Oomen, R. C. A. (2006), 'Properties of realized variance under alternative sampling schemes', *Journal of Business and Economic Statistics* **24**(2), 219–237.

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E. and Goldberger, A. L. (1994), 'Mosaic organization of dna nucleotides', *Phys. Rev. E* **49**, 1685–1689.

Rachev, S. T. and Mittnik, S. (2000), *Stable Paretian Models in Finance*, John Wiley & Sons Ltd., England.

Sánchez-Granero, M. J., Fernández-Martínez, M. and Trinidad-Segovia, J. E. (2012), 'Introducing fractal dimension algorithms to calculate the hurst exponent of financial time series', *The European Physical Journal B* **85**(3).

Stock, J. (1988), 'Estimating continuous time processes subject to time deformation', *Journal of the American Statistical Association* **83**, 77–84.

Wu, Z. (2012), 'On the intraday periodicity duration adjustment of high-frequency data', *Journal of Empirical Finance* **19**, 282–291.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005), 'A tale of two time scales: Determining integrated volatility with noisy high frequency data', *Journal of the American Statistical Association* **100**, 1394–1411.

A Appendix: Proofs

Proof of Proposition 2.2. Let $t' = t/c$. Then, for all $t \geq 0$ and for all $\Delta > 0$, it holds that

$$U(t + c\Delta) - U(t) = U\big(c(\tilde{t} + \Delta)\big) - U(c\tilde{t}). \tag{A.22}$$

From the unifractal assumption of U, we get that

$$U\big(c(\tilde{t} + \Delta)\big) - U(c\tilde{t}) \stackrel{d}{=} c^H \left(U\big(\tilde{t} + \Delta\big) - U(\tilde{t})\right). \tag{A.23}$$

Given that we assume that the increments of U are stationary, we get that

$$c^H \left(U\big(\tilde{t} + \Delta\big) - U(\tilde{t})\right) \stackrel{d}{=} c^H \left(U\big(t + \Delta\big) - U(t)\right), \tag{A.24}$$

which completes the proof. □

Proof of Theorem 2.7. Let $\mathcal{T}$ be such that it is contained in the second component of $\mathfrak{D}$. Based on the definition of θ in Equation (1), we get that, for all $t, h \geq 0$,

$$\theta(t + h) - \theta(t) = \lambda\left([0, t + h]\right) - \lambda\left([0, t]\right) = \lambda\left([t, t + h]\right). \tag{A.25}$$

As $\lambda\left([t, t + h]\right) \geq 0$, then $\theta(t + h) \geq \theta(t)\ \forall t, h \geq 0$ almost surely, i.e., $\theta(t)$ has almost surely non-decreasing paths.

In order to show that $\theta(t)$ is a multifractal process, we make use of the second definition of the multifractality in Section 2.1 and show that (a) it has moment-scaling relationships as the ones given in Equation (5) and (b) it has stationary increments.

a) From Assumption 2.6 we get that for all $q \in \mathcal{Q}$ and for all $t \in \mathcal{T}$,

$$\mathbb{E}\left[\theta(t)^q\right] = \mathbb{E}\left[\lambda\left([0, t]\right)^q\right] = c_\lambda(q) t^{\tau_\lambda(q)+1}, \tag{A.26}$$

where $\mathcal{Q}$ is some interval on the real line and there exists $\varepsilon > 0$ such that $(-\varepsilon, 1] \in \mathcal{Q}$ and $c_\lambda(q)$ and $\tau_\lambda(q)$ are inherited from the measure λ.

b) From Equation (A.25), we get that

$$\theta(t + h) - \theta(t) = \lambda\left([t, t + h]\right) \quad \text{and} \tag{A.27}$$

$$\theta(s + h) - \theta(s) = \lambda\left([s, s + h]\right), \tag{A.28}$$

for all $(s, h), (t, h) \in \mathfrak{D}$.

Define $\eta[s, t] := \log(\lambda[s, t])$ and let $M_{\eta[t,t+h]}(q)$ be its moment generating function. Then, for all $q \in (-\varepsilon, \varepsilon)$, it holds that

$$M_{\eta[t,t+h]}(q) = \mathbb{E}\left[\exp(q \cdot \eta[t, t + h])\right] = \mathbb{E}\left[\exp(q \cdot \log(\lambda[t, t + h]))\right] \tag{A.29}$$

$$= \mathbb{E}\left[\lambda[t, t + h]^q\right] = c_\lambda(q) h^{\tau_\lambda(q)+1}, \tag{A.30}$$

and analogously

$$M_{\eta[s,s+h]}(q) = \mathbb{E}\left[\lambda[s, s + h]^q\right] = c_\lambda(q) h^{\tau_\lambda(q)+1}, \tag{A.31}$$

i.e., $M_{\eta[t,t+h]}(q) = M_{\eta[s,s+h]}(q)$ for all $q \in (-\varepsilon, \varepsilon)$. As the moment generating function on any interval $(-\varepsilon, \varepsilon)$ uniquely charactrizes the distribution of the underlying variable, we get that the distributions of $\eta[s, s + h]$ and $\eta[t, t + h]$ are identical for all s, t and for all $h > 0$. As the logarithm is a strictly increasing function, we also get that the distributions of $\lambda[s, s + h]$ and $\lambda[t, t + h]$ are identical for all s, t and for all $h > 0$, which implies that the process $\theta(t)$ defined in Equation (1) has stationary increments.

□

Proof of Theorem 2.8. For all $t \in \mathcal{T}$, it holds that

$$\mathbb{E}\left[\,|p(t)|^q \,|\, \theta(t) = u\right] = \mathbb{E}\left[\,|U(u)|^q \,|\, \theta(t) = u\right]. \tag{A.32}$$

For the unifractal process U, it holds that $U(u) \overset{d}{=} u^H U(1)$ and therefore

$$\mathbb{E}\left[\,|U(u)|^q \,|\, \theta(t) = u\right] = \mathbb{E}\left[\,|u^H U(1)|^q \,\middle|\, \theta(t) = u\right] = \theta(t)^{Hq}\,\mathbb{E}[|U(1)|^q], \tag{A.33}$$

as $U(t)$ and $\theta(t)$ are independent. By the law of iterated expectations, we get that

$$\begin{aligned}
\mathbb{E}\left[|p(t)|^q\right] &= \mathbb{E}\left[\mathbb{E}\left[\,|p(t)|^q \,|\, \theta(t) = u\right]\right] && \text{(A.34)}\\
&= \mathbb{E}\left[\theta(t)^{Hq}\mathbb{E}\left[|U(1)|^q\right]\right] && \text{(A.35)}\\
&= \mathbb{E}\left[\theta(t)^{Hq}\right]\mathbb{E}\left[|U(1)|^q\right]. && \text{(A.36)}
\end{aligned}$$

From the Theorem 2.7, we know that $\theta(t)$ is a multifractal process, i.e. $\mathbb{E}\left[\theta(t)^{Hq}\right] = c_\lambda(Hq)\, t^{\tau_\lambda(Hq)+1}$ and thus

$$\mathbb{E}\left[\theta(t)^{Hq}\right]\mathbb{E}\left[|U(1)|^q\right] = c_\theta(Hq)\mathbb{E}\left[|U(1)|^q\right] \cdot t^{\tau_\lambda(Hq)+1}, \tag{A.37}$$

which implies that the process $p(t)$ satisfies the multi-scaling law typical to multifractality with $\tau_p(q) = \tau_\lambda(Hq)$ and $c_p(q) = c_\lambda(Hq)\mathbb{E}\left[|U(1)|^q\right]$.

In the following, we show that the process $p(t) = U\big(\theta(t)\big)$ has stationary increments: i.e., for all $s, t \geq 0$, $t > s$ and for all $h > 0$, it holds that

$$U\big(\theta(t+h)\big) - U\big(\theta(s+h)\big) \quad \overset{d}{=} \quad U\big(\theta(t)\big) - U\big(\theta(s)\big). \tag{A.38}$$

In order to prove Equation (A.38), we show that

$$\begin{aligned}
U\big(\theta(t+h)\big) - U\big(\theta(s+h)\big) \quad &\overset{d}{=} \quad U\big(\theta(t+h) - \theta(s+h)\big) - U(0) && \text{(A.39)}\\
&\overset{d}{=} \quad U\big(\theta(t) - \theta(s)\big) - Y(0) && \text{(A.40)}\\
&\overset{d}{=} \quad u\big(\theta(t)\big) - U\big(\theta(s)\big). && \text{(A.41)}
\end{aligned}$$

Formally, the equality in distribution in (A.39) can be shown by conditioning on $\theta(s+h)$. For that, we assume that the real-valued stochastic process $p(t)$ is equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ and is defined on some complete probability space $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$. Let the measure of $\theta(s+h)$ be denoted by $\nu_{\theta(s+h)}$. Then, for all $s, t \geq 0$, $t > s$, for all $h > 0$ and for all $A \in \mathcal{B}(\mathbb{R})$, it holds that

$$\begin{aligned}
&\mathbb{P}\left(U\big(\theta(t+h)\big) - U\big(\theta(s+h)\big) \in A\right)\\
&= \int_{\mathbb{R}_+} \mathbb{P}\left(U\big(\theta(t+h)\big) - U\big(\delta\big) \in A \,\middle|\, \theta(s+h) = \delta\right) \mathrm{d}\nu_{\theta(s+h)}(\delta) && \text{(A.42)}\\
&= \int_{\mathbb{R}_+} \mathbb{P}\left(U\big(\theta(t+h) - \delta\big) - U\big(0\big) \in A \,\middle|\, \theta(s+h) = \delta\right) \mathrm{d}\nu_{\theta(s+h)}(\delta) && \text{(A.43)}\\
&= \mathbb{P}\left(U\big(\theta(t+h) - \theta(s+h)\big) - U\big(0\big) \in A\right), && \text{(A.44)}
\end{aligned}$$

where we use in (A.43) the fact that the process $U(t)$ has stationary increments. The third equality in (A.41) can be shown equivalently by conditioning on $\theta(s)$.

For the equality in (A.40), it holds that for all $s, tt \geq 0$, $t > s$, for all $h > 0$ and for all $A \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}
&\mathbb{P}\left(U\big(\theta(t+h)-\theta(s+h)\big)-U\big(0\big)\in A\right)\\
&=\int_{\mathbb{R}_+}\mathbb{P}\left(U\big(\eta\big)-U\big(0\big)\in A\big|\,\theta(t+h)-\theta(s+h)=\eta\right)\mathrm{d}\nu_{\theta(t+h)-\theta(s+h)}(\eta) &\text{(A.45)}\\
&=\int_{\mathbb{R}_+}\mathbb{P}\left(U\big(\eta\big)-U\big(0\big)\in A\right)\mathrm{d}\nu_{\theta(t)-\theta(s)}(\eta) &\text{(A.46)}\\
&=\int_{\mathbb{R}_+}\mathbb{P}\left(U\big(\eta\big)-U\big(0\big)\in A\big|\,\theta(t)-\theta(s)=\eta\right)\mathrm{d}\nu_{\theta(t)-\theta(s)}(\eta) &\text{(A.47)}\\
&=\mathbb{P}\left(U\big(\theta(t)-\theta(s)\big)-U\big(0\big)\in A\right), &\text{(A.48)}
\end{aligned}$$

where we use in (A.46) the fact that the process $\theta(t)$ has stationary increments and it is independent of $U(t)$ and in (A.47) the fact that the processes $U(t)$ and $\theta(t)$ are independent. □

B Appendix: Tables

Table B.1: Descriptive statistics of the intraday returns for the foreign exchange rates over the window May 7, 2008 - July 20, 2016 and for the stocks over the window January 2, 2001-July 24, 2017 computed at 5 minute frequency and sampled by two sampling schemes.

	EURUSD		IBM		BAC	
	CTS	TTS	CTS	TTS	CTS	TTS
Mean	-5.49E-07	-5.49E-07	9.54E-06	9.54E-06	-1.10E-05	-1.10E-05
Median	0.0000	0.000	0.0000	0.0000	0.0000	0.0000
Maximum	0.0150	0.0137	0.0759	0.0759	0.0684	0.0765
Minimim	-0.0117	-0.0121	-0.0380	-0.0380	-0.0929	-0.0783
Std. dev	0.0004	0.0004	0.0014	0.0014	0.0025	0.0025
Skewness	0.1013	-0.0015	0.5781	0.5781	-0.7884	-0.1087
Kurtosis	29.6087	20.1280	48.6243	48.6243	76.1225	56.9787

Table B.2: Descriptive statistics of $\widehat{q}_{p,d}$ for the foreign exchange rates over the window May 7, 2008 - July 20, 2016 and for the stocks over the window January 2, 2001-July 24, 2017

Sampling scheme	CTS			TTS		
Probability p	1%	2.5%	5%	1%	2.5%	5%
	EURUSD					
mean	-6.9670	-7.2699	-7.5406	-7.0042	-7.2884	-7.5406
variance	0.1778	0.1713	0.1743	0.1840	0.1732	0.1696
skewness	-0.0710	-0.0461	-0.0593	-0.0072	-0.0196	-0.0340
kurtosis	3.6626	3.5845	3.5237	3.6346	3.5856	3.4557
JB test-statistic	40.6535	31.0038	25.5263	35.6722	30.4994	18.7972
	IBM					
mean	-5.8222	-6.0944	-6.3459	-5.8254	-6.0991	-6.3494
variance	0.2497	0.2297	0.2642	0.2479	0.2296	0.2275
skewness	0.5722	0.6603	0.6946	0.5468	0.6244	0.6576
kurtosis	3.6888	3.7661	3.6402	3.6367	3.7629	3.6778
JB test-statistic	306.556	400.517	402.02	275.191	367.986	376.185
	BAC					
mean	-5.5273	-5.8033	-6.0581	-5.5334	-5.8069	-6.0575
variance	0.4382	0.3920	0.4583	0.4369	0.3965	0.3834
skewness	0.7855	0.8503	0.8179	0.7952	0.8300	0.8301
kurtosis	4.1658	4.3489	4.1832	4.2087	4.3102	4.3156
JB test-statistic	657.632	809.626	700.317	685.622	768.469	771.082

Table B.3: Descriptive statistics of $\widehat{es}_{p,d}$ for the foreign exchange rates over the window May 7, 2008 - July 20, 2016 and for the stocks over the window January 2, 2001-July 24, 2017

Sampling scheme	CTS			TTS		
Probability p	1%	2.5%	5%	1%	2.5%	5%
	EURUSD					
mean	-6.7249	-6.9328	-7.1361	-6.7673	-6.9683	-7.1625
variance	0.2018	0.1779	0.1687	0.2091	0.1847	0.1728
skewness	0.0256	-0.0529	-0.0802	0.0762	0.0057	-0.0238
kurtosis	3.6180	3.6077	3.6256	3.4211	3.5194	3.5613
JB test-statistic	34.0504	33.6855	36.9249	17.7567	23.8947	28.0986
	IBM					
mean	-5.7586	-5.8774	-6.0389	-5.7617	-5.8807	-6.0428
variance	0.2642	0.2406	0.2253	0.2630	0.2384	0.2230
skewness	0.5408	0.6029	0.6707	0.5162	0.5785	0.6384
kurtosis	3.6064	3.7502	3.8106	3.5437	3.7139	3.7617
JB test-statistic	264.228	346.577	422.147	233.925	317.584	379.783
	BAC					
mean	-5.4636	-5.5838	-5.7493	-5.4703	-5.5891	-5.7522
variance	0.4583	0.4248	0.4024	0.4573	0.4234	0.3975
skewness	0.7563	0.8094	0.8516	0.7740	0.8084	0.8523
kurtosis	4.0682	4.2408	4.3109	4.1412	4.2564	4.3626
JB test-statistic	589.184	714.833	793.821	635.514	720.376	818.307

C Appendix: Figures

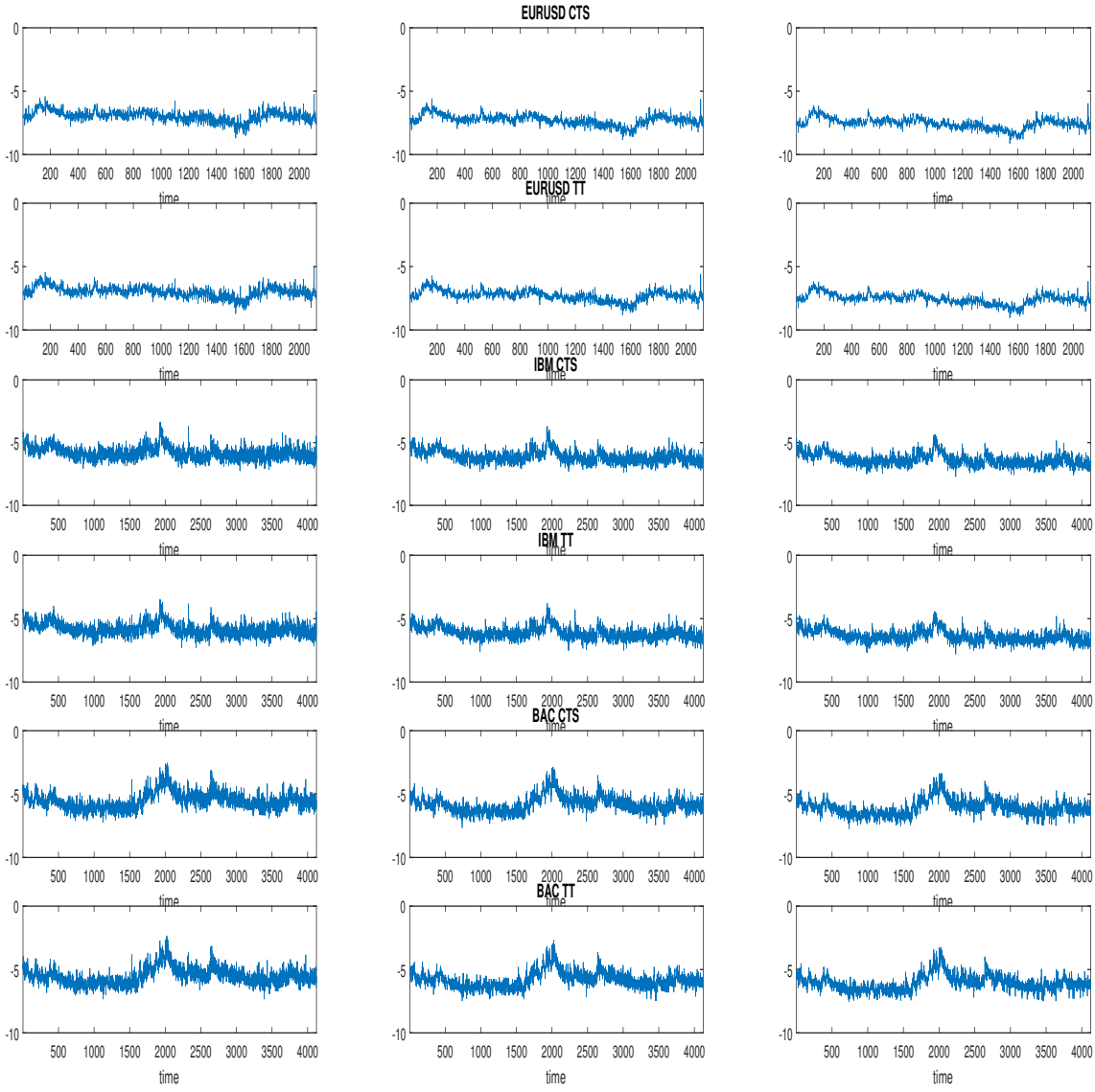

Figure C.1: Line graph of the log-transformation of the negative estimates of daily quantiles, $\widehat{q}_{p,d}$, computed from scaling up estimates of intraday quantiles estimated from data sampled in TTS and CTS (use the DMA estimator of Hurst coefficient and the frequency of 5 minute to sample the data). The first column corresponds to $p = 1\%$, the second column corresponds to $p = 2.5\%$ and the third column corresponds to $p = 5\%$.

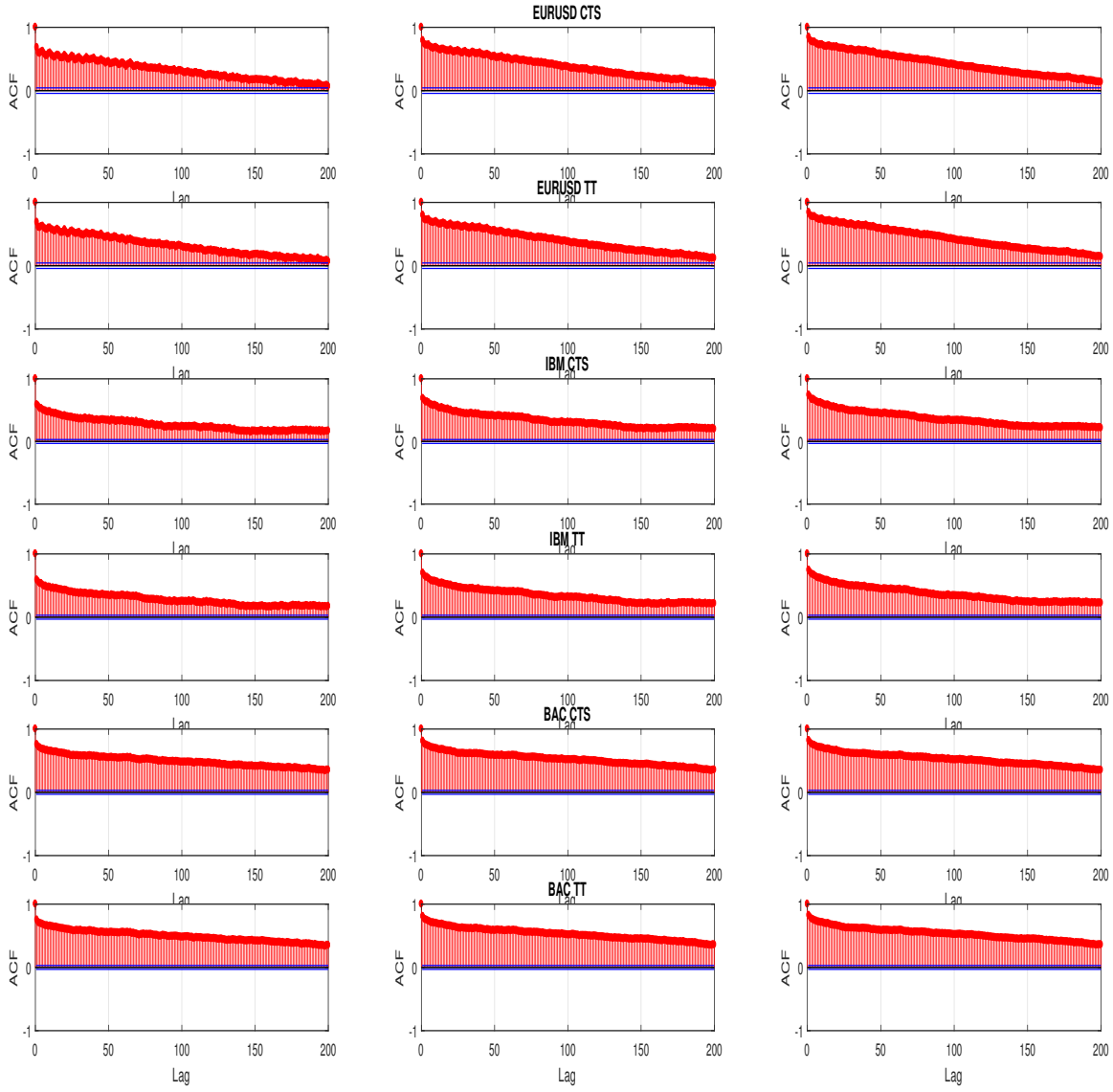

Figure C.2: ACF of the log-transformation of the negative estimates of daily quantiles, $\widehat{q}_{p,d}$, computed from scaling up estimates of intraday quantiles estimated from data sampled in TTS and CTS (use the DMA estimator of Hurst coefficient and the frequency of 5 minute to sample the data). The first column corresponds to $p = 1\%$, the second column corresponds to $p = 2.5\%$ and the third column corresponds to $p = 5\%$. The blue line represents the 95% confidence interval.

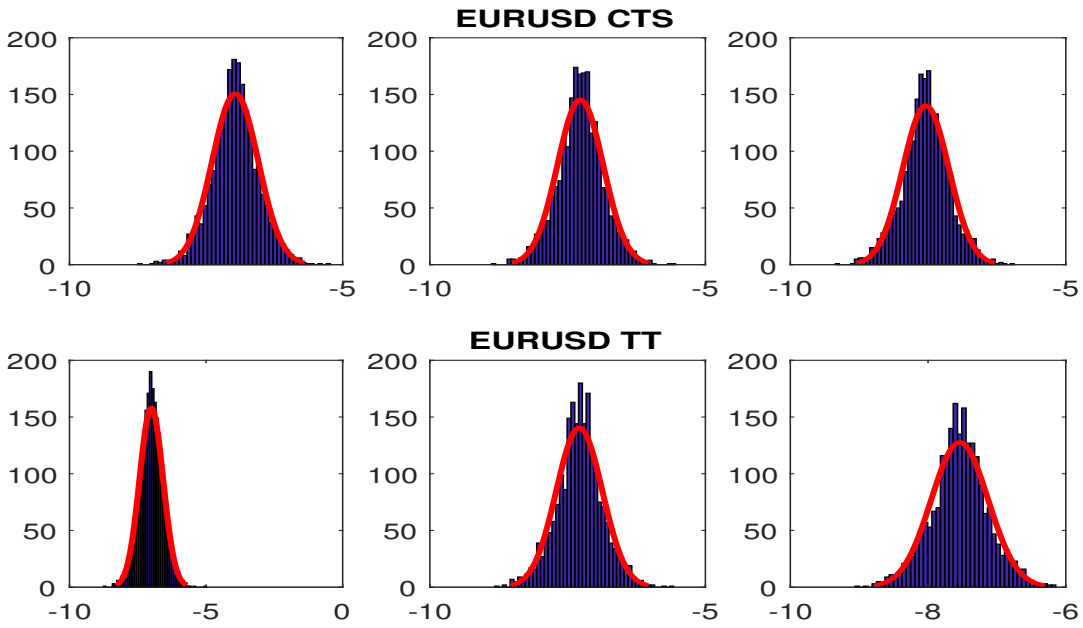

Figure C.3: Histogram of the log-transformation of the negative estimates of daily quantiles, $\widehat{q}_{p,d}$ of EURUSD, computed based on the daily estimates of the quantiles as described in the main text from scaling up estimates of intraday quantiles estimated from data sampled in TTS and CTS (use the DMA estimator of Hurst coefficient and the frequency of 5 minute to sample the data). The first column corresponds to $p = 1\%$, the second column corresponds to $p = 2.5\%$ and the third column corresponds to $p = 5\%$. The red line represents the density function of a normal distribution.

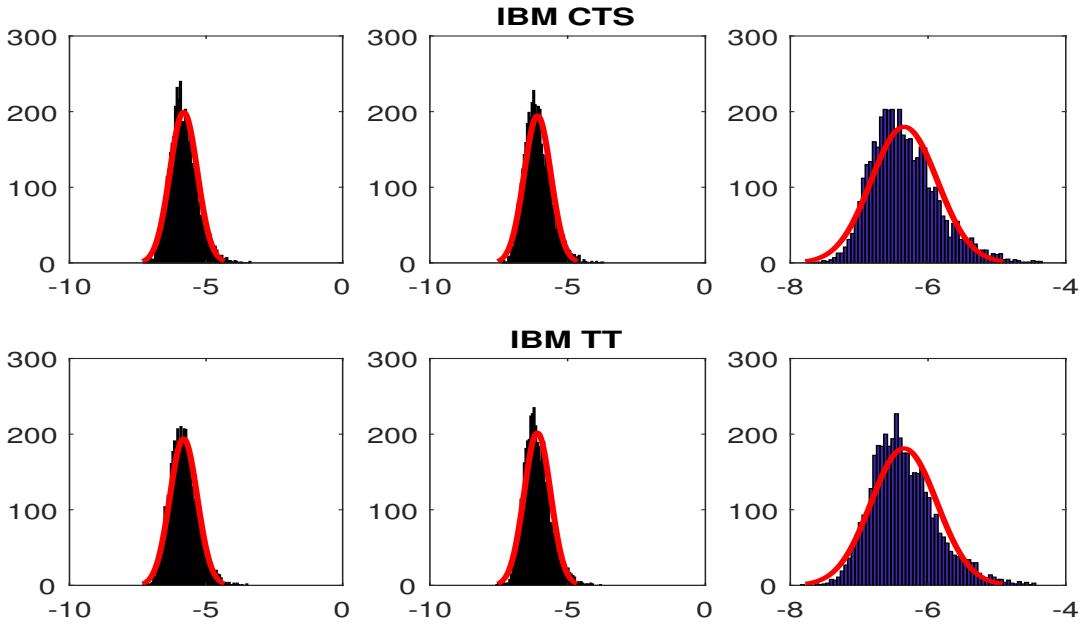

Figure C.4: Histogram of the log-transformation of the negative estimates of daily quantiles, $\widehat{q}_{p,d}$ of IBM, computed based on the daily estimates of the quantiles as described in the main text from scaling up estimates of intraday quantiles estimated from data sampled in TTS and CTS (use the DMA estimator of Hurst coefficient and the frequency of 5 minute to sample the data). The first column corresponds to $p = 1\%$, the second column corresponds to $p = 2.5\%$ and the third column corresponds to $p = 5\%$. The red line represents the density function of a normal distribution.

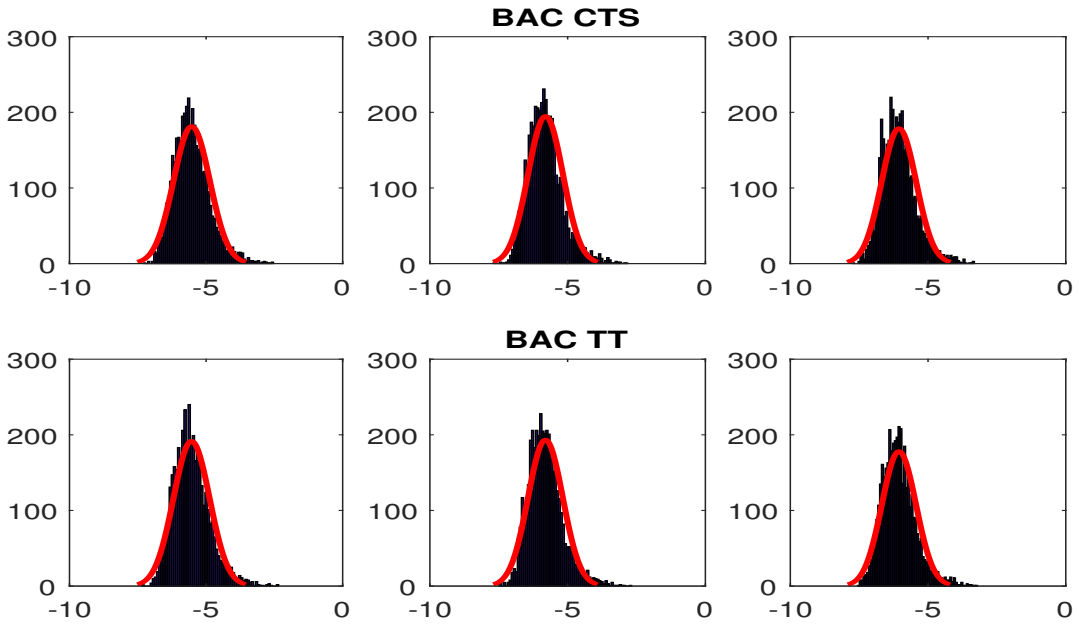

Figure C.5: Histogram of the log-transformation of the negative estimates of daily quantiles, $\widehat{q}_{p,d}$ of BAC, computed based on the daily estimates of the quantiles as described in the main text from scaling up estimates of intraday quantiles estimated from data sampled in TTS and CTS (use the DMA estimator of Hurst coefficient and the frequency of 5 minute to sample the data). The first column corresponds to $p = 1\%$, the second column corresponds to $p = 2.5\%$ and the third column corresponds to $p = 5\%$. The red line represents the density function of a normal distribution.

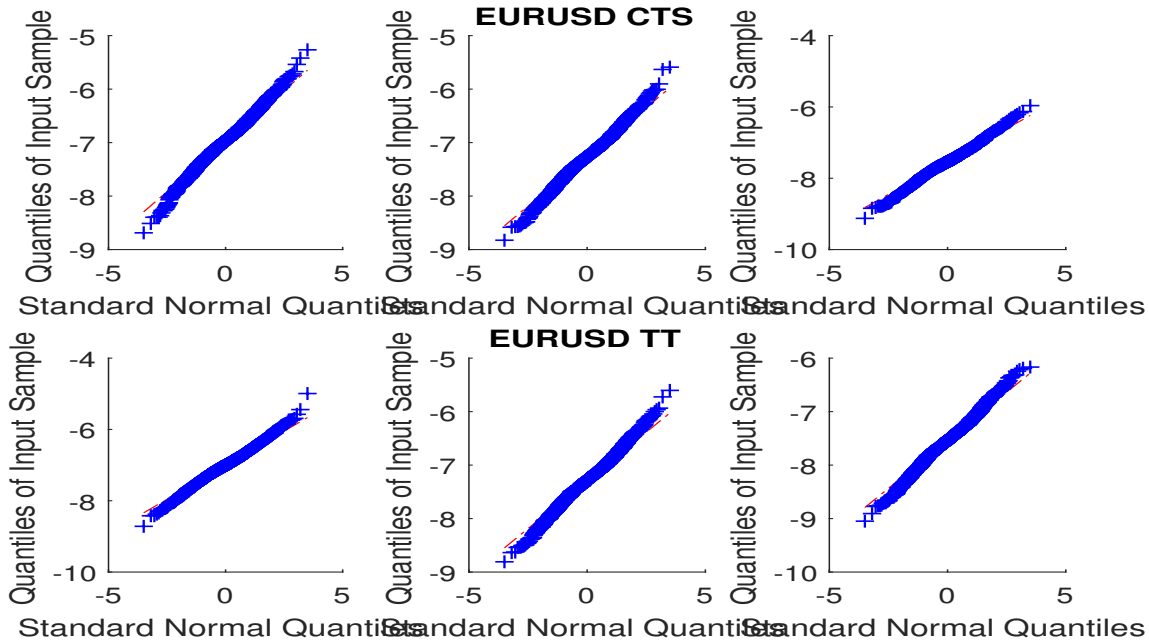

Figure C.6: QQ-plots of the log-transformation of the negative estimates of daily quantiles, $\widehat{q}_{p,d}$ of EURUSD, computed based on the daily estimates of the quantiles as described in the main text from scaling up estimates of intraday quantiles estimated from data sampled in TTS and CTS (use the DMA estimator of Hurst coefficient and the frequency of 5 minute to sample the data). The first column corresponds to $p = 1\%$, the second column corresponds to $p = 2.5\%$ and the third column corresponds to $p = 5\%$. The red line corresponds to the standard normal distribution.

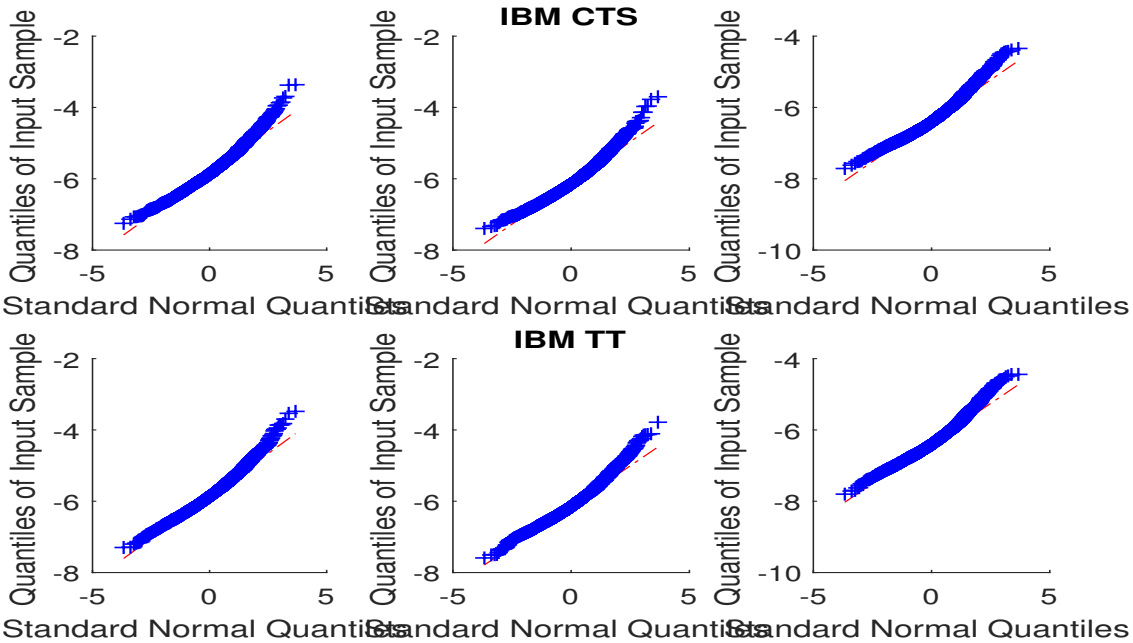

Figure C.7: QQ-plots of the log-transformation of the negative estimates of daily quantiles, $\widehat{q}_{p,d}$ of IBM, computed based on the daily estimates of the quantiles as described in the main text from scaling up estimates of intraday quantiles estimated from data sampled in TTS and CTS (use the DMA estimator of Hurst coefficient and the frequency of 5 minute to sample the data). The first

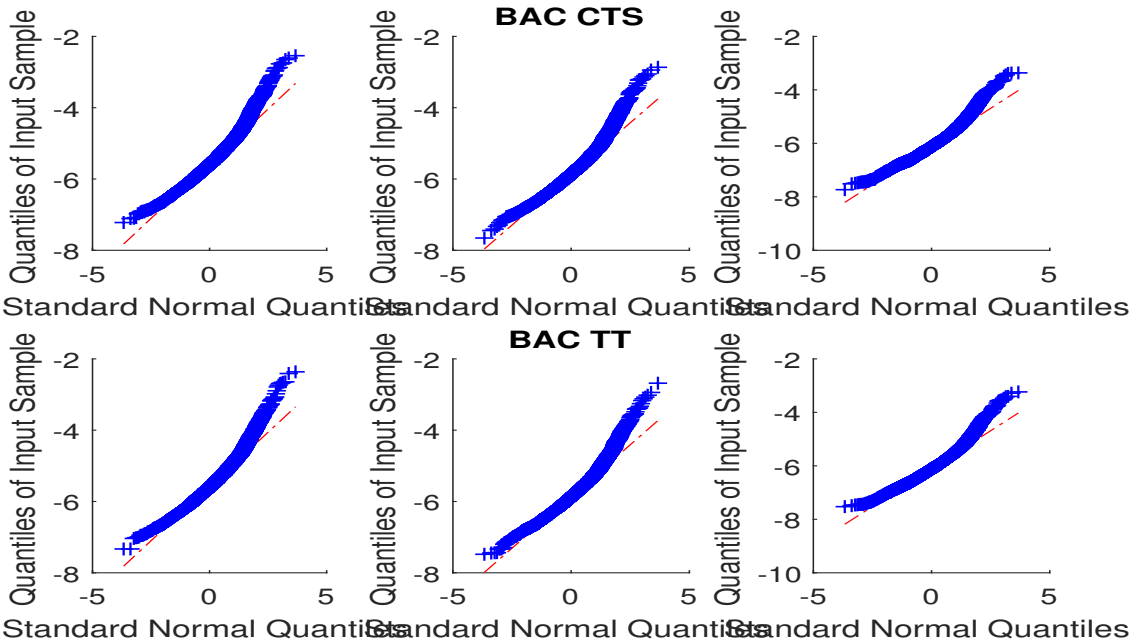

Figure C.8: QQ-plots of the log-transformation of the negative estimates of daily quantiles, $\widehat{q}_{p,d}$ of BAC, computed based on the daily estimates of the quantiles as described in the main text from scaling up estimates of intraday quantiles estimated from data sampled in TTS and CTS (use the DMA estimator of Hurst coefficient and the frequency of 5 minute to sample the data). The first column corresponds to $p = 1\%$, the second column corresponds to $p = 2.5\%$ and the third column corresponds to $p = 5\%$. The red line corresponds to the standard normal distribution.