Predicting Intraday Price Distributions at High Frequencies

Matti Antola
Valo Research and Trading

Tommi A. Vuorenmaa*
Valo Research and Trading

(July 15, 2013)
Revised: August 9, 2013

Abstract

We propose extensions to the continuous-time random walk (CTRW) framework so far mainly developed within the econophysics community. Using numerical methods, we extend the CTRW framework with more general marginal distributions than previously proposed. There are two main findings in this respect: First, modeling the returns with a Student-t distribution gives at least as good price distribution predictions as the other previously proposed distributions in this framework. Second, a mixed-Weibull waiting time distribution fits exceptionally well to our three-week long Nasdaq OMX data. When combined with standard financial econometric GARCH and ACD models and intraday seasonality filtering procedures new to the CTRW framework, our models deliver realistic intraday Value-at-Risk and Expected Shortfall predictions. Compared to the basic CTRW model, the total average performance improvement is typically around 70 percent. The effect of filtering out memory in waiting times is particularly noticeable: around 50 percent of the total performance improvement. Overall, our extensions and filtering methods make the CTRW framework a useful tool for intraday risk management (trading), especially at high frequencies of less than a minute or so.

Introduction

Modeling intraday dynamics of financial prices, and stock market prices in particular, is known to be extremely challenging. Their random character at daily (or larger) time-scales is further complicated by market microstructure related issues at the very shortest intraday time-scales. The consensus is that prediction of future price movements is nearly impossible – markets are thought to be

*Address correspondence to: tommi.vuorenmaa@valotrading.com
Postal address: Valo Research and Trading, Mannerheimintie 2 (Floor 6), FI-00100 Helsinki, Finland. The usual disclaimer applies.
We thank Aleksi Pitkäjärvi and Nima Reyhani at Valo Research and Trading for comments.

efficient in this sense. Thus, the best one might hope for is to predict the most likely price movements. Technically, one would like to predict the price distribution for arbitrary time-scales. That would be of interest not only from a trading perspective but, perhaps more importantly, also from a risk management perspective. To estimate the future price distribution with reasonable statistical precision one needs to assume that certain statistical characteristics are stable over time. If we are able to make realistic assumptions and can still solve the intraday price distribution model somehow, then various risk measures, such as Value-at-Risk (VaR) and Expected Shortfall (ES), would ideally be easily calculatable for any chosen intraday time-scale to control risk in real time.

Numerous smaller, but closely connected, challenges appear when trying to forecast intraday price distributions. A special characteristic of high frequency data is that trades and quotes are irregularly spaced in time: information arrives in bursts and market players react to new information with different speeds. For example, after a significant news, there is commonly a period of higher price turbulence where market players with different risk appetites battle against each other to finally agree on a new fair price level. Thus, periods of short waiting times can be expected to be correlated with information arrival [see, e.g., Easley and O'Hara (1992)]. It is then also quite natural to expect that return volatility would be clustered at intraday time-scales [see, e.g., Engle and Russell (1998)], implying that both waiting times between trades (or quotes) and volatility would display some sort of memory structure. To complicate matters, the turbulent time periods near market openings and closings are typically found to be distinctively different from the other more tranquil intraday time periods. These challenges, and more, await the modeler of intraday price distributions.

In this paper, we study the predictability of future intraday price distributions using the so-called continuous-time random walk (CTRW) framework, which explicitly allows for a simultaneous modeling of returns and waiting times. The CTRW framework was initially proposed by Montroll and Weiss (1965), and later introduced in the econophysics literature by Scalas, Gorenflo, and Mainardi (2000). A critical assumption of the CTRW framework is that both returns and waiting times are independent and identically distributed (i.i.d.) random variables. This assumption is well known to be unrealistic: for example, the idea behind all GARCH models in financial econometrics is to model return volatility as a time-dependent process [see, e.g., Bollerslev (1987)]. Analogously, ACD models assume that waiting times are time-dependent [see Engle and Russell (1998)]. Because the CTRW framework does not account for time-dependency, but does account for asynchronous waiting times, it allows us to study the effect of memory with regards to price distribution prediction.

The strength of the CTRW framework is that it allows for several quantities to be explicitly calculated. For example, price volatility can be shown to depend only on the waiting time distribution and the second moment of the return distribution [see Masoliver, Montero, and Weiss (2003)]. Another important factor is that the CTRW framework easily provides intraday price distribution estimates for any time-scale in wall-clock time, while taking advantage of all the available high-frequency trade time data.

To make the CTRW framework more realistic, we propose new extensions using a numerical approach. More explicitly, in this paper, we extend the basic CTRW framework by first assuming that the return distribution is better described by a Student-t distribution from the family of generalized hyperbolic distributions. The previous CTRW literature has focused on much simpler distributions, such as normal or double-exponential distributions. Second, we allow more flexibility in the waiting times by using a novel mixed-Weibull distribution. We then incorporate the well known "inverse-U" intraday seasonality in waiting times. Finally, we document the effect of data filtration with GARCH and ACD models. Thus, we effectively study the significance of non-Markovian waiting times and returns. Our empirical VaR and ES results show that the extensions we propose are indeed valuable and provide direction for further research.

This paper is constructed as follows. In Section 1, we describe the model and the marginal distributions we propose to use. In Section 2, we describe the data used in the empirical analysis. Section 3 collects the basic empirical results. In Section 4, we study the effect of intraday seasonality and temporal dependence filtering on VaR and ES risk measures. Section 5 concludes. We include technical material related to the numerical method in the appendix.

1 Model

1.1 General features

We first shortly describe the CTRW framework. For a more detailed introduction using a slightly different notation, see Masoliver et al. (2006).

We define the (logarithmic) price process as $X(t) = \log P(t)$, where $P(t)$ is the last traded price of a financial instrument which, in our case, is the stock price.[1] The times at which trades occur is a random variable, T_i. To simplify notation, the corresponding trade price, $X(T_i)$, is set to X_i. From these variables we calculate the (logarithmic) returns, $\Delta X_i = X_i - X_{i-1}$, and waiting times, $\Delta T_i = T_i - T_{i-1}$. In the standard CTRW model, returns and waiting times are assumed to be uncorrelated over time and with each other. Then the price distribution is defined by the marginal distributions of returns and waiting times.[2] Since it is assumed that $X(0) = 0$, we predict the future price distribution at the moment a trade takes place. Thus, the idea is to sample the cumulative return series in "trade time" – that is, using a clock that does not progress homogenously in "wall-clock" time.[3] Trade time, and related concepts,

[1] We subtract the price at $t = 0$ so that $X(0) = 0$. We could also use other definitions of returns, based on for example the mid-quote, or use altogether other sampling clocks, such as "lost time" [see Vuorenmaa (2012)]. We could also condition on other factors, such as the minimum cumulative volume. These additional factors could affect the predicted price distribution. In the current context, however, we find our definition to be natural.

[2] If the waiting times are not exponentially distributed, the CTRW model is not Markovian. Then the remaining waiting time depends on how much time has passed since the previous jump or, in more technical terms, the hazard function of waiting times is non-constant.

[3] In detail, we calculate the cumulative returns as follows: (1) Let (X_i, T_i) be the price and time pair for trade i. We start at $i = 1$ and find the largest index j satisfying $T_j < T_i + t$ (where

can be argued to be more natural for automated trading, such as high-frequency trading (HFT) [see, e.g., Easley, López de Prado, and O'Hara (2012)].[4] The predicted price distribution, on the other hand, is estimated in the humanly more natural wall-clock time. We next describe the relationship between the predicted price distribution and marginal distributions in more detail.

The two main ingredients of the CTRW framework, namely the return and waiting time distributions, are defined by the joint probability density function, $h(x,t)$,

$$h(x,t)dxdt = \mathbb{P}\left[x < \Delta X_i < x + dx;\ t < \Delta T_i < t + dt\right],$$

where the index i serves to remind that returns and waiting times could, in principle, be correlated. As we discuss later, though, here we assume independence on empirical grounds.[5] From the above joint probability distribution, the return and waiting time marginal densities, $f(x)$ and $g(t)$, respectively, are simple to derive as

$$\begin{aligned} f(x) &= \int_0^\infty h(x,t)dt = \mathbb{P}\left[x < \Delta X < x + dx\right], \\ g(t) &= \int_{-\infty}^\infty h(x,t)dx = \mathbb{P}\left[t < \Delta T < t + dt\right]. \end{aligned}$$

The crux of the CTRW framework is the time-dependent price distribution prediction, $p(x,t)$, defined as

$$p(x,t)dx = \mathbb{P}[x < X(t) < x + dx;\ p(x,0) = \delta(x)],$$

where $\delta(x)$ is the Dirac delta function satisfying $\delta(x) = 0$ for $x \neq 0$ and $\int_{-\infty}^{\infty} \delta(x)r(x)dx = r(0)$ for any integrable function $r(x)$. The knowledge of $p(x,t)$ allows to pin down the probability of the price to be in a certain range in the future. While this distribution describes the (macroscopic) price behavior, it is connected to the return and waiting time distributions through

$$p(x,t) = \delta(x)S(t) + \int_0^t \int_{-\infty}^\infty h(x - x', t - t')p(x', t')dx'dt', \tag{1}$$

where the survival function, $S(t) = \int_t^\infty g(t')dt'$, gives the probability for the price not changing for a priori known time t.[6] Notice that, by construction,

t is the time-scale). (2) Record $X_j - X_i$ as the resulting cumulative return. (3) Increase i by one step and go back to step 1. (4) End when $T_i + t > T_N$, where N is the last observation.

[4] Naturally, we could as well apply a more refined clock in sampling, for example the "lost time" clock suggested in Vuorenmaa (2012). This would have the additional benefit of automatically producing non-zero return series and deleting the first-lag autocorrelation in returns that may affect the returns filtration. Our preliminary work suggests that the results are largely unchanged compared to the more commonly applied trade time clock.

[5] The case of non-independent return and waiting time distributions is currently under study by the authors.

[6] The first term accounts for the possibility that the price has not yet moved, while the second term accounts for the possibility that the price is at location x' just before time t and then jumps to x at t.

$p(0, t)$ is affected by the Dirac delta function, whose contribution we disregard in the figures in the empirical section for simplicity.[7]

To make Eq. (1) useful, we solve it for p. Taking the Fourier and Laplace transforms with respect to x and t yields the general solution,

$$\begin{aligned} p(k,s) &= S(s)\frac{1}{1-h(k,s)} \\ &= \frac{1-g(s)}{s}\frac{1}{1-h(k,s)}, \end{aligned} \tag{2}$$

where k and s correspond to x and t in the frequency domain [see Masoliver et al. (2006)].[8] Further, assuming that the returns and waiting times are uncorrelated with each other gives the simple factorization, $h(k, s) = f(k)g(s)$, thus providing a link between the price and marginal distributions. To find $p(k, s)$ in the time domain, where we need it in practice, we calculate the Fourier-Laplace inverse of Eq. (2). In this paper, we study realistic choices of marginal distributions, for which $p(k, s)$ cannot be analytically inverted. Thus, we need to resort to numerical methods discussed in Appendix A.2.

1.2 Marginal distributions

Two marginal distributions are relevant in the CTRW framework: the return distribution, with density f, and the waiting time distribution, with density g. We next describe the choices we make for both of these distributions, and expand the universe of distributions previously studied in the CTRW framework using numerical methods.

1.2.1 Returns

In the finance literature, the shape of a proper return distribution has been tested at least since the 1960s when the normal distribution was first documented to produce unrealistically low tail probabilities [see, e.g., Mandelbrot (1963)]. Consequently, the so-called alpha-stable distributions were proposed as stock return models. These distributions do not (except in the normal distribution case) have a well-defined second moment and are thus theoretically difficult to maneuver, nor have they performed that well in empirical tests. Similarly, the main reason for the shortage of good return distribution candidates in the CTRW framework seems to be that the price distribution is not analytically tractable with most realistic return distributions. There are two benchmark return distributions that do allow for an analytic solution: normal distribution,

[7] Since we exclude zero returns in our empirical analysis, the survival function of the waiting time distribution gives the probability a price has not changed up to a specific time point t. Thus, for small enough time-scales, the CTRW model predicts a peak in the price distribution at $x = 0$. In reality, prices can either stay unchanged for an arbitrary time or revert back to its start level within that same time-scale due to price discreteness.

[8] We use k and s in the frequency domain, and x and t in the time domain.

with density $f_n(x) = e^{-x^2/2\sigma^2}/\sigma\sqrt{2\pi}$, and double-exponential distribution, with density $f_{de}(x) = e^{-|x|/\gamma}/2\gamma$.

In the CTRW literature, Masoliver et al. (2006) report a general result showing that at intermediate times, $t \approx \mathbb{E}[\Delta T]$, the tail of $p(x,t)$ is given by

$$p(x,t) \sim \frac{t}{\mathbb{E}[\Delta T]} f(x),$$

which motivates the use of heavier tail distributions than exponentially decaying tails. Masoliver, Montero and Weiss (2003) and Masoliver et al. (2006) suggest to use a heavier tail return distribution of the form[9]

$$f_{pl}(x) = \frac{\beta - 1}{2\gamma\beta\left(1 + \frac{|x|}{\gamma\beta}\right)^{\beta}}. \tag{3}$$

However, as we find this distribution to converge to the double-exponential distribution when $\beta \to \infty$, and the double-exponential fits reasonably well to our data, this power law distribution provides little extra value to us.

Outside the CTRW literature, a good candidate for returns (with heavier than normal tails) is provided by the generalized hyperbolic (GH) distribution family [see Barndorff-Nielsen (1977)]. The GH distributions have become quite popular in the financial econometrics literature during the last ten years or so due to their good performance over many fixed (even intraday) time-scales [see, e.g., Raible (2000)]. In its general form, the density of a symmetric GH return distribution is defined by [see, e.g., Eberlein and von Hammerstein (2002)]

$$f_{gh}(x) = \frac{\alpha^{\lambda}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^{\lambda}K_{\lambda}(\delta\alpha)} \left(\delta^2 + x^2\right)^{\frac{2\lambda-1}{4}} K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2 + x^2}\right),$$

where K_ν is the modified Bessel function of the second kind. In this paper, rather than applying the whole GH family, we concentrate on a special case: the (non-standard) zero-mean Student-t distribution with ν degrees of freedom,[10]

$$f_{st}(x) = \frac{1}{\sqrt{\nu}\sigma B\left(\frac{1}{2}, \frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}}. \tag{4}$$

The Student-t distribution is easier to estimate than the GH distribution family, and it has been successfully used in the context of many different financial time series and econometric models, for example GARCH [see, e.g., Bollerslev (1987)].

1.2.2 Waiting times

The modeling of waiting times – or durations, as they are often called in financial econometrics – is a much newer topic than the modeling of returns. This is

[9] In contrast to Masoliver, Montero and Weiss (2003), we scale γ with β.

[10] Student-t distribution is recovered from the GH distribution when $\lambda = -\nu/2$, $\delta = \sigma\sqrt{\nu}$, and the other parameters are zero.

mainly because accurate intraday data started to be available only in the 1990s. Engle and Russell (1998) were among the first to model the intertemporally correlated event arrival times. For example, the waiting times between trades are not homogenously spaced in time, but vary in random fashion. Interestingly, there is a special memory structure in that randomness: short waiting times tend to be followed by short waiting times and long waiting times tend to be followed by long waiting times. In other words, there is a tendency for waiting times to cluster similarly as returns do. The problem is, as explained in Section 1, that the standard CTRW framework assumes i.i.d. waiting times and thus does not account for any such clustering effects in a natural way although it allows for asynchronicity automatically.

In the CTRW literature, and more generally in the waiting time literature, the benchmark distribution is the exponential distribution, $g(t) = e^{-t/\mu}/\mu$. Because several empirical studies, both in financial econometrics and econophysics, report a bad fit to real data, other more realistic waiting times candidates have been proposed and studied in the 2000s. In the CTRW literature, the most well-known candidates include the Mittag-Leffler function [see Scalas, Gorenflo, and Mainardi (2000), and Mainardi et al. (2000)], certain power law distributions [see Masoliver, Montero, and Weiss (2003) and Masoliver et al. (2006)], scaled exponentials, and Weibull distributions [see Jiang, Chen, Zhou (2008)].

In this paper, we consider the Weibull distribution as our main candidate for modeling waiting times. Its density and survival functions are known to be

$$g_w(t) = \frac{\alpha}{\lambda}\left(\frac{t}{\lambda}\right)^{\alpha-1} e^{-\left(\frac{t}{\lambda}\right)^{\alpha}} \quad \text{and} \quad S_w(t) = e^{-\left(\frac{t}{\lambda}\right)^{\alpha}}. \tag{5}$$

The increased presence of trading algorithms dominate the very highest frequencies in modern markets [see, e.g., Vuorenmaa (2013)]. We accommodate for this by assuming two regimes of waiting times. The first, shorter than approximately one second ($= t_0$), is characterized by high-frequency and algorithmic trading, while the second regime is characterized by significantly slower (electronic or human) trading.[11] Thus, we propose to use a mixture of Weibull distributions:

$$g_{mw}(t) = p g_{w_1}(t) + (1-p) g_{w_2}(t), \tag{6}$$

where g_{w_i} are separate Weibull distributions with independent parameters approximately fitting the regions $\Delta T < t_0$ and $\Delta T > t_0$ respectively. Here $p \sim 0.3$ corresponds to the proportion of HFT related trades. To the best of our knowledge, we are the first to propose this in the CTRW literature.

1.2.3 Estimation method

The estimation of distribution parameters is commonly done by maximum likelihood or the method of moments. In this paper, we estimate the waiting time

[11] High-frequency trading (HFT) and algorithmic trading should not be used interchangeably. As explained in Vuorenmaa (2013), both types of fast electronic trading are subsets of automated trading, but they serve a different purpose: while HFT is mostly about market making and statistical arbitrage, algorithmic trading attempts to minimize trading costs of (mainly large institutional) buy and sell orders. Both types of trading affect waiting times.

and return distribution parameters in two steps. First, we require the mean waiting time and return variance to match their theoretical counterparts. This requirement is based on an analytic result that in the limit ($t \to \infty$, $k \to 0$), if the relevant moments exist, the CTRW price distribution converges to a Gaussian, irrespectively of the marginal distributions [see, e.g., Masoliver et al. (2006)]:[12]

$$p(k,t) \sim \exp\left\{-\frac{\mathbb{E}\left[\Delta X^2\right] tk^2}{2\mathbb{E}\left[\Delta T\right]}\right\}. \tag{7}$$

That is, we estimate the marginal distributions so that the resulting Gaussian price distribution matches the theoretical result, Eq. (7). Because the exponential, normal (with mean zero), and symmetric double-exponential distributions have only one free parameter, they now become fixed.

For the other more complicated marginal distributions with more parameters, we find an analytic relation for the parameters to enforce the moments. More precisely, for the distribution of Eq. (3) and the Student-t of Eq. (4), we use, respectively,

$$\gamma = \sqrt{\frac{\mathbb{E}\left[\Delta X^2\right]\Gamma(\beta)}{2\beta^2(\beta-1)\Gamma(\beta-3)}} \quad \text{and} \quad \nu = \frac{2\mathbb{E}\left[\Delta X^2\right]}{\mathbb{E}\left[\Delta X^2\right]-\sigma^2}.$$

For the Weibull of Eq. (5) and mixed-Weibull of Eq. (6), we use, respectively,

$$\lambda = \frac{\mathbb{E}\left[\Delta T\right]}{\Gamma(1+1/\alpha)} \quad \text{and} \quad \lambda_2 = \frac{\mathbb{E}\left[\Delta T\right]-\lambda_1 p\Gamma(1+1/\alpha_1)}{(1-p)\Gamma(1+1/\alpha_2)}.$$

When β is large, the distribution of Eq. (3) converges to the double-exponential distribution. As this is empirically validated by the authors (but unreported), in the empirical section we only discuss the double-exponential results. Notice also that as $\nu \to \infty$, the Student-t distribution converges to the normal distribution.

In the second step of estimation, for the distributions with remaining free parameters, we use a non-linear least squares method to estimate the parameters by matching the theoretical cumulative distribution function with the empirical one. This method is better known as the method of minimum distance. More specifically, we minimize $\sum_i [F(x_i)-\hat{F}_n(x_i)]^2$ with respect to all free parameters, where $F(x)$ and $\hat{F}_n(x)$ are the theoretical and empirical distribution functions. For the mixed-Weibull distribution, we initialize the minimization algorithm on parameter estimates of two Weibull distributions satisfying $\Delta T < 1$ (in seconds) and $\Delta T > 1$, respectively. For the initial value of p, we use the fraction of waiting times with $\Delta T < 1$. For the other distributions, the initial parameter values are inconsequential in our experience. In the empirical results, we only report the standard errors for the variables estimated by this method.

The above estimation method requires an analytic form for the cumulative distribution function, which poses a problem for the family of GH distributions generally. For this reason, we resort to using only a special case of GH, the Student-t distribution, and defer the more general solution to another paper.

[12] That the price distribution is approximately Gaussian at long enough time-scales is numerically validated by the authors.

2 Data

We use three weeks of Nasdaq OMX trade price data from March 2013 for the following six stocks (tickers in parentheses): Apple (AAPL), Boeing (BA), Chevron (CVX), Google (GOOG), IBM (IBM), and 3M (MMM). These six stocks are chosen based on their relatively large prices compared to the other Dow Jones Industrial Average index stocks. The minimum tick size is one cent for all of our sample stocks and the trades are time-stamped to the precision of nanoseconds. We use GOOG as our primary example, for which the price hovers around \$800 with a typical bid-ask spread of about \$0.25 (see Figure 1).

We use the following data filtering mechanism. We concentrate on trades taking place within the continuous trading hours, that is, from 9:30am to 4:00pm (EST). We decide to remove the first and last fifteen minutes of the continuous trading session to account for these highly turbulent intraday periods. The opening period is quite standardly removed, especially in the analysis of waiting times [see, e.g., Engle and Russell (1998)], while the closing period is removed here as a precaution to decrease the effect of intraday seasonality. We exclude zero returns and merge trades taking place at the same nanosecond. Merging of trades with the same time-stamp is standard practice in the analysis of waiting times [see, e.g., Vuorenmaa (2009)]. Zero returns are excluded as they would create problems for the continuous distributions in the CTRW framework, and because they can be expected not to carry significant information of the stock.[13] The total number of removed and saved trades for GOOG are reported in Table 1. Almost half of the returns are excluded from our analysis as zero returns. The data filtering mechanism is described in more detail in Appendix A.1.

The CTRW model, in its standard form (as described in Section 1), relies on continuous distributions. In reality, however, stock prices live on a discrete grid. We consider only stocks with a relatively small tick size. Table 2 shows that most of the changes for GOOG are larger than eight ticks: 43, 39, and 47 percent for Weeks 10, 11, and 12, respectively. The minimum tick size does not appear to be restrictive for GOOG. For the cheaper stocks, however, this is not as clear. For example, MMM, with an average price of \$105, has a significant portion (66 percent) of price changes with one tick apart. Similarly, BA, an even cheaper stock with an average price of \$83, but with significantly higher number of trades, may be restricted by the tick size as well. Finally, notice that while GOOG has an average number of observations among the six stock we analyze (around 5 000 trades per day), AAPL has around four times more, making it a potentially lucrative target for HFT. Thus, for the robustness of our CTRW price distribution predictions, AAPL is another noteworthy case.

[13] Here we depart from the analysis of Dionne et al. (2009), for example, who argue otherwise. We agree with them in that the efficient market hypothesis implies that stock prices change when new information arrives in the market. But when prices are stale, no new information has entered the market and the previous stock price reflects the current information.

3 Empirical analysis

We use the following shorthand notation throughout the empirical section: normal (N), double-exponential (DE), Student-t (ST), exponential (E), Weibull (W), and mixed-Weibull (MW). We first analyze our primary example, GOOG, for Weeks 10, 11, and 12, separately. We then compare these results to the other stocks for the full three-week period. Later, in Section 4, we measure the CTRW models' performance using the 99 percent VaR and ES statistics.

3.1 Return distribution

Descriptive statistics for the pre-filtered returns are reported in Table 3 (Panel 1). We observe that the standard deviation of returns for GOOG, Week 12, is noticeably larger than for Weeks 10 or 11. Week 12 also experiences the largest negative return between consecutive trades (0.13 percent). Otherwise, the three weeks appear similar. The other five stocks we analyze do not differ significantly in these respects from GOOG. The most volatile stocks are BA and MMM. All six return series are unconditionally non-Gaussian (test statistics not reported).

The unconditional parameter estimates for GOOG Weeks 10, 11, and 12 are reported separately in Table 4 (Panel 1). The parameter estimates differ quite significantly between the three weeks, but stay within a reasonable range. The most noteworthy result of this table is that the Student-t distribution converges to the normal distribution for CVX for the full three-week long data (Panel 2). As we discuss later in more detail, a more advanced data filtering procedure, which accounts for intraday seasonality and temporal dependency in the second moment, makes the data more Gaussian. After such filtering, we find that the Student-t distribution converges to the normal distribution in four out of six cases (Panel 3). The two cases for which no convergency is observed, GOOG and AAPL, suggest that fat-tailed distributions are sometimes necessary.

We collect the root-mean-squared deviation (RMSD) statistics in Table 5. Panel 1 confirms that for GOOG the normal distribution has the largest deviation from the empirical distribution. For the other stocks (Panel 2), the Student-t distribution outperforms (or fits equally well) both the normal and double-exponential distributions. For AAPL and GOOG, the normal distribution does the worst, while for BA, CVX, and MMM, the normal distribution outperforms the double-exponential by a small margin. Excluding zero returns should act in favor of the normal distribution. After more advanced filtering procedures, the double-exponential distribution performs even worse (Panel 3).

Figure 2 illustrates the empirical return distributions for GOOG for three separate weeks. It descriptively confirms the above results. The normal distribution fits the body of the distribution quite well, but it consistently underestimates in the tails. The Student-t and double-exponential distributions do much better in the tails, the former slighly overestimating in the far end.

Naturally, parameter estimation is hindered by our relatively small sample size. The fact that the Student-t distribution converges to the normal distribution in many cases supports this conjecture: without sufficient amount of data in the tails, higher degree of flexibility is not well utilized. The parameter estimates depend on what methodology is applied. In principle, we could weight different regions of the empirical distribution according to our risk preference and more clearly separate different distribution fits from each other. It is not, however, our goal in this paper to make definitive conclusions about the performance of distributions behaving similarly in the tails. In the current context, it is more important for us to study how the predicted price distribution is affected by the choice of reasonable marginal distributions at different time-scales.

3.2 Waiting time distribution

Table 3 (Panel 2) collects the descriptive statistics for waiting times. A couple of numbers deserve a special mention. For GOOG, we find the mean waiting time for Week 12 (11 sec) to be around twice as large as for Week 10 (6.1 sec). This corresponds to the higher reported number of trades for Week 10 (see Table 1). For the most actively traded stock, AAPL, the mean waiting time over the three-week period is only 2.4 sec, while for the least actively traded stock, MMM, the mean waiting time is 24 sec . However, the minimum waiting time is similar for all the six stocks, being the smallest for GOOG ($0.92\mu s$) and largest for BA ($6.3\mu s$). Notice that zero waiting times do not exist because we have merged together trades with the same time-stamp (as explained in Section 2). There are significant waiting time gaps for all the six stocks. The largest waiting time is for MMM (440 sec), indicating that a fat-tailed distribution may sometimes be necessary in modeling the waiting times. These results are as expected.

Table 4 also includes the different waiting time distribution parameter estimates. For GOOG, these estimates can be observed to fluctuate quite significantly between the three weeks we analyze. The estimates stay within reasonable limits for all the six stocks, however. From the perspective of predicting the price distribution at high frequencies, it is worthwhile to pay attention to the mixed-Weibull model. As described above (in Section 1.2.2), the mixed-Weibull distribution has two different time regimes: one for short times (less than approx. 1 sec) and another for longer times. For GOOG and AAPL, we find the estimated probability of a short waiting time, $\widehat{p}$, to be around 30 percent. For the less actively traded stocks, this probability is much lower: for MMM (the least actively traded stock), it is only around 15 percent. AAPL and GOOG also have noticeably smaller scale parameter estimates for the first regime, $\widehat{\lambda}_1$, and AAPL also for the second regime, $\widehat{\lambda}_2$. Finally, we find the estimates for the two shape parameters, $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$, to be rather stable and significantly smaller than one (the theoretical value for the exponential distribution). The Weibull and mixed-Weibull shape parameter estimates also imply that the hazard function is not constant over time, again contradicting the exponential distribution.[14]

The RMSD results in Table 5 confirm the above reasoning: for GOOG (Panel 1), the weekly statistics document clear evidence against the exponential distri-

[14] We do not explicitly consider the possibility of non-monotonic hazard functions here [see, e.g., Vuorenmaa (2009)] although the mixed-Weibull distribution accommodates for it.

bution. Panel 2 confirms the results for the other five stocks. All these results rank the Weibull distribution as the second best option with about halved deviation from the empirical distribution. Most interestingly, we consistently find the smallest RMSD for the mixed-Weibull distribution. The RMSD is around ten times, or sometimes even more, smaller than for the Weibull distribution. This gives some concrete indication of how well the the idea of two regimes works. It is worth noticing that, similarly as for returns, more advanced filtering techniques increase the performance of the standardly applied exponential distribution. However, this performance increase is not substantial when compared to the performance of the Weibull and mixed-Weibull distributions.

Figure 3 illustrates the empirical waiting time distributions for GOOG. This figure confirms descriptively that the exponential distribution tends to overestimate at the smallest times and underestimate at the largest times. The shape of the Weibull distribution is better, but it tends to overestimate in the tails. The mixed-Weibull distribution produces the best fit. The survival functions are plotted in Figure 4, which show three different 'knees' at approximately 10^{-4} sec, 10^{-3} sec, and 1 sec. This also motivates the use of two different regimes in the mixed-Weibull distribution. While the Weibull distribution is unable to fit these knees, the mixed-Weibull is able to do so, except perhaps for the first knee where we observe a small departure from the empirical survival function. Based on these results, we subsequently focus on the Weibull and mixed-Weibull results.

3.3 Price distribution

As mathematically formulated in Section 1, the CTRW framework attempts to make predictions for the price distribution utilizing both trade and wall-clock time. We now compare the different CTRW model predictions to the observed empirical price development using time-scales of reasonable widths. Since we are mostly interested in the high-frequency domain, we focus on price distribution predictions over a relatively small time-scale, which we here set to 10 sec.

We illustrate the difference between the CTRW model price distribution prediction and the empirical price distribution with a short (10 sec) and a long (20 min) time-scale. In the figures, we show four different marginal distribution combinations based on the results of the previous section. For comparison, we include exponential waiting times with normally distributed returns as our fifth model, although neither marginal distribution was found to be very realistic above. As an example, Figure 5 shows the prediction for GOOG for the three separate weeks with a time-scale of 10 sec. As expected for such a small time-scale, the exponential-normal model seriously underestimates the tails. On the other hand, all four of our proposed models perform quite well, especially for Weeks 10 and 12. The Weibull-Student-t model performs the best in the tail regions, which is also expected since the Student-t distribution accommodates for fatter tails than the double-exponential distribution. For Week 12, however, we find a clear asymmetry. For that particular week, the right-hand tail happens to be much thinner than the left-hand tail.[15] For Week 11, the best fitting model underestimates the negative tail density by a factor of ten or so. In the figures, the yellow vertical lines correspond to the empirical 99 percent VaR on both the negative and positive realm. The problem in the negative tail appears only in the region exceeding that VaR value. The associated underestimation of risk can be expected to be better captured by the 99 percent ES rather than the same confidence level VaR since ES describes the tail risk more accurately. We use both commonly applied risk metrics in the next section for diagnosing problems.

The price distribution prediction results with a larger time-scale of 20 min for GOOG are illustrated in Figure 6. For all the three weeks separately, the five models are tightly jammed together. This is an expected consequence of the asymptotic result towards Gaussianity we discussed in Section 1. In practice, this means that it does not matter what ingredients we put in the CTRW model, the predicted price distribution is much the same once the time-scale is large enough. In reality, though, we find more density mass near the zero returns than expected by the CTRW framework. Another region where the models fail is the tail. More precisely, for Week 10, the tails for GOOG are slightly underestimated, while for Week 11 the difference is more significant. For Week 12, the right-hand tail is overestimated. The first percentiles (the vertical yellow lines) indicates that most of the problems appear near or beyond this limit.

VaR and ES metrics are commonly applied in finance, but typically require a relatively large number of observations. Thus, we do not expect the above results to be reliable enough for decision making. In what follows, we extend the CTRW framework in other dimensions than just fine-tuning the marginal distributions. We then quantify the performance using the VaR and ES metrics with the full three-week long data for all the six stocks. We believe this is a practically meaningful way to quantify how good the CTRW models really are.

4 Other extensions

We next study whether the detected points of failure of the CTRW framework, particularly the sometimes poor fit in the tails of the price distribution, could be accounted for. We posit two hypotheses of the likely cause, neither of which have been considered before in this framework to the best of our knowledge. We first consider the effect of intraday seasonality and derive a prediction formula for it. We then study the effect of temporal dependence in waiting times and returns by filtering them out by using standard financial econometrics methods.

[15] At larger time-scales (one day or larger), this phenomena is commonly known as the "leverage effect" and says that negative returns tend to be of larger magnitude than positive returns. We have not attempted to incorporate this feature into the CTRW framework, because it may be of less importance at high frequencies on which we concentrate.

4.1 Deseasonalization

Average waiting times in stock markets typically vary significantly over a trading day in a predictable manner: short waiting times are expected near the opening and closing of the day while long waiting times are expected outside these turbulent time periods [see, e.g., Vuorenmaa (2009)]. Thus, the expected value of waiting times, $\mathbb{E}[\Delta T]$, has an "inverse-U" intraday pattern, while for volatility it is the other way around.

In the CTRW framework, the diurnal variation can be taken into account by conditioning the relevant functions on time of the day, τ :

$$f(x,\tau), \quad g(t,\tau), \quad h(x,t,\tau), \quad \text{and} \quad p(x,t,\tau).$$

Usually, in the econometric literature, the diurnal pattern is estimated using cubic or piecewise linear splines, kernel methods, dummy variables, or linear regression on time. Here we simply partition the day into different intervals and divide out the seasonal variation:

$$\Delta\hat{T}_i = \Delta T_i \frac{\mathbb{E}[\Delta T]}{\mathbb{E}[\Delta T]_{j(i)}}, \tag{8}$$

where $\mathbb{E}[\Delta T]_{j(i)}$ is the average duration in period j containing observation i. Table 6 reports $\mathbb{E}[\Delta T]_{j(i)}$ for the six stocks we analyze. The first and last fifteen minutes produce the lowest mean waiting times, in line with our expectations. We also confirm the inverse-U pattern, with a maximum at 12:45pm–1:45pm.

We fit the waiting time distribution, $\hat{g}(t)$, to the deseasonalized data. Time dependence can be reinstated in the waiting time distribution by

$$g(t,\tau) = \frac{\mathbb{E}[\Delta T]}{\mathbb{E}[\Delta T]_{k(\tau)}} \hat{g}\left(t\frac{\mathbb{E}[\Delta T]}{\mathbb{E}[\Delta T]_{k(\tau)}}\right),$$

where $k(\tau)$ denotes period k containing the time of the day, τ. We may also reinstate the seasonality in the time-dependent cumulative returns distribution. Assuming $\hat{p}(x,t)$ is the price distribution using the deseasonalized waiting time distribution $\hat{g}(t)$,

$$p(x,t,\tau) = \hat{p}\left(x, t\frac{\mathbb{E}[\Delta T]}{\mathbb{E}[\Delta T]_{k(\tau)}}\right). \tag{9}$$

The effect is demonstrated in Figure 7, showing that the predicted price distribution is wider for the morning hours than mid-day. The effect of seasonality also becomes stronger at larger time-scales (here 20 min). Thus, intraday seasonality can be expected to have a significant impact on the VaR and ES statistics, as we verify later. We then account for intraday seasonality using Eq. (8) rather than Eq. (9). This is simply because we analyze the data ex post. In real-time risk management or trading, the latter approach would be preferred.

As already described in Section 2, our first choice of dealing with seasonality is to remove the first and last fifteen minutes before the application of the CTRW model. This is partly for convenience. When compared to the seasonality adjustment method outlined just above, we find both ways to improve the results, although data removal is a safer choice. This may be due to two reasons: First, there are simply not enough data to find a stable relationship between time and the different variables. Second, there are more important factors to consider in the CTRW framework than seasonality. In the light of empirical evidence, we believe that the latter reason rings truer. It appears that the first and last fifteen minutes of trading consist of rapid price movements that clearly violate the inherent CTRW model assumption of i.i.d. returns and waiting times. In the last section, we study the effect of these time periods, too. Before that, however, we study how important the i.i.d. assumption is over the calmer trading period, excluding the problematic opening and closing periods.

4.2 GARCH/ACD-filtering

The CTRW framework does not allow for any temporal dependency in waiting times or returns. However, stock market data exhibit both types of dependency, strong for short lags and significant even at much longer lags. In more technical terms, the sample autocorrelation function (ACF) decays as a power law. This phenomenon is termed long-range dependency [see, e.g., Beran (1992)]. The existence of significant autocorrelation in the second moment of returns, known as volatility clustering, is regarded as one of the key stylized facts of stock market returns. To improve performance, we attempt to filter out temporal dependency from returns and waiting times before the estimation of CTRW.

The modeling of returns is standardly done as follows. The simplest form of the generalized autoregressive conditional heteroskedastic (GARCH) model [see Bollerslev (1986)], with one lag, zero mean, and i.i.d. normal errors, is

$$\Delta X_i = v_i^{1/2}\epsilon_i,$$

$$v_i = \omega + \alpha \Delta X_{i-1}^2 + \beta v_{i-1},$$

where the constraints $\omega > 0$, $\alpha \geq 0$, and $\beta \geq 0$ ensure the conditional variance is non-negative, and $\alpha + \beta < 1$ guarantees stationarity.

The standard model for waiting times is the autoregressive conditional duration (ACD) model [see Engle and Russell (1998)]. We use a simple specification for waiting times between each trade, each coupled with i.i.d. Weibull errors,

$$\Delta T_i = w_i \varepsilon_i,$$

$$w_i = \varkappa + \gamma \Delta T_{i-1} + \delta w_{i-1},$$

where $\varkappa > 0$, $\gamma \geq 0$, and $\delta \geq 0$ in an analogous fashion to the GARCH model for returns.[16] Similarly to the GARCH model, numerous extensions to the ACD

[16] We use a Matlab package for ACD estimation, available in public domain (by Marcelo Perlin): https://sites.google.com/site/marceloperlin/matlab-code/estimation-and-simulation-of-acd-models-in-matlab (Accessed 1.6.2013). We modify it to take into account the stationarity requirement. We use the unconditional mean waiting time as a starting value.

model exist, but we refrain from using them to be conservative.[17]

We use the residuals of these models as the input data in the CTRW framework.[18] These residuals are much less autocorrelated than the original return and waiting time series, as measured by the Ljung-Box test statistic (see Table 7). The effect of the Weibull-ACD(1,1) model is particularly noticeable: in some cases, for example MMM, the sample autocorrelation cannot be easily distinguished from a Gaussian white noise series. The Normal-GARCH(1,1) does not appear to fit the return data so well. As an example, for GOOG and AAPL, a very significant amount of autocorrelation remains after filtering ($LB > 700$). This is expected to a large extent. Volatility clustering is not typically nearly as strong at high frequencies – especially in trade time – as it is at lower frequencies in wall-clock time (e.g., a day), where the GARCH models are commonly applied. The untypically low GARCH parameter β estimates ($\widehat{\beta} < 0.8$), except for MMM and CVX, prove that the volatility clustering effect is indeed weak.

The sample ACF shown in Figure 8 illustrates this point: while the Weibull-ACD(1,1) model does a very good job in removing autocorrelation from the original (unfiltered) series, the Normal-GARCH(1,1) model does not. For GOOG, for example, the absolute return ACF stays significantly above the upper confidence bound. This does not, of course, imply that no other GARCH model would not remove temporal dependence more efficiently, but such a study is out of the scope of this paper. In fact, there exists numerous ACD and GARCH models accounting for special features of financial time series [for a review, see, e.g., Teräsvirta (2009)]. Also, GARCH models have been modified high-frequency data in mind [see, e.g., Dionne, Duchesne, and Pacurar (2009)].

4.3 Performance evaluation

In this last empirical section, we evaluate the performance of different CTRW model specifications. We evaluate how the different marginal distributions affect performance and how much deseasonalization and GARCH/ACD-filtering add value. The performance evaluation is done with metrics familiar from financial risk management: VaR and ES statistics are commonly applied to evaluate the riskiness of an asset portfolio. We use these metrics to evaluate how different our CTRW model predictions are from the empirically realized values.

Typically, VaR and ES statistics are calculated for the negative (left-hand) tail. We also consider the positive tail. As yardsticks for aggregate performance across all stocks, we use both the root-squared-mean deviation (RSMD) and error in the mean. The closer these yardsticks are to zero, the better is the prediction of the CTRW model. Thus, in our view the best model specification is such that the realized risk is equal, or very close, to the estimated (predicted) risk. In other words, in the best possible scenario, $VaR_{est}/VaR_{emp} = 1$ and $ES_{est}/ES_{emp} = 1$, for both the negative and the positive tail separately. In case of an equal performance between different CTRW model specifications, we

[17] The only departure from the standard ACD model is the use of Weibull errors instead of exponential errors. We find the Weibull errors to significantly improve the fit.

[18] We scale the residuals to keep $\mathbb{E}[\Delta X^2]$ and $\mathbb{E}[\Delta T]$ constant.

prefer overestimation to underestimation of risk on the grounds that the latter is potentially more dangerous from the point of view of prudent risk management.

There are two different types of seasonality adjustments we make. The first option is to simply remove the first and last fifteen minutes of trading. The second option is to apply the seasonality adjustment procedure described in Section 4.1. We find that both procedures are useful even on a stand-alone basis, but the result is best when they are combined. In the tables (unless otherwise mentioned), we exclude the opening and closing time periods and then make the seasonality adjustment.

We first describe the Value-at-Risk results. Table 8 reports the performance of the different model specifications in terms of 99 percent VaR *before* deseasonalization and GARCH/ACD-filtering. We find significant underestimation of risk for most stocks, but its magnitude is related to the time-scale. For example, at the shortest price distribution prediction time-scale we consider (10 sec), the performance of all our models is bad for BA, particularly with exponentially distributed waiting times and normally distributed returns (E-N). A longer time-scale (2 min) tends to make the fit worse for the other stocks. With the largest time-scale (20 min), the results get a bit better again, with for example GOOG performing exceptionally well in both tails. The results for the largest time-scale demonstrate the aforementioned asymptotic result that all the model predictions are similar to each other – and to E-N, in particular.

Table 9 reports the VaR results *after* deseasonalization and GARCH/ACD-filtering. The E-N results are generally improved across all our time-scales. For example, at the shortest time-scale, the underestimation found for BA decreases from 0.4 (left-hand tail) and 0.5 (right-hand tail) to 0.75 and 0.67, respectively. Similarly, the model with mixed-Weibull distributed waiting times and double-exponentially distributed returns (MW-DE) improves from 0.55 and 0.68 to 0.99 and 0.89, respectively. Basically, all models perform much better, particularly at the middle-sized and largest time-scales. The estimates now seem to be well-centered around the empirical values on average. To better grasp the performance gain due to deseasonalization and GARCH/ACD-filtering, we visualize the tabled VaR values in Figure 9. The increased amount of white and green space in the right-hand side of the figure shows the effect of deseasonalization and GARCH/ACD-filtering. Red space (indicating underestimation of risk), on the other hand, appears to dominate the left-hand side of the figure.

Expected Shortfall can be argued to more accurately represent tail risk than VaR, because ES is the expected loss beyond a certain limit (here 1 percent). Thus, we repeat the above study for ES. The results in Tables 10 and 11 are similar to the VaR results presented above, but provide somewhat clearer evidence of the improved performance due to deseasonalization and GARCH/ACD-filtering. At the smallest time-scale, for example, the E-N model performance for BA rises from 0.27 (left-hand tail) to 0.70. Better performing models are the MW-DE model (left-hand tail 0.97) and W-DE (left-hand tail 0.99). At the largest time-scale, the E-N model performance for GOOG rises from 0.93 (left-hand tail) to 0.97. The W-DE model is again very close to the realized empirical value. The tabled ES values are illustrated in Figure 10. There is a clear shift from a red to a white-greenish color, representing the improved CTRW model performance due to deseasonalization and GARCH/ACD-filtering.

Thus far we have documented how the different CTRW models compare to each other, with and without deseasonalization and GARCH/ACD-filtering. Now we break down the effects of the individual components. More precisely, we seek to know what extensions matter the most. To make the results transparent, we calculate the RMSD and error in mean values for each CTRW model specification. Table 12 shows RMSD and error in mean values in terms of VaR. The column labeled "Filtering" denotes the three different filtering methods we apply: deseasonalization (D), GARCH (G), and ACD (A). Generally speaking, the more filtering is introduced, the smaller the deviations from the empirical values become. If all filtering methods are turned on (denoted DGA), the performance statistics decrease to about half or more of the unfiltered values. For example, RMSD (error in mean) for the best performing model (MW-ST) at the smallest time-scale decreases from 0.21 (-0.15) to 0.09 (-0.01). Notably, ACD-filtering has the largest effect across all the CTRW model specifications. For the same MW-ST model, its effect is around 50 percent. The effect is even larger at the middle-sized time-scale. The error in mean values are more dependent on the time-scale, however: at the smallest time-scale, introducing ACD-filtering causes the error in mean values to sometimes change the sign from a negative to a positive value, but this does not take place at the larger time-scales. MW-DE and MW-ST typically perform the best at the smallest time-scale, W-DE and W-ST at middle sized time-scales, while all the models perform equally well at the largest time-scale. With all filtering turned on, the RMSD metric gives roughly the same performance at all time-scales with most variation across the models at the smallest time-scale. While the models perform the best at the smallest time-scale, the differences between our four model specifications (i.e., excluding E-N) are however too small for making definitive conclusions.

Table 13 reports the corresponding results for ES. These results are perhaps less satisfying but more stable than the VaR results. Thus, they potentially reveal meaningful differences between the models. For example, for the best performing model (MW-DE) at the smallest time-scale, RMSD (error in mean) decreases from 0.27 (-0.21) to 0.13 (-0.01). It is striking that at the smallest time-scale, the error in mean is zero for the W-ST model after the ACD-filtering is introduced. The other models do very well in this respect as well, except E-N, for which the RMSD and error in mean values are consistently the largest. Deseasonalization and GARCH-filtering do not appear to be that important invidually nor together. The effect of deseasonalization, however, turns out to be more significant when the time-scale is middle sized or large.

Finally, in Tables 14 (VaR) and 15 (ES), we report the performance of the different CTRW model specifications *without* removing the opening and closing periods. The difference to the previous tables is that now the effect of seasonality removal is more visible. Without any filtering methods, the models have larger RMSD and error in mean values. Thus, while it is not necessary to remove the problematic opening and closing periods for good performance, deseasonalization and GARCH/ACD-filtering is required.

In summary, the above results show that the ACD-filtering component is by far the most important. In ranking, it is followed by deseasonalization especially at the middle-sized or large time-scales. The least important component appears to be GARCH-filtering. This is in line with what we would expect: standard GARCH models are not ideally suited for high-frequency data. ACD models, on the other hand, are specifically developed such data in mind.

5 Conclusions

In this paper, we describe the so-called continuous-time random walk (CTRW) framework in its generality. For the first time in literature, we take a detailed look at different CTRW model predictions by comparing them to the realized empirical price distributions. More exactly, we evaluate how accurate the intraday price distribution predictions are. We show that the simple no-memory structure of the CTRW framework limits the attainable accurary of the intraday price distribution predictions. Inspired by financial econometrics practice, we propose a few extentions to the standard CTRW and demonstrate their usefulness from a risk management perspective. The enhanced intraday price distribution predictions match much better to the realized empirical price distributions than the basic CTRW predictions do. We focus on evaluating the performance of the different CTRW models using the conventional financial risk management metrics of Value-at-Risk (VaR) and Expected Shortfall (ES).

The extensions we make are as follows. First, we extend the universe of potentially useful marginal distributions using numerical methods. For returns, we find the Student-t distribution of the family of generalized hyperbolic (GH) distributions to add much value over the normal distribution. The double-exponential distribution performs similarly to the Student-t distribution. For waiting times, we propose the mixed-Weibull distribution. In total, there are five different CTRW model specifications we compare: exponential-normal (E-N), Weibull-Student-t (W-ST), Weibull-Double-Exponential (W-DE), mixed-Weibull-Student-t (MW-ST), and mixed-Weibull-double-exponential (MW-DE). Typically, the winner in terms of the most realistic VaR and ES values, is MW-DE, and the runner-up is MW-ST. The differences are however quite small, except for E-N, which is consistently the worst performer. Although the newly proposed mixed-Weibull distribution clearly fits the body of the waiting time distribution better than Weibull does, the predicted fatter-tailed price distribution of the W-DE (W-ST) matters more in terms of the tail-sensitive VaR and ES statistics. We find the fat-tailedness of W-DE to be attractive when the turbulent opening and closing trading periods are included in the analysis.

Second, we make extensions with regards to intraday seasonality adjustment and temporal dependence. We find that seasonality matters especially for waiting times. The VaR and ES statistics improve at least a few percentages when seasonality is explicitly accounted for. There are two types of temporal dependence we consider. We filter out temporal dependence in returns using the standard generalized autoregressive conditional heteroskedastic (GARCH) model. We find its effect to be quite low, which is is understandable because our data are asynchronously spaced at high frequencies. The asynchronous nature is precisely the reason why the waiting time filtering performs very well. Of all our extentions, we find the standard autoregressive conditional duration (ACD) model to have the most important effect on the VaR and ES statistics.

In practical terms, we find our enhanced CTRW model specifications to improve performance over the standard CTRW model by around 40 percent. This increase is mostly due to more accurate tail modeling. Filtering out temporal dependence by the ACD model adds another 30 percentages compared to the basic model. Thus, in total, the performance gain we find is around 70 percent.

As the tails of the price distribution are potentially very important in financial decision making, our results should prove valuable for improving intraday risk management. The value is heightened by the fact that automated trading, and its subclass high-frequency trading (HFT) in particular, make up a significant portion of the total daily volume in some markets. By construction, the CTRW framework extracts information from all the asynchronously spaced observations that happen at subsecond time-scales and turn them to humanly understandable predictions in wall-clock time for any desired time-scale. This should be valuable for risk averse institutional investors executing large orders.

There are several directions for future research. First, more realistic marginal distributions could be applied. We find the GH distribution family to provide good candidates in this respect. In terms of waiting times, more work could be done with respect to mixtures of distributions along the lines of the proposed mixed-Weibull, for example. The effect of temporal dependence provides another interesting avenue. For example, GARCH effects can be better accounted for by applying specialized GARCH models for high-frequency data. The next step would be to include the non-Markovian effects in the CTRW framework, especially for waiting times. Finally, the independence assumption between returns and waiting times that was assumed in this paper can be relaxed in a mathematically rigorous fashion. We have made some progress along these lines and are keen to continue further.

A Appendix

A.1 Data pre-filtering

The logic of our pre-filtering algorithm is the following:

1. Import data for a given date.
2. Remove data outside continuous trading hours, 9:30am and 4pm. (We also discard data outside of 9:45am and 3:45pm unless otherwise mentioned.)
3. Merge trades time-stamped at the same nanosecond and save the last trade price.
4. Form the logarithmic price series.
5. Calculate the return series. Zero returns are removed by merging with the next increment until the return is non-zero.
6. Time-stamp each increment with date and time.
7. Append the dataset to the full incremental dataset.
8. Go back to step one.
9. After each date is imported, cumulatively sum the full incremental dataset, starting from $X(0) = 0$.

A.2 Numerical solution

A.2.1 Method

The algorithm we use to calculate $p(x, t)$ from the marginal distributions $f(x)$ and $g(t)$ works as follows. First, we calculate $[f(k_i)]$ and $[g(s_j)]$ at pre-defined values of k and s.[19] The Laplace transform of the Weibull distribution is not available in a convenient analytic form, so we calculate the values $[g(s_j)]$ by numerically integrating $g(t)e^{-st}$. Similarly, we use Matlab's Fast Fourier Transform (FFT) to calculate $[f(k_i)]$ from $f(x)$. Second, we use Eq. (2) to calculate $[p(k_i, s_j)]$ and then obtain its Fourier and Laplace inverses. The Laplace inverse is based on a method introduced by Valsa and Brancik (1998), which turns the integral over s into a finite sum over the discrete locations s_j.[20] For the Fourier inverse we use Matlab's inverse-FFT algorithm. With these two methods we can apply the inverse Fourier and Laplace algorithms to $[p(k_i, s_j)]$ in either order.

We next describe the Laplace and Fourier inversion algorithms. We then check the validity of this method against two known analytic solutions: (1) normal returns and exponential waiting times [see Scalas, Gorenflo, and Mainardi

[19] Here $i = 1...N_x(= 32768)$ and $j = 1...N_s(= 40)$. The accuracy can be arbitrarily increased by increasing N_x and N_s. The values of s_j depend on t.

[20] We have built our routine on top of a publicly available INVLAP Matlab package, available at http://www.mathworks.com/matlabcentral/fileexchange/ 32824-numerical-inversion-of-laplace-transforms-in-matlab. (Accessed 1.2.2013).

(2004)], and (2) coupled double-exponential returns and exponential waiting times [see Masoliver et al. (2006)].

A.2.2 Laplace transform inversion

To explain the idea behind the numerical Laplace inverse method introduced by Valsa and Brancik (1998), we examine the Bromwich integral form of the Laplace inverse:

$$g(t) = \int_{\gamma-i\infty}^{\gamma+i\infty} \frac{1}{2\pi i} e^{st} g(s) ds,$$

where $g(s)$ is assumed regular for $\mathrm{Re}(s) > 0$ and vanishing for $|s| \to \infty$. The integration contour is along the imaginary axis, with $\mathrm{Re}(s) = \gamma$. In this method, the exponential e^{st} is first approximated by

$$e^{st} \approx \frac{1}{e^{-st} + e^{-2a} e^{st}} = \frac{e^a}{2\cosh(a - st)} = \pi e^a \sum_{n=0}^{\infty} \frac{(-1)^n (n + \frac{1}{2})}{(n + \frac{1}{2})^2 \pi^2 + (a - st)^2},$$

which is accurate in the limit $a \gg \gamma t$. Changing the order of integration and summation, we get

$$g(t) \approx \pi e^a \sum_{n=0}^{\infty} \int_{\gamma-i\infty}^{\gamma+i\infty} \frac{1}{2\pi i} \frac{(-1)^n (n + \frac{1}{2})}{(n + \frac{1}{2})^2 \pi^2 + (a - st)^2} g(s) ds.$$

Since $g(s)$ vanishes at infinity, the integration contour can be closed with an infinite semicircle around the positive real axis. The key point is that since $g(s)$ has no poles in this area, the integral can be calculated as a sum over the residues, which are located at zeros of the denominator:

$$s_{\pm} = \frac{a \pm i(n + \frac{1}{2})\pi}{t}.$$

Thus, this method turns the integral over s into an infinite sum over t-dependent discrete locations s_i. Truncation of this sum is another source of numerical error. In this regard, the method includes further optimization of the sum for faster convergence.

A.2.3 Fourier transform inversion

The Fourier inverse integral can easily be discretized and written as a discete inverse-FFT transformation. We first consider the forward Fourier transform. Assuming $f(x)$ to be non-zero only in the interval $0 \leq x \leq L_x$, we write the Fourier transform as

$$f(k) = \int_{-\infty}^{\infty} e^{ikx} f(x) dx = \int_0^{L_x} e^{ikx} f(x) dx.$$

Discretizing with grid spacing δx and δk, and defining $x_n = n\delta x$ and $k_n = n\delta k$, gives

$$f(k_m) = \delta x \sum_{n=0}^{N_x-1} \exp(ix_n k_m) f(x_n) = \delta x \sum_n \left(e^{\frac{-2\pi i}{N_x}}\right)^{-mn\frac{\delta k \delta x N_x}{2\pi}} f(x_n).$$

where $N_x = L_x/\delta x$. We then require $\delta k \delta x N_x / 2\pi \equiv 1$, and denote $\omega_{Nx} = \exp(-2\pi i/N_x)$, giving

$$f(k_m) = \delta x \sum_{n=0}^{N_x-1} \omega_{N_x}^{-mn} f(x_n) = \delta x \mathrm{fft}_n(f(x_n), m),$$

where $\mathrm{fft}_n(f(x_n), m)$ is the standard FFT. Assuming ifft_m is the inverse-FFT, we have

$$f_n = \frac{1}{\delta x} \mathrm{ifft}_m(f(k_m), n).$$

A.2.4 Analytic solutions

For the first analytic solution, we use exponentially distributed waiting times and normally distributed returns. Then the solution to Eq. (1) can be written as an infinite sum over the number of jumps up to time t [see Scalas, Gorenflo, and Mainardi (2004)],

$$p(x,t) = \sum_{n=0}^{\infty} P(n,t) f_n(x),$$

where $P(n,t)$ gives the probability of n jumps at time t and $f_n(x)$ gives the returns distribution after n jumps. The function P is given by the Poisson distribution,

$$P(n,t) = \frac{(t/\mu)^n}{n!} e^{-t/\mu}.$$

The function $f_n(x)$ is the Fourier inverse transform of $f(k)^n$,

$$\begin{aligned} f_0(x) &= \delta(x) \\ f_n(x) &= \frac{1}{\sqrt{2\pi}\sigma\sqrt{n}} \exp\left(-\frac{x^2}{2n\sigma^2}\right). \end{aligned}$$

Thus,

$$p(x,t) = e^{-t/\mu} \left[\delta(x) + \sum_{n=1}^{\infty} \frac{(t/\mu)^n}{n!} \frac{1}{\sqrt{2\pi}\sigma\sqrt{n}} \exp\left(-\frac{x^2}{2n\sigma^2}\right)\right].$$

When comparing this solution to the numeric one, we truncate the sum at N terms. N is taken much larger than the expected number of jumps up to time t: $N \approx 10t/\mu$. We discover a very good convergence between the analytic and numerical solutions (see the left-hand panel of Figure 11). At relevant density values, the analytic and numerical densities agree point-by-point to the accuracy of 1/1000.

Another analytic solution can be found for exponentially distributed waiting times and double-exponentially distributed returns. Also the waiting times and returns are non-trivially coupled, so that the multivariate distribution is

$$h(x,t) = \frac{1}{2\gamma\sqrt{\pi\mu t}} \exp\left(-\frac{\mu x^2}{4t\gamma^2} - \frac{t}{\mu}\right).$$

Masoliver et al. (2006) study this case and find

$$p(x,t) = e^{-t/\mu}\left(\delta(x) + \frac{1}{\gamma\sqrt{\pi}}\int_0^{\sqrt{t/\mu}} e^{\xi^2 - x^2/4\xi^2\gamma^2}\, d\xi\right).$$

We fit the model on real data and compare the two solutions, showing a good convergence (see the right-hand panel of Figure 11). At relevant values of the density, the analytic and numerical densities agree point-by-point to the accuracy of 1/100. The slight decrease in performance is probably related to the complexity of the model, given that the waiting times and returns are coupled.

Table 1: Pre-filtering statistics.

Stock	Trades	$\Delta T \neq 0$	$\Delta X \neq 0$	$\overline{P}$
GOOG Week 10	36248	28612	17794	829
GOOG Week 11	22947	17971	11395	824
GOOG Week 12	18312	15050	9449	812
AAPL	308543	228681	134425	437
BA	82377	49409	23074	83
CVX	78486	54542	24919	119
GOOG	77507	61633	38638	824
IBM	71573	48427	26687	211
MMM	37896	25986	13536	105

Note: $\Delta X \neq 0$ is the number of valid observations.

Table 2: Trades in the range of [i,j] ticks.

Stock	$(\Delta Y)_1^1$	$(\Delta Y)_2^3$	$(\Delta Y)_4^7$	$(\Delta Y)_8^\infty$
GOOG Week 10	0.17	0.16	0.24	0.43
GOOG Week 11	0.2	0.17	0.24	0.39
GOOG Week 12	0.16	0.14	0.22	0.47
AAPL	0.29	0.28	0.28	0.15
BA	0.74	0.23	0.03	0
CVX	0.71	0.26	0.04	0
GOOG	0.18	0.16	0.24	0.43
IBM	0.52	0.34	0.12	0.02
MMM	0.66	0.28	0.05	0

Table 3: Descriptive statistics of returns and waiting times.

Stock	Min	1st Q.	Med	3rd Q.	Max	Mean	Std
Panel 1 (Returns)							
GOOG Week 10	−1	−0.072	0.012	0.072	0.82	0.00056	0.14
GOOG Week 11	−0.83	−0.061	−0.012	0.061	0.97	−0.0015	0.13
GOOG Week 12	−1.3	−0.086	−0.012	0.086	0.93	−0.0018	0.17
AAPL	−1.2	−0.069	0.022	0.069	0.95	0.00038	0.13
BA	−1.3	−0.12	0.12	0.12	1.5	0.0016	0.2
CVX	−1	−0.084	0.083	0.084	0.85	0.0012	0.14
GOOG	−1.3	−0.072	−0.012	0.072	0.97	−0.00062	0.15
IBM	−1.1	−0.048	0.046	0.049	0.84	0.0012	0.12
MMM	−1.6	−0.095	0.094	0.095	1.8	$-5.1e-05$	0.18
Panel 2 (W. times)							
GOOG Week 10	$1.2e-06$	0.01	1.3	6.7	200	6.1	–
GOOG Week 11	$9.2e-07$	0.012	2.2	12	210	9.5	–
GOOG Week 12	$2.7e-06$	0.027	3.7	15	250	11	–
AAPL	$1.1e-06$	0.0098	0.53	2.8	110	2.4	–
BA	$6.3e-06$	0.55	5.6	18	360	14	–
CVX	$1.9e-06$	0.76	6.4	18	250	13	–
GOOG	$9.2e-07$	0.013	2	9.8	250	8.4	–
IBM	$1.9e-06$	0.47	5.1	16	220	12	–
MMM	$1.6e-06$	2.3	12	32	440	24	–

Note: Panel 1 statistics multiplied by 1000. Panel 2 statistics in seconds.

Table 4: Marginal distribution parameter estimates.

	Normal	D-Exp	Student-t		Exp	Weibull		Mixed-Weibull				
Stock	$\hat{\sigma}$	$\hat{\gamma}$	$\hat{\sigma}$	$\hat{\nu}$	$\hat{\mu}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{p}$	$\hat{\lambda}_1$	$\hat{\alpha}_1$	$\hat{\lambda}_2$	$\hat{\alpha}_2$
Panel 1 (Before)												
GOOG Week 10	0.00014	0.0001	0.0001 (1e−07)	3.9	6.1	1.8	0.396 (0.0007)	0.3046 (0.0002)	0.00413 (1e−05)	0.5223 (0.0007)	6.5	0.6653 (0.0003)
GOOG Week 11	0.00013	$9.3e-05$	$9.08e-05$ (2e−07)	3.8	9.5	3	0.409 (0.001)	0.3114 (0.0002)	0.00495 (2e−05)	0.523 (0.001)	11	0.7175 (0.0004)
GOOG Week 12	0.00017	0.00012	0.0001158 (2e−07)	3.7	11	5.2	0.476 (0.001)	0.2743 (0.0002)	0.00379 (2e−05)	0.516 (0.001)	13	0.7618 (0.0004)
Panel 2 (Before)												
AAPL	0.00013	$8.9e-05$	$9.692e-05$ (8e−08)	5	2.4	0.81	0.4171 (0.0002)	0.29013 (6e−05)	0.004042 (3e−06)	0.5954 (0.0002)	2.6	0.66746 (8e−05)
BA	0.0002	0.00014	0.000192 (2e−06)	57	14	8.6	0.5662 (0.0006)	0.1827 (0.0001)	0.02014 (5e−05)	0.5905 (0.0009)	14	0.7477 (0.0001)
CVX	0.00014	0.0001	0.00014	∞	13	9.3	0.6357 (0.0008)	0.1872 (0.0001)	0.01988 (6e−05)	0.604 (0.001)	15	0.857 (0.0002)
GOOG	0.00015	0.0001	0.00010036 (9e−08)	3.8	8.4	2.7	0.41 (0.0005)	0.2982 (0.0001)	0.004238 (8e−06)	0.5223 (0.0005)	9.3	0.6868 (0.0002)
IBM	0.00012	$8.7e-05$	$9.78e-05$ (4e−07)	5.3	12	7.7	0.5812 (0.0007)	0.18907 (9e−05)	0.01098 (2e−05)	0.677 (0.001)	13	0.7782 (0.0001)
MMM	0.00018	0.00013	0.000173 (3e−06)	23	24	18	0.6774 (0.0008)	0.13125 (9e−05)	0.0167 (6e−05)	0.625 (0.002)	25	0.8206 (0.0001)
Panel 3 (After)												
AAPL	0.00013	$8.9e-05$	$9.797e-05$ (5e−08)	5.1	2.4	0.95	0.4445 (0.0003)	0.30277 (6e−05)	0.005402 (4e−06)	0.5459 (0.0002)	2.9	0.75214 (9e−05)
BA	0.0002	0.00014	0.0002	∞	14	9.6	0.6142 (0.0007)	0.1682 (0.0001)	0.02623 (8e−05)	0.599 (0.001)	15	0.7943 (0.0002)
CVX	0.00014	0.0001	0.00014	∞	13	9.7	0.6599 (0.0009)	0.1893 (0.0001)	0.02488 (8e−05)	0.579 (0.001)	15	0.8964 (0.0002)
GOOG	0.00015	0.0001	0.00010244 (7e−08)	3.9	8.4	3.1	0.4362 (0.0006)	0.3132 (0.0001)	0.00624 (1e−05)	0.4787 (0.0005)	11	0.7787 (0.0002)
IBM	0.00012	$8.7e-05$	0.00012	∞	12	8.3	0.6149 (0.0008)	0.19596 (9e−05)	0.01446 (3e−05)	0.634 (0.001)	14	0.8407 (0.0002)
MMM	0.00018	0.00013	0.00018	∞	24	19	0.718 (0.001)	0.1401 (0.0001)	0.029 (0.0001)	0.563 (0.001)	26	0.8878 (0.0002)

Note: Panels 1, 2 and Panel 3 denote before and after deseasonalization and GARCH/ACD-filtering, respectively. Parameters estimated by a non-linear least squares method as explained in the main text, with standard errors in parenthesis.

Table 5: Root-mean-squared deviation of the best fit.

Stock	Normal	D-Exp	ST	Exp	Weibull	MW
Panel 1 (Before)						
GOOG Week 10	0.039	0.014	0.017	0.208	0.074	0.005
GOOG Week 11	0.042	0.018	0.02	0.197	0.082	0.006
GOOG Week 12	0.043	0.014	0.019	0.17	0.082	0.005
Panel 2 (Before)						
AAPL	0.033	0.028	0.021	0.202	0.061	0.004
BA	0.069	0.08	0.069	0.137	0.048	0.003
CVX	0.073	0.088	0.073	0.11	0.053	0.004
GOOG	0.042	0.012	0.017	0.2	0.078	0.005
IBM	0.052	0.052	0.045	0.131	0.054	0.003
MMM	0.072	0.082	0.072	0.096	0.038	0.002
Panel 3 (After)						
AAPL	0.033	0.032	0.021	0.183	0.067	0.004
BA	0.091	0.124	0.091	0.117	0.045	0.004
CVX	0.083	0.118	0.083	0.102	0.053	0.004
GOOG	0.039	0.012	0.016	0.182	0.084	0.005
IBM	0.052	0.078	0.052	0.118	0.056	0.004
MMM	0.07	0.103	0.07	0.083	0.04	0.003

Note: Panels 1, 2 and Panel 3 denote before and after deseasonalization and GARCH/ACD-filtering, respectively. $RMSD = \sqrt{\sum_i (F(x_i) - \widehat{F}_n(x_i))^2/n}$, where $F(x_i)$ is the theoretical CDF and $\widehat{F}_n(x_i)$ is the empirical CDF.

Table 6: Mean waiting time in each time-slot (in seconds).

Stock	09:30–09:45	09:45–10:45	10:45–11:45	11:45–12:45	
AAPL	0.9	1.5	2.5	2.9	
BA	5.8	8.9	12.8	17	
CVX	7.1	9.2	12.4	15.3	
GOOG	3	4.7	8.9	10.7	
IBM	7.7	9.2	11.7	14.8	
MMM	15	15.8	25	28.5	
	12:45–13:45	13:45–14:45	14:45–15:45	15:45–16:00	Aggr.
AAPL	3.4	3.1	2.2	0.9	2.1
BA	18.4	17.9	14.7	8.8	13
CVX	18.5	14.5	11.9	4.8	11.8
GOOG	12.9	10.8	7.9	2.5	7.2
IBM	15.6	13.9	10.3	4.1	11.1
MMM	34.4	27.6	21.1	9.5	22.1

Table 7: GARCH and ACD model parameter estimates.

	GARCH(1,1)			Dependence		ACD(1,1)				Dependence	
Stock	$\widehat{\omega}$	$\widehat{\alpha}$	$\widehat{\beta}$	LB_b	LB_a	$\widehat{\varkappa}$	$\widehat{\gamma}$	$\widehat{\delta}$	Weibull	LB_b	LB_a
AAPL	$4e-09$ $(4e-09)$	0.104 (0.002)	0.674 (0.002)	6800	760	0.113 (0.006)	0.131 (0.005)	0.868 (0.005)	0.3934 (0.0009)	26000	72
BA	$2e-08$ $(1e-08)$	0.136 (0.008)	0.426 (0.008)	2000	420	0.008 (0.005)	0.091 (0.004)	0.908 (0.004)	0.527 (0.003)	6900	110
CVX	$2e-09$ $(8e-09)$	0.078 (0.004)	0.818 (0.004)	2500	98	0.371 (0.046)	0.114 (0.011)	0.883 (0.011)	0.518 (0.003)	4600	86
GOOG	$8e-09$ $(8e-09)$	0.134 (0.004)	0.499 (0.004)	2800	770	0.644 (0.059)	0.123 (0.01)	0.875 (0.01)	0.351 (0.002)	7000	48
IBM	$3e-09$ $(8e-09)$	0.149 (0.005)	0.675 (0.004)	6000	410	0.309 (0.042)	0.098 (0.01)	0.899 (0.01)	0.482 (0.003)	5600	59
MMM	0 $(1e-08)$	0.062 (0.003)	0.889 (0.003)	1000	36	0.429 (0.062)	0.103 (0.01)	0.894 (0.01)	0.568 (0.005)	4000	35

Note: LB_b and LB_a refer to the Ljung-Box (20 lags) test statistic before and after filtering, respectively.

Table 8: Performance in terms of VaR (99 percent) before filtering.

Time-scale	Stock	Tail	MW-DE	MW-ST	W-DE	W-ST	E-N
10 sec	AAPL	Left	0.77	0.77	0.85	0.85	0.62
		Right	0.74	0.73	0.81	0.82	0.6
	BA	Left	0.55	0.52	0.57	0.53	0.4
		Right	0.68	0.64	0.71	0.67	0.5
	CVX	Left	1.12	1.05	1.15	1.08	0.86
		Right	0.94	0.88	0.97	0.91	0.72
	GOOG	Left	0.85	0.85	0.97	0.97	0.6
		Right	1.01	1.01	1.14	1.15	0.7
	IBM	Left	0.92	0.89	0.95	0.93	0.68
		Right	0.94	0.91	0.97	0.95	0.7
	MMM	Left	1.06	0.99	1.06	0.99	0.8
		Right	1.06	1	1.06	1	0.8
2 min	AAPL	Left	0.79	0.8	0.83	0.83	0.77
		Right	0.64	0.64	0.67	0.67	0.62
	BA	Left	0.58	0.57	0.6	0.59	0.53
		Right	0.58	0.57	0.6	0.59	0.53
	CVX	Left	0.83	0.81	0.85	0.84	0.78
		Right	0.83	0.81	0.85	0.83	0.77
	GOOG	Left	0.79	0.81	0.86	0.87	0.72
		Right	0.88	0.9	0.96	0.97	0.81
	IBM	Left	0.64	0.64	0.66	0.66	0.59
		Right	0.69	0.69	0.71	0.71	0.64
	MMM	Left	0.77	0.75	0.79	0.77	0.7
		Right	0.77	0.75	0.78	0.76	0.7
20 min	AAPL	Left	0.92	0.92	0.93	0.93	0.92
		Right	0.69	0.69	0.69	0.69	0.68
	BA	Left	0.83	0.83	0.84	0.83	0.82
		Right	0.61	0.61	0.61	0.61	0.6
	CVX	Left	0.94	0.93	0.94	0.94	0.93
		Right	0.92	0.92	0.93	0.92	0.91
	GOOG	Left	0.99	1	1.01	1.02	0.98
		Right	0.96	0.96	0.98	0.98	0.95
	IBM	Left	0.71	0.71	0.72	0.72	0.71
		Right	0.78	0.78	0.79	0.79	0.78
	MMM	Left	0.84	0.84	0.85	0.84	0.83
		Right	0.77	0.76	0.77	0.76	0.75

Note: Ratio between the estimated and empirical VaR (VaR_{est}/VaR_{emp}) before deseasonalization and GARCH/ACD-filtering.

Table 9: Performance in terms of VaR (99 percent) after filtering.

Time-scale	Stock	Tail	MW-DE	MW-ST	W-DE	W-ST	E-N
10 sec	AAPL	Left	0.92	0.91	1.01	1.01	0.76
		Right	0.91	0.91	1.01	1	0.76
	BA	Left	0.99	0.93	1.02	0.96	0.75
		Right	0.89	0.83	0.91	0.86	0.67
	CVX	Left	1.21	1.13	1.24	1.16	0.95
		Right	1.05	0.99	1.08	1.01	0.83
	GOOG	Left	1.04	1.03	1.18	1.18	0.76
		Right	1.11	1.11	1.26	1.26	0.81
	IBM	Left	1.05	0.99	1.09	1.02	0.81
		Right	1.04	0.98	1.08	1.02	0.8
	MMM	Left	1.13	1.05	1.12	1.05	0.88
		Right	1.16	1.08	1.15	1.08	0.9
2 min	AAPL	Left	0.96	0.96	1	1	0.94
		Right	0.85	0.85	0.87	0.88	0.82
	BA	Left	0.86	0.84	0.88	0.86	0.8
		Right	0.83	0.81	0.85	0.83	0.77
	CVX	Left	1.04	1.02	1.07	1.04	0.98
		Right	1.01	0.99	1.03	1.01	0.95
	GOOG	Left	1	1.02	1.09	1.1	0.93
		Right	1.08	1.09	1.17	1.18	1
	IBM	Left	0.83	0.81	0.85	0.83	0.77
		Right	0.92	0.9	0.94	0.93	0.86
	MMM	Left	1.04	1	1.06	1.02	0.95
		Right	0.96	0.93	0.98	0.95	0.88
20 min	AAPL	Left	1.09	1.09	1.1	1.09	1.09
		Right	0.79	0.79	0.79	0.79	0.79
	BA	Left	0.93	0.92	0.93	0.93	0.92
		Right	0.81	0.81	0.82	0.81	0.8
	CVX	Left	1.03	1.03	1.03	1.03	1.02
		Right	0.96	0.95	0.96	0.96	0.95
	GOOG	Left	1.05	1.05	1.07	1.07	1.04
		Right	1.19	1.19	1.21	1.21	1.18
	IBM	Left	0.84	0.83	0.84	0.84	0.83
		Right	0.94	0.93	0.94	0.94	0.93
	MMM	Left	0.94	0.94	0.95	0.94	0.93
		Right	0.91	0.91	0.92	0.91	0.9

Note: Ratio between the estimated and empirical VaR (VaR_{est}/VaR_{emp}) after deseasonalization and GARCH/ACD-filtering.

Table 10: Performance in terms ES (99 percent) before filtering.

Time-scale	Stock	Tail	MW-DE	MW-ST	W-DE	W-ST	E-N
10 sec	AAPL	Left	0.75	0.76	0.83	0.84	0.59
		Right	0.69	0.7	0.76	0.77	0.54
	BA	Left	0.39	0.36	0.41	0.37	0.27
		Right	0.63	0.58	0.65	0.6	0.44
	CVX	Left	1.02	0.92	1.03	0.94	0.75
		Right	0.75	0.67	0.76	0.69	0.55
	GOOG	Left	0.78	0.82	0.87	0.92	0.52
		Right	0.98	1.03	1.09	1.16	0.65
	IBM	Left	0.8	0.8	0.82	0.82	0.57
		Right	0.84	0.83	0.86	0.86	0.6
	MMM	Left	0.93	0.85	0.93	0.84	0.67
		Right	0.9	0.82	0.9	0.81	0.64
2 min	AAPL	Left	0.78	0.79	0.82	0.82	0.75
		Right	0.55	0.56	0.58	0.58	0.53
	BA	Left	0.53	0.52	0.55	0.54	0.47
		Right	0.57	0.56	0.59	0.58	0.51
	CVX	Left	0.78	0.75	0.8	0.77	0.71
		Right	0.74	0.71	0.76	0.74	0.68
	GOOG	Left	0.67	0.71	0.74	0.77	0.6
		Right	0.89	0.94	0.97	1.02	0.79
	IBM	Left	0.59	0.6	0.61	0.62	0.54
		Right	0.66	0.67	0.68	0.69	0.6
	MMM	Left	0.77	0.73	0.78	0.75	0.67
		Right	0.69	0.65	0.7	0.67	0.6
20 min	AAPL	Left	0.92	0.92	0.93	0.93	0.92
		Right	0.58	0.58	0.58	0.58	0.57
	BA	Left	0.83	0.82	0.84	0.83	0.81
		Right	0.56	0.56	0.57	0.57	0.55
	CVX	Left	0.89	0.89	0.9	0.89	0.88
		Right	0.93	0.93	0.94	0.93	0.92
	GOOG	Left	0.94	0.96	0.97	0.98	0.93
		Right	0.86	0.88	0.89	0.9	0.85
	IBM	Left	0.63	0.63	0.63	0.63	0.62
		Right	0.78	0.78	0.78	0.79	0.77
	MMM	Left	0.81	0.8	0.81	0.8	0.79
		Right	0.81	0.81	0.82	0.81	0.79

Note: Ratio between the estimated and empirical ES (ES_{est}/ES_{emp}) before deseasonalization and GARCH/ACD-filtering.

Table 11: Performance in terms of ES (99 percent) after filtering.

Time-scale	Stock	Tail	MW-DE	MW-ST	W-DE	W-ST	E-N
10 sec	AAPL	Left	0.91	0.92	1.01	1.02	0.74
		Right	0.89	0.9	0.98	0.99	0.72
	BA	Left	0.97	0.88	0.99	0.9	0.7
		Right	0.67	0.61	0.69	0.62	0.49
	CVX	Left	1.21	1.1	1.23	1.12	0.91
		Right	0.92	0.83	0.93	0.85	0.69
	GOOG	Left	0.98	1.03	1.09	1.16	0.68
		Right	1.07	1.12	1.2	1.27	0.75
	IBM	Left	0.98	0.89	1.01	0.92	0.72
		Right	1	0.91	1.03	0.94	0.74
	MMM	Left	1.1	0.98	1.09	0.97	0.81
		Right	1.14	1.02	1.13	1	0.84
2 min	AAPL	Left	0.96	0.96	1	1	0.92
		Right	0.79	0.79	0.82	0.83	0.76
	BA	Left	0.77	0.74	0.79	0.76	0.69
		Right	0.81	0.78	0.83	0.81	0.73
	CVX	Left	1.03	1	1.06	1.02	0.95
		Right	0.9	0.87	0.93	0.9	0.83
	GOOG	Left	0.91	0.95	0.99	1.03	0.82
		Right	1.06	1.11	1.16	1.21	0.96
	IBM	Left	0.79	0.77	0.82	0.79	0.73
		Right	0.89	0.86	0.91	0.88	0.81
	MMM	Left	1.02	0.97	1.04	0.99	0.92
		Right	0.9	0.85	0.91	0.87	0.8
20 min	AAPL	Left	1.1	1.1	1.11	1.11	1.09
		Right	0.64	0.64	0.64	0.64	0.64
	BA	Left	0.91	0.91	0.92	0.91	0.9
		Right	0.78	0.78	0.79	0.79	0.77
	CVX	Left	0.9	0.9	0.91	0.9	0.89
		Right	0.98	0.98	0.99	0.98	0.97
	GOOG	Left	0.98	0.99	1	1.01	0.97
		Right	1.19	1.21	1.22	1.24	1.18
	IBM	Left	0.84	0.84	0.85	0.84	0.83
		Right	0.92	0.92	0.93	0.92	0.91
	MMM	Left	0.88	0.87	0.89	0.88	0.87
		Right	0.93	0.92	0.93	0.93	0.91

Note: Ratio between the estimated and empirical ES (ES_{est}/ES_{emp}) after deseasonalization and GARCH/ACD-filtering.

Table 12: Performance statistics in terms of VaR (99 percent).

Time-scale	Filtering	MW-DE	MW-ST	W-DE	W-ST	E-N
Panel 1 (RMSD metric)						
10 sec	***	0.2	0.21	0.18	0.19	0.36
	D**	0.19	0.2	0.18	0.19	0.34
	G	0.2	0.22	0.17	0.19	0.37
	DG*	0.19	0.21	0.16	0.18	0.35
	**A	0.13	0.11	0.16	0.13	0.21
	D*A	0.14	0.11	0.17	0.14	0.2
	*GA	0.1	0.09	0.13	0.11	0.21
	DGA	0.11	0.09	0.14	0.12	0.21
2 min	***	0.28	0.29	0.26	0.27	0.33
	D**	0.23	0.24	0.21	0.22	0.28
	G	0.27	0.28	0.25	0.26	0.32
	DG*	0.22	0.23	0.2	0.22	0.27
	**A	0.11	0.11	0.1	0.11	0.15
	D*A	0.11	0.11	0.11	0.11	0.14
	*GA	0.1	0.11	0.09	0.1	0.14
	DGA	0.1	0.11	0.1	0.11	0.14
20 min	***	0.21	0.21	0.2	0.2	0.21
	D**	0.16	0.16	0.16	0.16	0.16
	G	0.21	0.21	0.2	0.2	0.21
	DG*	0.17	0.17	0.16	0.17	0.17
	**A	0.13	0.13	0.13	0.13	0.13
	D*A	0.13	0.13	0.13	0.13	0.13
	*GA	0.11	0.12	0.11	0.12	0.12
	DGA	0.12	0.12	0.12	0.13	0.12
Panel 2 (Error in mean)						
10 sec	***	−0.12	−0.15	−0.07	−0.1	−0.34
	D**	−0.1	−0.13	−0.05	−0.08	−0.32
	G	−0.14	−0.17	−0.09	−0.13	−0.35
	DG*	−0.12	−0.15	−0.06	−0.1	−0.33
	**A	0.05	0.02	0.11	0.07	−0.19
	D*A	0.06	0.02	0.12	0.08	−0.18
	*GA	0.03	−0.01	0.09	0.04	−0.2
	DGA	0.04	−0.01	0.1	0.05	−0.2
2 min	***	−0.27	−0.27	−0.24	−0.24	−0.32
	D**	−0.21	−0.21	−0.17	−0.18	−0.26
	G	−0.25	−0.26	−0.22	−0.23	−0.31
	DG*	−0.2	−0.21	−0.16	−0.17	−0.25
	**A	−0.07	−0.08	−0.04	−0.04	−0.13
	D*A	−0.06	−0.07	−0.03	−0.03	−0.12
	*GA	−0.06	−0.07	−0.03	−0.04	−0.12
	DGA	−0.05	−0.06	−0.02	−0.03	−0.11
20 min	***	−0.17	−0.17	−0.16	−0.16	−0.18
	D**	−0.1	−0.1	−0.09	−0.09	−0.11
	G	−0.17	−0.17	−0.16	−0.17	−0.18
	DG*	−0.11	−0.11	−0.11	−0.11	−0.12
	**A	−0.06	−0.06	−0.06	−0.06	−0.07
	D*A	−0.05	−0.05	−0.04	−0.04	−0.05
	*GA	−0.06	−0.06	−0.05	−0.05	−0.07
	DGA	−0.04	−0.05	−0.04	−0.04	−0.05

Note: The column "Filtering" denotes deseasonalization (D), GARCH (G), and ACD (A). $RMSD = \sqrt{\sum_i \left(VaR_{est_i}/VaR_{emp_i} - 1\right)^2/n}$ and $Mean = \sum_i \left(VaR_{est_i}/VaR_{emp_i} - 1\right)/n$.

Table 13: Performance statistics in terms of ES (99 percent).

Time-scale	Filtering	MW-DE	MW-ST	W-DE	W-ST	E-N
Panel 1 (RMSD metric)						
10 sec	***	0.27	0.29	0.24	0.27	0.44
	D**	0.26	0.28	0.24	0.27	0.43
	G	0.26	0.29	0.23	0.27	0.44
	DG*	0.25	0.28	0.23	0.27	0.43
	**A	0.13	0.13	0.13	0.15	0.29
	D*A	0.13	0.14	0.13	0.15	0.28
	*GA	0.13	0.14	0.13	0.15	0.29
	DGA	0.13	0.15	0.14	0.16	0.29
2 min	***	0.33	0.33	0.31	0.31	0.39
	D**	0.3	0.31	0.28	0.29	0.36
	G	0.31	0.33	0.29	0.3	0.38
	DG*	0.28	0.3	0.26	0.28	0.34
	**A	0.16	0.16	0.14	0.15	0.22
	D*A	0.15	0.16	0.14	0.15	0.21
	*GA	0.14	0.16	0.13	0.14	0.2
	DGA	0.14	0.15	0.13	0.15	0.19
20 min	***	0.24	0.24	0.24	0.24	0.25
	D**	0.21	0.21	0.21	0.21	0.22
	G	0.24	0.24	0.23	0.24	0.25
	DG*	0.2	0.21	0.2	0.21	0.21
	**A	0.16	0.16	0.16	0.16	0.17
	D*A	0.17	0.18	0.17	0.18	0.18
	*GA	0.14	0.15	0.14	0.14	0.15
	DGA	0.16	0.16	0.16	0.16	0.16
Panel 2 (Error in mean)						
10 sec	***	−0.21	−0.23	−0.17	−0.19	−0.43
	D**	−0.2	−0.22	−0.16	−0.18	−0.42
	G	−0.2	−0.24	−0.16	−0.2	−0.42
	DG*	−0.2	−0.24	−0.16	−0.2	−0.41
	**A	−0.02	−0.05	0.03	0	−0.27
	D*A	−0.01	−0.05	0.03	0	−0.27
	*GA	−0.02	−0.07	0.03	−0.02	−0.27
	DGA	−0.01	−0.07	0.03	−0.02	−0.27
2 min	***	−0.31	−0.32	−0.28	−0.29	−0.38
	D**	−0.27	−0.28	−0.24	−0.25	−0.34
	G	−0.3	−0.31	−0.27	−0.28	−0.36
	DG*	−0.26	−0.27	−0.23	−0.24	−0.33
	**A	−0.12	−0.13	−0.09	−0.09	−0.19
	D*A	−0.11	−0.12	−0.08	−0.08	−0.19
	*GA	−0.1	−0.12	−0.07	−0.08	−0.18
	DGA	−0.1	−0.11	−0.06	−0.08	−0.17
20 min	***	−0.2	−0.2	−0.2	−0.2	−0.22
	D**	−0.15	−0.15	−0.14	−0.14	−0.16
	G	−0.2	−0.2	−0.19	−0.2	−0.21
	DG*	−0.15	−0.15	−0.14	−0.14	−0.16
	**A	−0.09	−0.09	−0.08	−0.08	−0.1
	D*A	−0.09	−0.09	−0.08	−0.08	−0.1
	*GA	−0.09	−0.09	−0.08	−0.08	−0.1
	DGA	−0.08	−0.08	−0.07	−0.07	−0.09

Note: The column "Filtering" denotes deseasonalization (D), GARCH (G), and ACD (A). $RMSD = \sqrt{\sum_i \left(ES_{est_i}/ES_{emp_i} - 1\right)^2/n}$ and $Mean = \sum_i \left(ES_{est_i}/ES_{emp_i} - 1\right)/n$.

Table 14: Performance statistics in terms of VaR (99 percent), 9:30am-4:00pm.

Time-scale	Filtering	MW-DE	MW-ST	W-DE	W-ST	E-N
Panel 1 (RMSD metric)						
10 sec	***	0.24	0.25	0.2	0.21	0.41
	D**	0.18	0.19	0.16	0.16	0.34
	G	0.23	0.25	0.19	0.21	0.41
	DG*	0.19	0.21	0.16	0.18	0.36
	**A	0.11	0.09	0.14	0.12	0.21
	D*A	0.11	0.09	0.14	0.12	0.2
	*GA	0.09	0.08	0.1	0.09	0.23
	DGA	0.09	0.08	0.11	0.1	0.23
2 min	***	0.35	0.35	0.33	0.33	0.4
	D**	0.26	0.26	0.24	0.24	0.31
	G	0.34	0.34	0.31	0.32	0.38
	DG*	0.25	0.26	0.23	0.24	0.3
	**A	0.14	0.14	0.12	0.12	0.19
	D*A	0.12	0.12	0.1	0.1	0.17
	*GA	0.12	0.13	0.1	0.11	0.17
	DGA	0.11	0.12	0.09	0.1	0.15
20 min	***	0.3	0.3	0.29	0.29	0.3
	D**	0.18	0.18	0.18	0.18	0.19
	G	0.27	0.27	0.27	0.27	0.28
	DG*	0.2	0.2	0.19	0.2	0.2
	**A	0.14	0.14	0.14	0.13	0.14
	D*A	0.13	0.13	0.13	0.13	0.13
	*GA	0.14	0.14	0.14	0.14	0.14
	DGA	0.13	0.13	0.13	0.13	0.13
Panel 2 (Error in mean)						
10 sec	***	−0.19	−0.21	−0.14	−0.16	−0.39
	D**	−0.11	−0.13	−0.06	−0.08	−0.32
	G	−0.2	−0.23	−0.15	−0.18	−0.4
	DG*	−0.14	−0.18	−0.09	−0.13	−0.35
	**A	0.04	0.02	0.1	0.08	−0.19
	D*A	0.05	0.02	0.1	0.08	−0.19
	*GA	0	−0.04	0.06	0.01	−0.22
	DGA	0.01	−0.03	0.07	0.02	−0.22
2 min	***	−0.33	−0.33	−0.31	−0.31	−0.38
	D**	−0.24	−0.24	−0.21	−0.21	−0.29
	G	−0.32	−0.33	−0.29	−0.3	−0.37
	DG*	−0.23	−0.24	−0.2	−0.21	−0.29
	**A	−0.11	−0.11	−0.09	−0.08	−0.17
	D*A	−0.09	−0.09	−0.06	−0.06	−0.15
	*GA	−0.09	−0.11	−0.06	−0.08	−0.15
	DGA	−0.08	−0.09	−0.05	−0.06	−0.13
20 min	***	−0.27	−0.27	−0.27	−0.27	−0.28
	D**	−0.13	−0.13	−0.12	−0.12	−0.14
	G	−0.25	−0.25	−0.24	−0.25	−0.26
	DG*	−0.15	−0.15	−0.15	−0.15	−0.16
	**A	−0.11	−0.1	−0.1	−0.1	−0.11
	D*A	−0.06	−0.06	−0.06	−0.06	−0.07
	*GA	−0.09	−0.1	−0.09	−0.09	−0.1
	DGA	−0.07	−0.07	−0.07	−0.07	−0.08

Note: The column "Filtering" denotes deseasonalization (D), GARCH (G), and ACD (A). $RMSD = \sqrt{\sum_i \left(VaR_{est_i}/VaR_{emp_i} - 1\right)^2/n}$ and $Mean = \sum_i \left(VaR_{est_i}/VaR_{emp_i} - 1\right)/n$.

Table 15: Performance statistics in terms of ES (99 percent), 9:30am-4:00pm.

Time-scale	Filtering	MW-DE	MW-ST	W-DE	W-ST	E-N
Panel 1 (RMSD metric)						
10 sec	***	0.35	0.34	0.32	0.31	0.52
	D**	0.29	0.28	0.26	0.26	0.47
	G	0.31	0.34	0.28	0.31	0.49
	DG*	0.27	0.3	0.24	0.28	0.45
	**A	0.12	0.11	0.09	0.11	0.33
	D*A	0.12	0.11	0.1	0.12	0.32
	*GA	0.12	0.15	0.1	0.15	0.31
	DGA	0.12	0.15	0.11	0.15	0.31
2 min	***	0.44	0.42	0.42	0.4	0.49
	D**	0.35	0.34	0.33	0.32	0.41
	G	0.39	0.4	0.37	0.37	0.44
	DG*	0.31	0.32	0.29	0.3	0.37
	**A	0.22	0.2	0.19	0.17	0.28
	D*A	0.2	0.18	0.17	0.16	0.26
	*GA	0.17	0.18	0.14	0.15	0.23
	DGA	0.15	0.16	0.13	0.14	0.21
20 min	***	0.35	0.34	0.34	0.33	0.35
	D**	0.23	0.23	0.23	0.22	0.24
	G	0.32	0.32	0.32	0.32	0.33
	DG*	0.24	0.24	0.23	0.24	0.24
	**A	0.19	0.18	0.18	0.17	0.19
	D*A	0.18	0.17	0.18	0.17	0.19
	*GA	0.17	0.17	0.17	0.17	0.18
	DGA	0.17	0.17	0.17	0.17	0.18
Panel 2 (Error in mean)						
10 sec	***	−0.32	−0.31	−0.29	−0.28	−0.51
	D**	−0.26	−0.25	−0.22	−0.2	−0.45
	G	−0.28	−0.32	−0.25	−0.28	−0.48
	DG*	−0.23	−0.27	−0.19	−0.23	−0.44
	**A	−0.09	−0.08	−0.05	−0.03	−0.32
	D*A	−0.08	−0.06	−0.03	−0.02	−0.31
	*GA	−0.07	−0.11	−0.02	−0.07	−0.3
	DGA	−0.06	−0.1	−0.01	−0.06	−0.3
2 min	***	−0.42	−0.41	−0.4	−0.38	−0.47
	D**	−0.33	−0.31	−0.31	−0.29	−0.39
	G	−0.37	−0.38	−0.34	−0.35	−0.43
	DG*	−0.29	−0.3	−0.26	−0.27	−0.35
	**A	−0.2	−0.18	−0.18	−0.16	−0.27
	D*A	−0.18	−0.16	−0.15	−0.13	−0.24
	*GA	−0.15	−0.16	−0.11	−0.13	−0.21
	DGA	−0.12	−0.14	−0.09	−0.11	−0.19
20 min	***	−0.32	−0.32	−0.32	−0.31	−0.33
	D**	−0.18	−0.17	−0.17	−0.16	−0.19
	G	−0.3	−0.3	−0.29	−0.29	−0.31
	DG*	−0.18	−0.18	−0.17	−0.17	−0.19
	**A	−0.15	−0.14	−0.14	−0.13	−0.16
	D*A	−0.13	−0.12	−0.12	−0.11	−0.14
	*GA	−0.12	−0.13	−0.12	−0.12	−0.14
	DGA	−0.11	−0.11	−0.1	−0.11	−0.12

Note: The column "Filtering" denotes deseasonalization (D), GARCH (G), and ACD (A).
$RMSD = \sqrt{\sum_i \left(ES_{est_i}/ES_{emp_i} - 1\right)^2 / n}$ and $Mean = \sum_i \left(ES_{est_i}/ES_{emp_i} - 1\right)/n$.

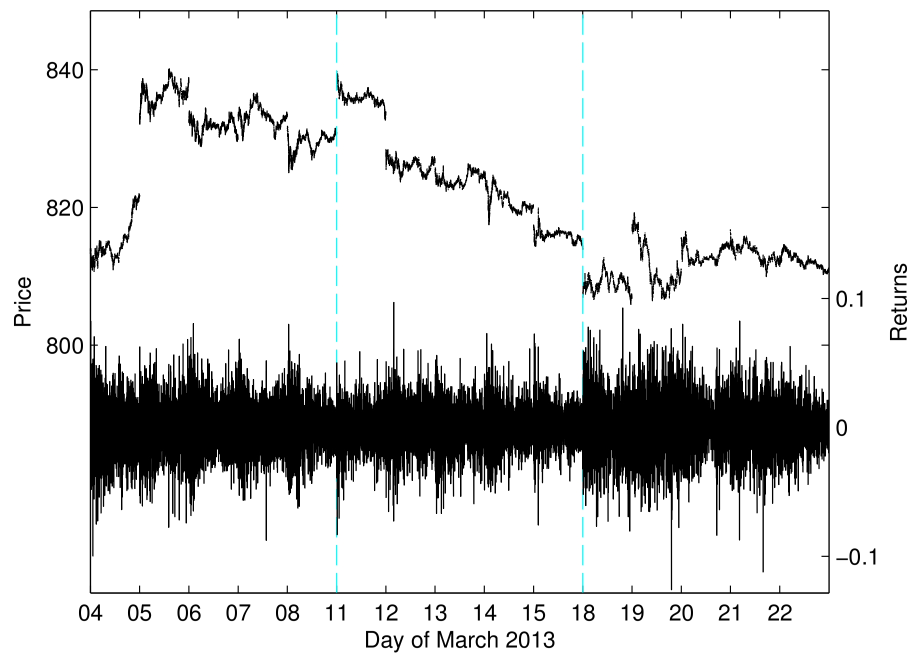

Figure 1: Price and returns for GOOG for Weeks 10, 11, and 12.

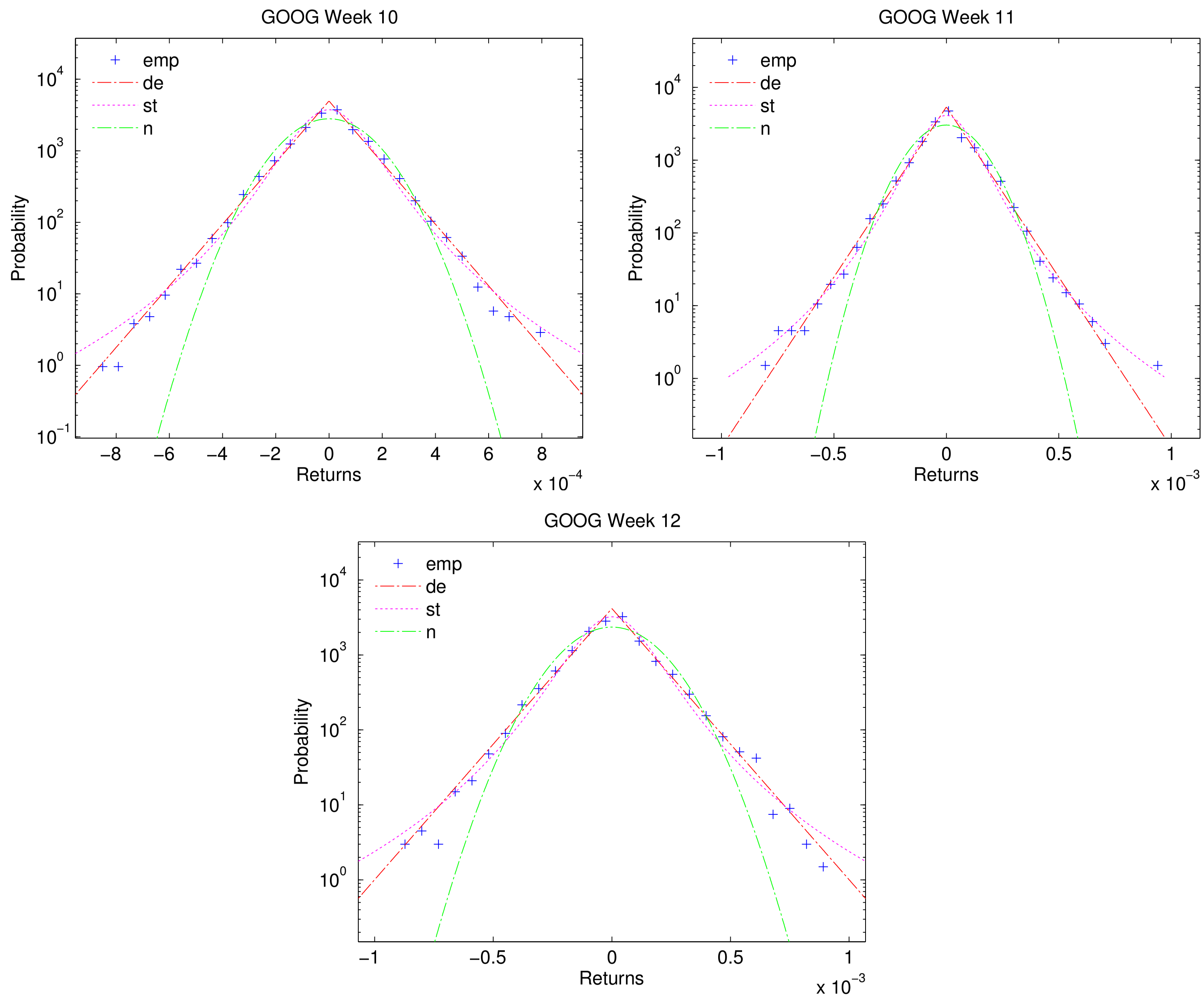

Figure 2: Return histograms with normal, double-exponential, and Student-t distributions for GOOG.

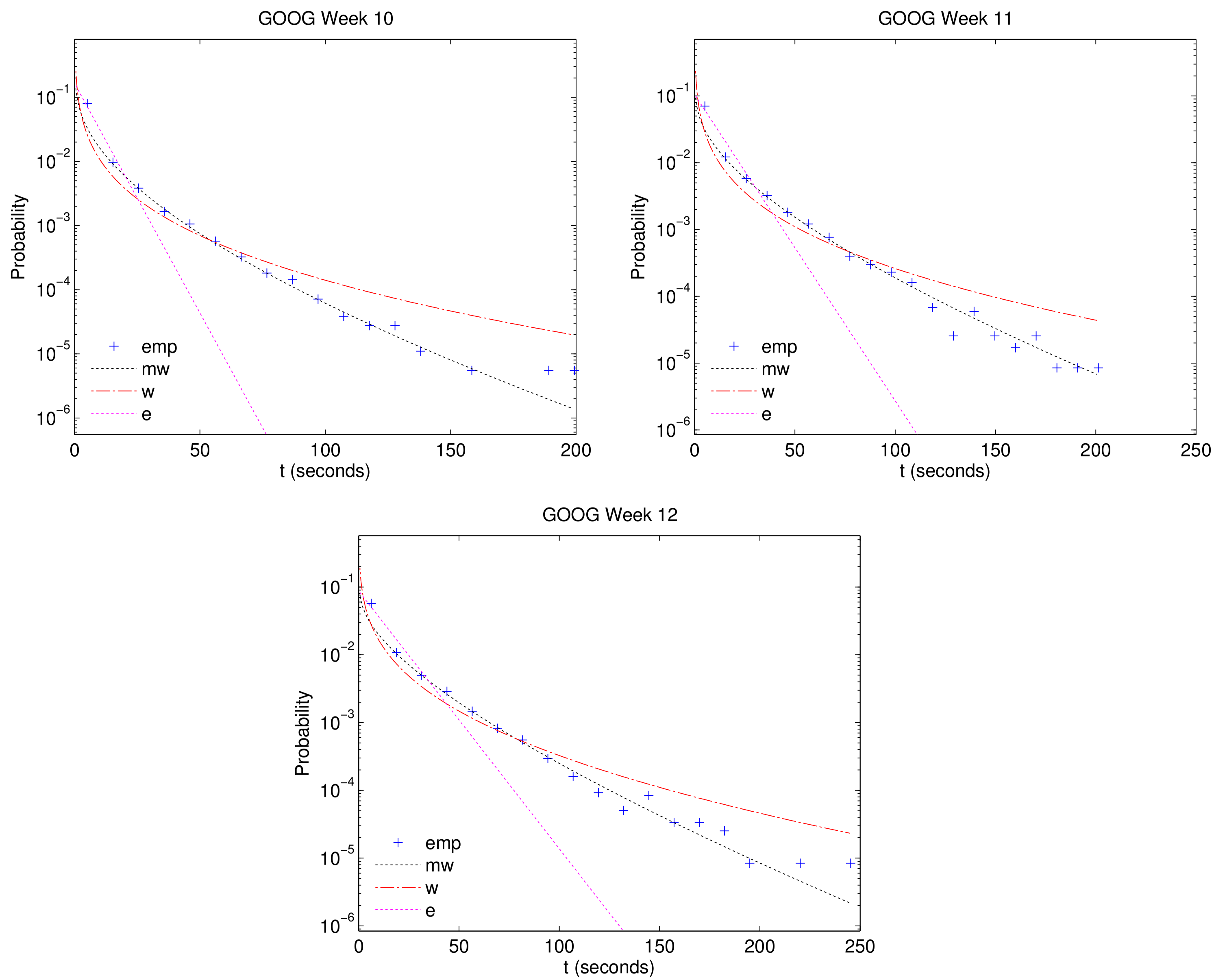

Figure 3: Normalized histograms with fitted probability density functions for GOOG.

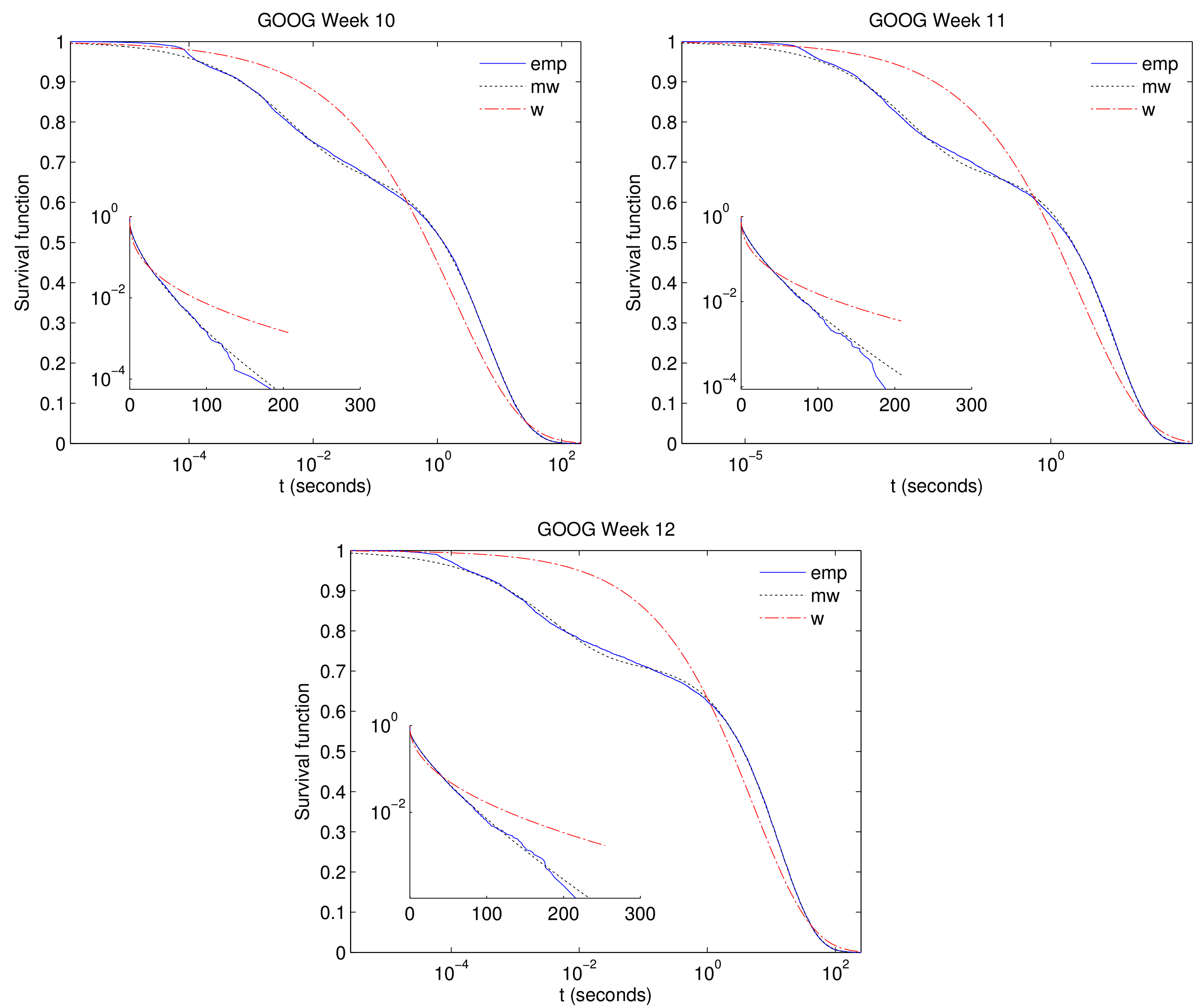

Figure 4: Survival probability for waiting times with Weibull and mixture-Weibull fits for GOOG. The main plot has a logarithmic x-axis while the inset plot has a logarithmic y-axis.

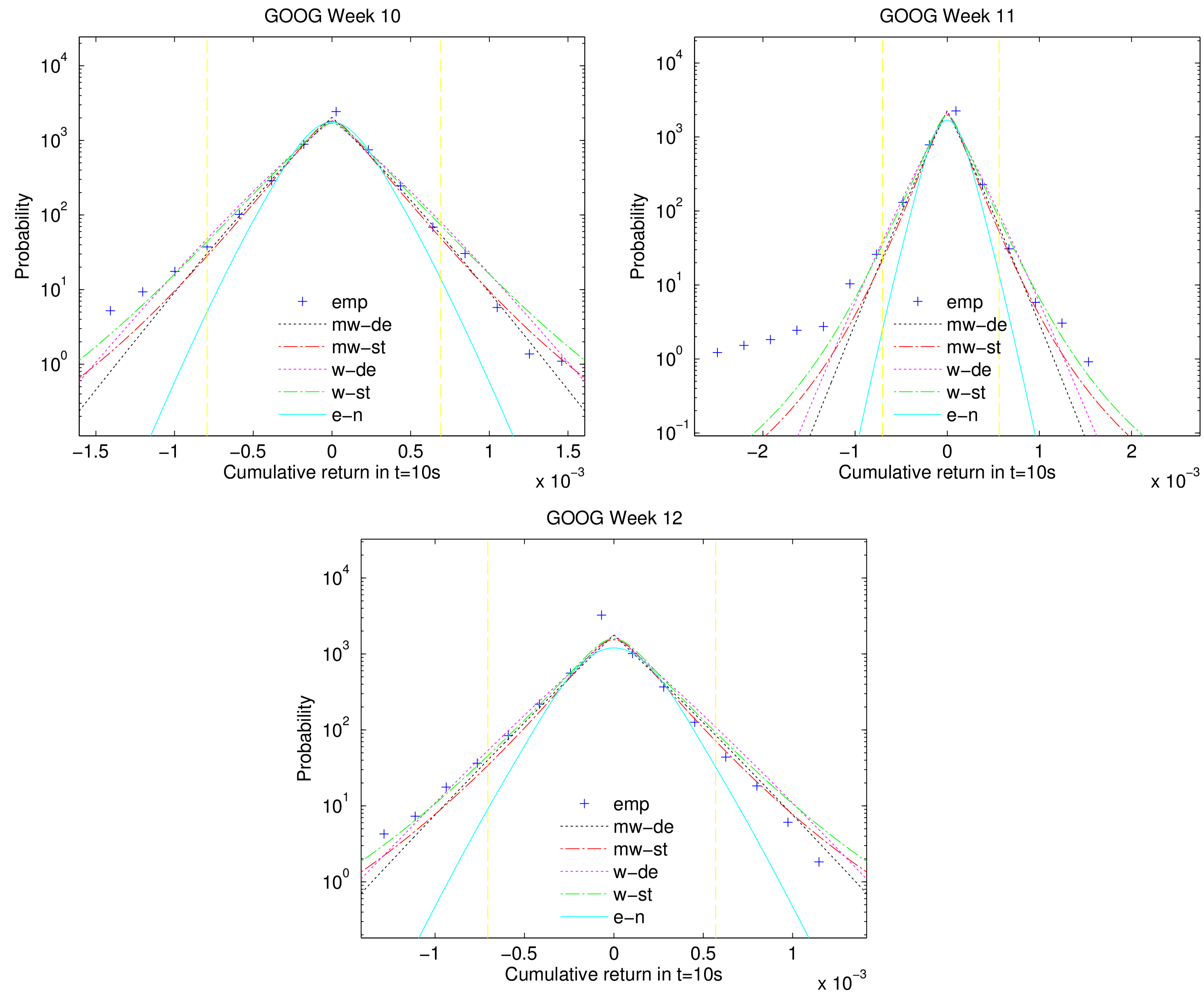

Figure 5: Comparison of the empirical and theoretical results based on p(x,t=10sec) for GOOG. The yellow vertical line corresponds to the first percentile to either negative or positive direction.

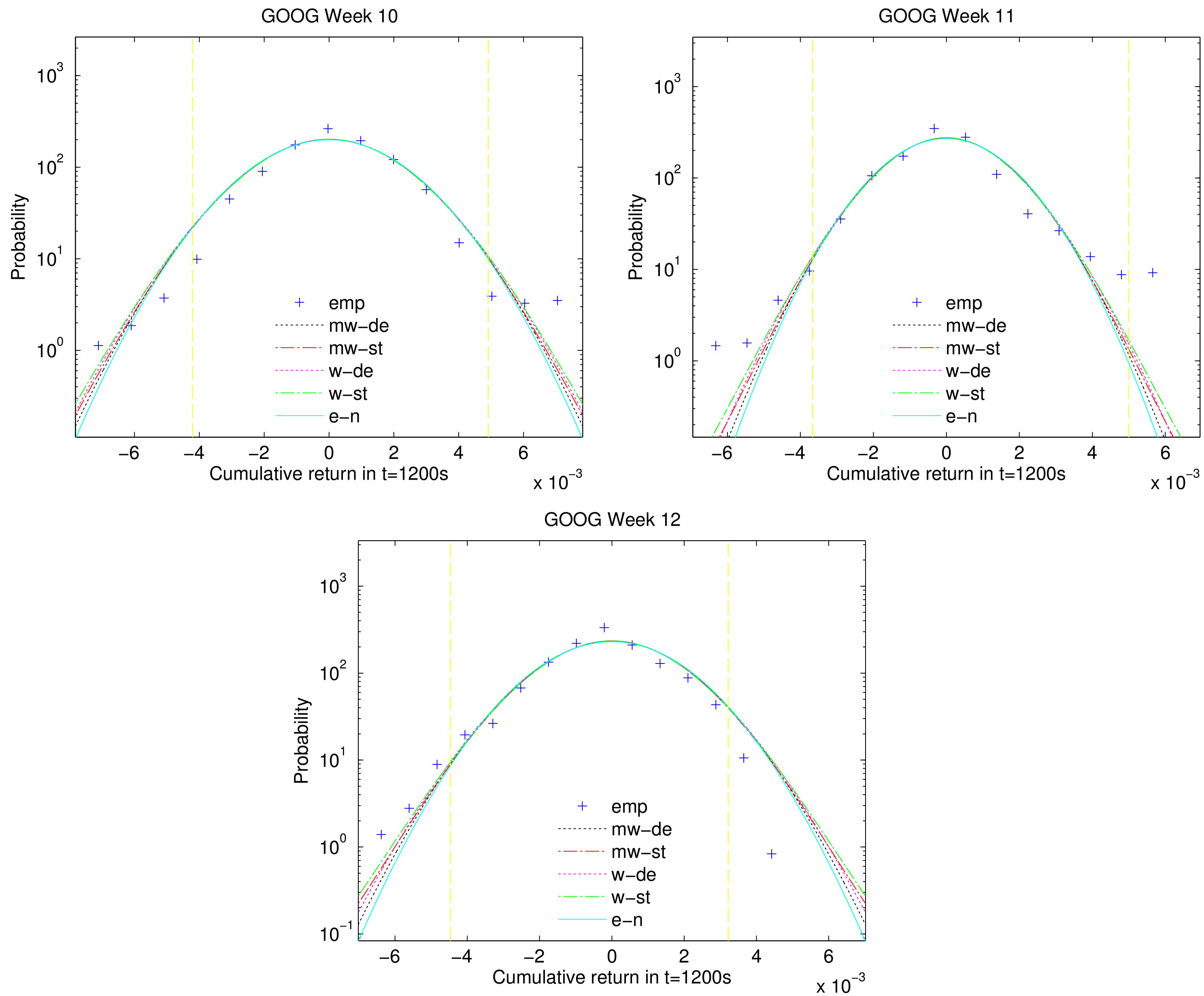

Figure 6: Comparison of the empirical and theoretical results based on p(x,t=20min) for GOOG. The yellow vertical line corresponds to the first percentile to either negative or positive direction.

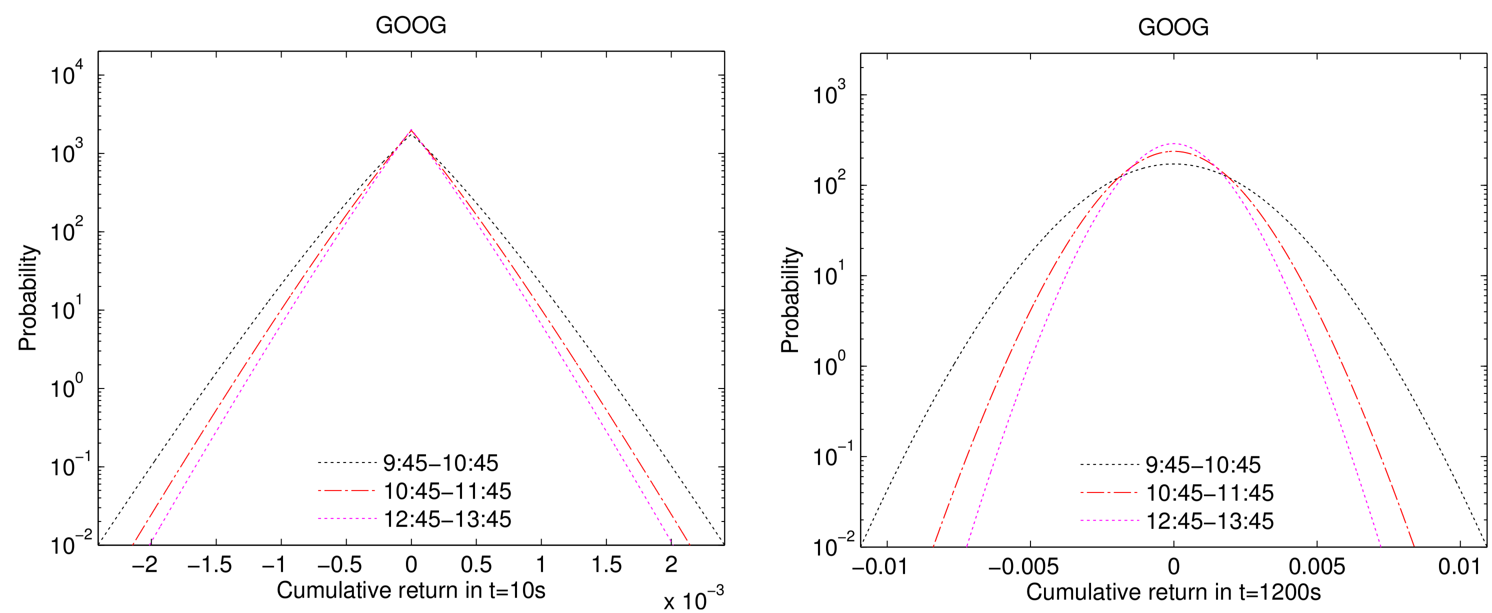

Figure 7: The predicted price distribution at 10sec (left-hand side) and 20min (right-hand side) and three different time-slots.

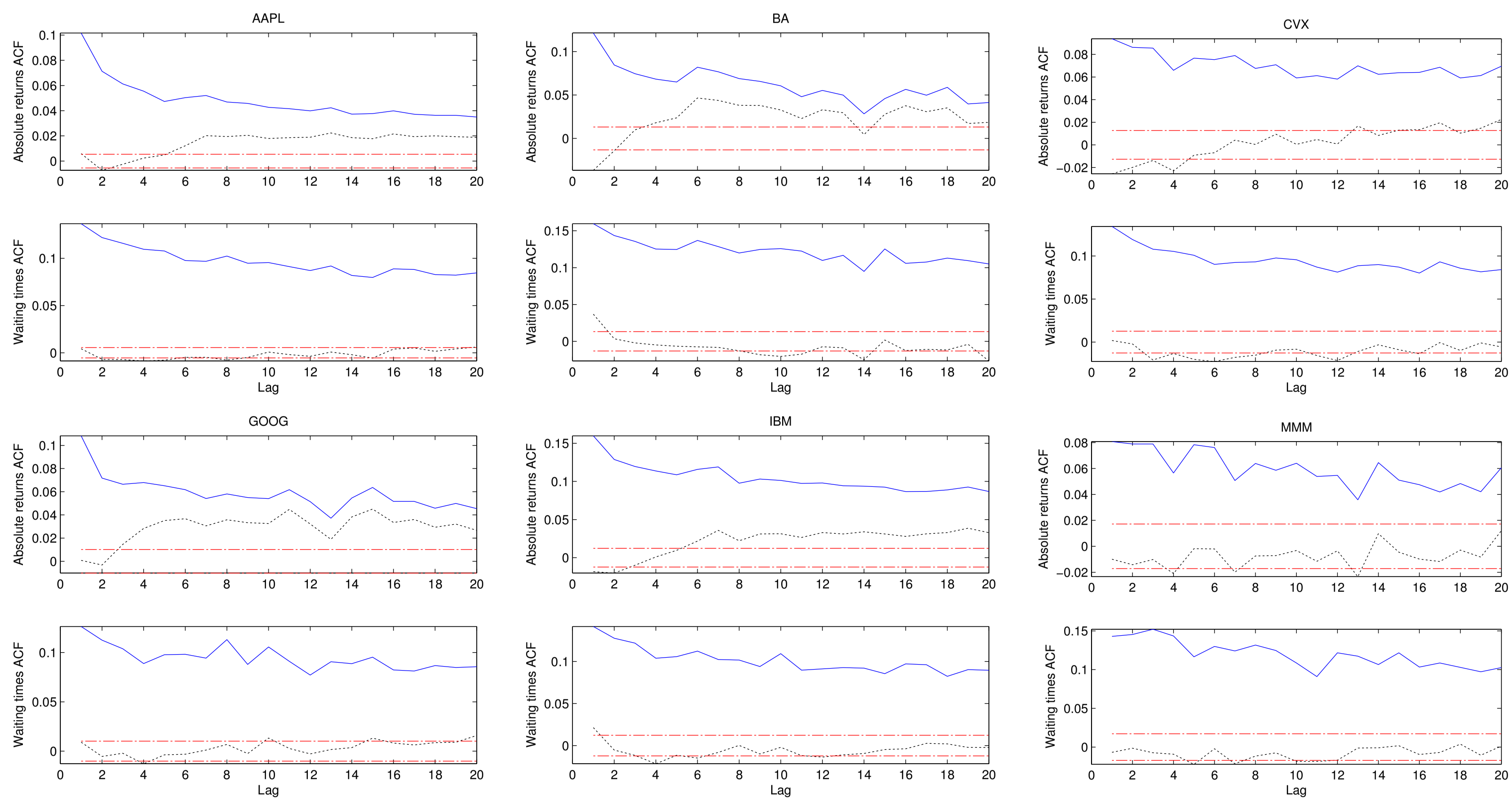

Figure 8: The ACF of absolute returns and waiting times. The blue lines denote data before deseasonalization and GARCH/ACD-filtering. The horizontal confidence bounds are 95 percent for a white noise process.

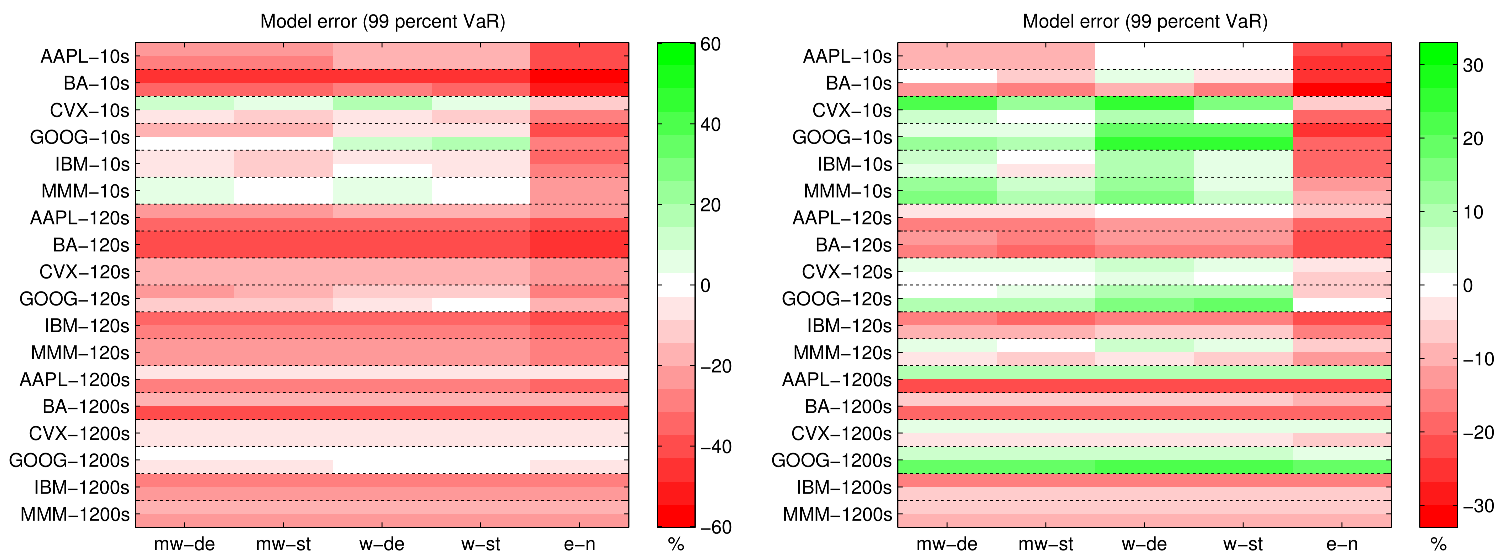

Figure 9: $VaR_{pred}/VaR_{emp} - 1$ in percentages before (left panel) and after (right panel) deseasonalization and GARCH/ACD-filtering. For each stock and time-scale, the upper row is the left tail and lower row the right tail.

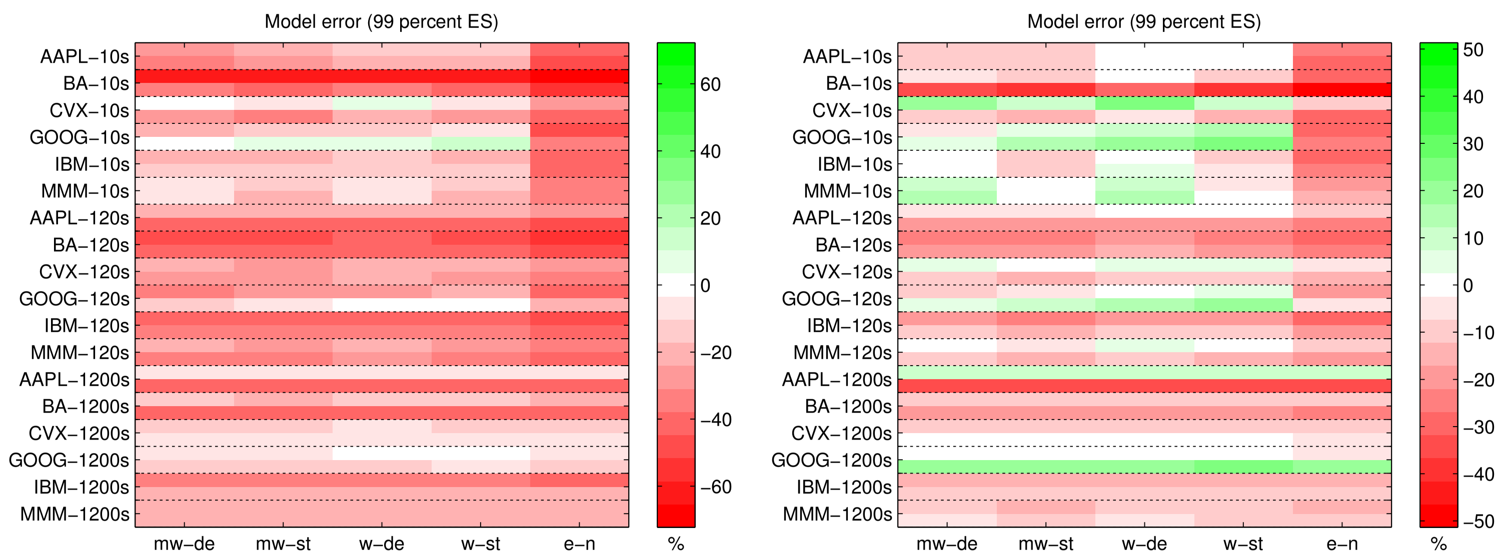

Figure 10: $ES_{pred}/ES_{emp} - 1$ in percentages before (left panel) and after (right panel) deseasonalization and GARCH/ACD-filtering. For each stock and time-scale, the upper row is the left tail and lower row the right tail.

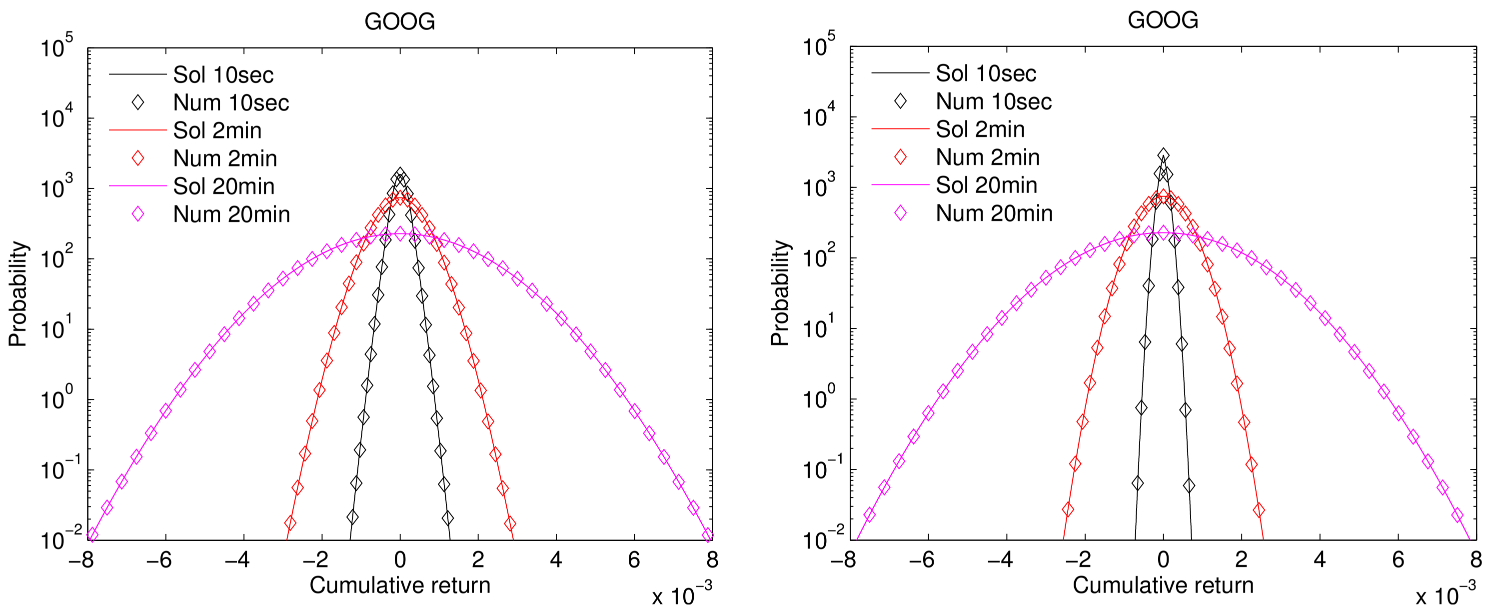

Figure 11: Comparison of analytic and numerical solutions. The parameter estimates used are for GOOG.

References

Barndorff-Nielsen, O. (1977). "Exponentially Decreasing Distributions for the Logarithm of Particle Size," *Proceedings of the Royal Society of London* A **353**, 401–419.

Beran, Jan. (1992). "Statistical Methods for Data with Long-Range Dependence," *Statistical Science* **7**, 404–416.

Bollerslev, Tim. (1987). "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics* **69**, 542–547.

— (1986). "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics* **31**, 307–327.

Dionne, Georges, Pierre Duchesne, and Maria Pacurar. (2009). "Intraday Value at Risk (IVaR) Using Tick-by-Tick Data with Application to the Toronto Stock Exchange," *Journal of Empirical Finance* **16**, 777–792.

Easley, David, and Maureen O'Hara. (1992). "Time and the Process of Security Price Adjustment," *Journal of Finance* **47**, 577–605.

Easley, David, Marcos M. López de Prado, and Maureen O'Hara. (2012). "The Volume Clock: Insights into the High-Frequency Paradigm," *Journal of Portfolio Management* **39**, 19–29.

Eberlein, Ernst. (2010). "Generalized Hyperbolic Models," Encyclopedia of Quantitative Finance, R. Cont (ed.), John Wiley & Sons Ltd., 833–836.

Eberlein, Ernst, and Ernst August v. Hammerstein. (2002). "Generalized Hyperbolic and Inverse Gaussian Distributions: Limiting Cases and Approximation of Processes," in Dalang, Robert C., Marco Dozzi, Francesco Russo (eds), Seminar on Stochastic Analysis, Random Fields and Applications IV, Progress in Probability 58, Birkhäuser, 221–264.

Engle, Robert. F, and Jeffrey R. Russell. (1998). "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica* **66**, 1127–1162.

Jiang, Zhi-Qiang, Wei Chen, and Wei-Xing Zhou. (2008). "Scaling in the Distribution of Intertrade Durations of Chinese Stocks," *Physica* A **387**, 5818–5825.

Mainardi, Francesco, Marco Raberto, Rudolf Gorenflo, and Enrico Scalas. (2000). "Fractional Calculus and Continuous-Time Finance. II: the Waiting-Time Distribution," *Physica* A **287**, 468–481.

Mandelbrot, Benoit. (1963). "The Variation of Certain Speculative Prices," *Journal of Business* **36**, 394–419.

Masoliver, Jaume, Miquel Montero, and George H. Weiss. (2003). "Continuous-Time Random-Walk Model for Financial Distributions," *Physical Review* E **67**, 021112.

Masoliver, Jaume, Miquel Montero, Josep Perelló, and George H. Weiss. (2006). "The Continuous Time Random Walk Formalism in Financial Markets," *Journal of Economic Behavior Organization* **61**, 577–598.

Montroll, Elliott W., and George H. Weiss. (1965). "Random Walks on Lattices II," *Journal of Mathematical Physics* **6**, 167–181.

Raible, Sebastian. (2000). "Lévy Processes in Finance: Theory, Numerics, and Empirical Facts," Ph.D. dissertation, University of Freiburg.

Scalas, Enrico, Rudolf Gorenflo and Francesco Mainardi. (2004). "Uncoupled Continuous-Time Random Walks: Solution and Limiting Behavior of the Master Equation," *Physical Review* E **69**, 011107.

— (2000). "Fractional Calculus and Continuous-Time Finance," *Physica* A **284**, 376-384.

Teräsvirta, Timo. (2009). "An Introduction to Univariate GARCH Models," in T. G. Andersen, R. A. Davis, J.-P. Kreiβ, and T. Mikosch (eds.), Handbook of Financial Time Series. Springer.

Valsa, Juraj, and Lubomir Brancik. (1998). "Approximate Formulae for Numerical Inversion of Laplace Transforms," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* **11**, 153–166.

Vuorenmaa, Tommi A. (2013). "The Good, the Bad, and the Ugly of Automated High-Frequency Trading," *Journal of Trading* **8**, 58–74.

— (2012). "Liquidity, Activity, and Dependence on Interlinked Trading Venues," Working Paper, Valo Research and Trading.

— (2009). "A q-Weibull Autoregressive Conditional Duration Model with an Application to NYSE and HSE Data," Working Paper, University of Helsinki.