More accurate volatility estimation and forecasts using price durations*

Ingmar Nolte[†] Stephen J. Taylor[‡] Xiaolu Zhao[§]

This version: July, 2016

Abstract

We investigate price duration based variance estimators that have long been ignored in the literature. We show i) how price duration based estimators can be used for the estimation and forecasting of the integrated variance of an underlying semi-martingale price process and ii) how they are affected by a) important market microstructure noise components such as the bid/ask spread, irregularly spaced observations in discrete time and discrete price levels, as well as b) price jumps. We develop i) a simple-to-construct non-parametric estimator and ii) a parametric price duration estimator using autoregressive conditional duration specifications. We provide guidance how these estimators can best be implemented in practice by optimally selecting a threshold parameter that defines a price duration event. We provide simulation evidence that price duration estimators give lower RMSE's than competing estimators and forecasting evidence that they extract relevant information from high-frequency data better and produce more accurate forecasts than competing realized volatility and option-implied variance estimators, when considered in isolation or as part of a forecasting combination setting.

Keywords: Price durations; Volatility estimation; High-frequency data; Market microstructure noise components; Forecasting.

---

*We would like to thank Torben Andersen, Sandra Nolte and all participants of the faculty seminar at Reading, the Konstanz-Lancaster 2015 and the ESRC-NWDTC 2015 workshops, the SoFiE 2016 and IAAE 2016 conferences for helpful comments.

[†]Department of Accounting and Finance, Lancaster University Management School, Lancaster University, UK. Email: i.nolte@lancaster.ac.uk.

[‡]Department of Accounting and Finance, Lancaster University Management School, Lancaster University, UK. Email: s.taylor@lancaster.ac.uk.

[§]Corresponding author. Department of Accounting and Finance, Lancaster University Management School, Lancaster University, UK. Email: x.zhao@lancaster.ac.uk

1 Introduction

Precise volatility estimates are indispensable for many applications in finance. We focus on price duration based variance estimators, that in contrast to GARCH, realized volatility (RV) type and option-implied variance estimators have received very little attention in the literature so far. We show how price duration based estimators can be used to estimate and forecast the integrated variation (IV) of an underlying semi-martingale process. We investigate how market microstructure noise effects, such as the bid/ask spread, irregularly spaced price observations and price discreteness, and also price jumps, affect, individually and jointly, price duration based integrated variance estimators in terms of bias and efficiency.

Within the class of price duration variance estimators we develop i) a simple-to-construct non-parametric estimator and ii) a parametric price duration estimator on the basis of dynamic autoregressive conditional duration (ACD) specifications. We show how these estimators can be robustified against market microstructure noise (MMS) influences by optimally choosing the threshold parameter that determines the size of the price change which defines a price duration event. Through simulation evidence, we show that the price duration estimators produce lower RMSE's. Within a forecasting setup we provide evidence that price duration variance estimators extract relevant information from (high-frequency) data better, and produce more accurate variance forecasts, than competing RV-type and option-implied variance estimators, when considered either in isolation or as part of a forecasting combination.

Over the last decade RV-type quadratic variation estimators[1] following Andersen, Bollerslev, Diebold, and Ebens (2001) and Barndorff-Nielsen and Shephard (2002) have become the standard tool for the construction of daily variance estimators by exploiting intra-day high-frequency data. In the presence of MMS noise three main approaches for the estimation of the integrated variance exist. The sub-sampling method of Zhang, Mykland, and Ait-Sahalia (2005) and Ait-Sahalia, Mykland, and Zhang (2011) combines RV estimators computed on different return sampling frequencies and gives rise to the two-scale and multi-scale realized variance estimators. The Least Squares based IV estimation framework of Nolte and Voev (2012) is related to this and allows for the joint estimation of IV and the moments of market noise. Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) develop the class of realized kernel estimators and Podolskij and Vetter (2009) and Jacod, Li, Mykland, Podolskij, and Vetter (2009) introduce the pre-averaging based IV estimators.

Essentially RV-type variance estimators are based on the idea of aggregating, over a daily horizon, say, squared (log-) price changes computed on *fixed* intra-day *intervals*, say

[1] In the absence of price jumps we simply refer to integrated variance estimators.

of 2 minutes. Hence they impose structure on the time-dimension, but keep the outcomes in the price domain flexible. Price duration based variance estimators are based on the opposite consideration: here structure is imposed on the price domain by *fixing* the price change *size*, but allowing the time to generate such price changes (price durations) to vary. From an information point of view, price durations condition on the complete history of the price process after a previous price event, while RV-type estimators can be and actually are constructed from a sparser information set that only requires knowledge of the prices at the start and end of an interval. It is precisely this potential information advantage that makes price duration based variance estimators attractive and it is surprising that over the last two decades only a handful of studies analysed them in any depth. A notable but neglected working paper by Andersen, Dobrev, and Schaumburg (2008) provides analytic results for diffusion processes which shows that duration estimators are much more efficient than RV estimators. A further attractive feature of price duration based variance estimation is that in its parametric form, i.e. with a parametric form assumption for the dynamic price duration process, not only an integrated variance estimator but also a local (intra-day, spot) variance estimator can be obtained.

After Cho and Frees (1988) the next reference introducing price duration variance estimators is Engle and Russell (1998), which includes ACD specifications. Gerhard and Hautsch (2002) and more recently Tse and Yang (2012) also develop price duration based variance estimators using ACD specifications to govern the price duration dynamics. All three ACD studies start from a point process concept to construct volatility estimators, but do not relate the estimators to a desirable underlying theoretical concept such as the integrated variation of a Brownian semi-martingale process. These studies also provide little guidance on the practical task of selecting a good price change threshold when MMS noise effects are present, which is important for implementation. Our study fills these gaps.

The derivation of duration based volatility estimators in this paper is initially done in a pure diffusion setting. Following Engle and Russell (1998) and Tse and Yang (2012), we approximate the integrated variance of the diffusion process by that of a step process, whose conditional instantaneous variance can be related to the conditional intensity function of the price duration. The integral of the instantaneous variance of this step process provides an estimate of IV and the estimation error goes to zero as the threshold size approaches zero. We then consider the effect that transaction prices are either bid or ask prices and rely on Monte Carlo evidence to analyse the joint influence of bid/ask spreads, irregularly spaced discrete trading times and discrete price levels, as well as price jumps, upon our duration based integrated variance estimators. We find, on the basis of both simulations and empirical evidence, that the existence of bid and ask prices biases the duration based variance estimates upwards while discrete time transactions yield downward biases. Both effects diminish for a large enough and increasing price change threshold parameter. Other sources of biases are end of day effects, discrete prices and potential jumps. Their magnitudes are quantified either theoretically or through Monte Carlo evidence. It is noteworthy that price duration variance estimators possess by construction some robustness regarding large price jump events.

To compare the accuracy and the information content of price duration based estimators with estimators from RV and also option-implied classes, we conduct a comprehensive forecasting study. We perform both individual and combination forecasts, on 20 DJIA stocks over 11 years from 2002 to 2012, over three horizons, one day, one week, and one month. We find that the duration based class of variance estimators generally perform better than RV type and option-implied estimators. The parametric price duration estimators, in isolation, yield more accurate forecasts than their non-parametric counterparts and all other estimators (RV and option-implied type) over all three horizons. However, no individual estimator alone seems to subsume all relevant information and combining forecasts from the three considered classes of estimators significantly improves the forecast accuracy. Our findings confirm the theoretical prediction of Andersen, Dobrev, and Schaumburg (2008) that duration based variance estimators contain more relevant information than RV-type estimators. Our results also contribute to the debate in the volatility forecasting literature about the accuracy of high-frequency estimators relative to option-implied estimators. While Blair, Poon, and Taylor (2001), Jiang and Tian (2005), Giot and Laurent (2007), and Busch, Christensen, and Nielsen (2011) find that option-implied estimators provide the most accurate volatility forecasts for stock indices, the opposite conclusion, favouring high-frequency estimators is supported in Bali and Weinbaum (2007), Becker, Clements, and White (2007) and Martin, Reidy, and Wright (2009). Our univariate forecasts provide clear evidence that high-frequency estimators (of which duration based estimators are best) are more accurate than option-implied alternatives for our sample period and our sample of 20 DJIA stocks.

The rest of the paper is organized in the following way: Section 2 lays out the theoretical foundations for the duration based integrated variance estimators and includes a theoretical discussion on market microstructure noise components. Section 3 describes the high-frequency data used subsequently and provides descriptive results that motivate the simulation study. Section 4 contains the simulation study that assesses the effects of market microstructure noise components on our duration based integrated variance estimators, provides guidance on the choice of a preferred price change threshold value, and compares the accuracy and efficiency of the duration based estimator with competing estimators. Section 5 contains the empirical analysis of our estimators including a discussion on the construction of the parametric duration based integrated variance estimators and empirical evidence on the choice of a preferred price change threshold value. Section 6 contains the forecasting study and Section 7 concludes.

2 Theoretical Foundations

In Section 2.1 we provide the theoretical foundations for parametric and non-parametric duration based integrated variance estimators in a pure diffusion setting in the absence of MMS noise. Section 2.2 provides theoretical results for duration based integrated variance estimators in the presence of bid and ask transaction prices and price jumps. The analysis of further market microstructure noise components and their interplay is deferred to the simulation study in Section 4.

2.1 Duration based integrated variance estimators: pure diffusion setting

Initially we assume that the efficient log-price, X_t, follows a pure diffusion process with no drift, represented by

$$dX_t = \sigma_{X,t} dB_t. \tag{1}$$

For each trading day and a selected threshold δ, a set of event times $\{t_d, d = 0, 1, ...\}$ is defined in terms of absolute cumulative price changes exceeding δ, by $t_0 = 0$ and

$$t_d = \inf_{t>t_{d-1}} \{|X_t - X_{t_{d-1}}| = \delta\}, \quad d \geq 1. \tag{2}$$

Let $x_d = t_d - t_{d-1}$ denote the time duration between consecutive events and let $\mathcal{I}_{d-1}$ denote the complete price history up to time t_{d-1}. For the conditional distribution $x_d|\mathcal{I}_{d-1}$, we denote the density function by $f(x_d|\mathcal{I}_{d-1})$, the cumulative density function by $F(x_d|\mathcal{I}_{d-1})$ and the intensity (or hazard) function by $\lambda(x_d|\mathcal{I}_{d-1}) = f(x_d|\mathcal{I}_{d-1})/(1 - F(x_d|\mathcal{I}_{d-1}))$.

Following Engle and Russell (1998) and Tse and Yang (2012), duration based variance estimators rely on a relationship between the conditional intensity function and the conditional instantaneous variance of a step process. The step process $\{\tilde{X}_t, t \geq 0\}$ is defined by $\tilde{X}_t = X_t$ when $t \in \{t_d, d \geq 0\}$ and by $\tilde{X}_t = \tilde{X}_{t_{d-1}}$ whenever $t_{d-1} < t < t_d$. The conditional instantaneous variance of $\tilde{X}_t$ equals

$$\sigma^2_{\tilde{X},t} = \lim_{\Delta \to 0} \frac{1}{\Delta} \operatorname{var}(\tilde{X}_{t+\Delta} - \tilde{X}_t | \mathcal{I}_{d-1}), \quad t_{d-1} < t < t_d. \tag{3}$$

As Δ approaches zero we may ignore the possibility of two or more events between times t and $t + \Delta$, so that the only possible outcomes for $\tilde{X}_{t+\Delta} - \tilde{X}_t$ can be assumed to be 0, δ and $-\delta$. The probability of a non-zero outcome is determined by $\lambda(x|G_{d-1})$ and consequently

$$\sigma^2_{\tilde{X},t} = \delta^2 \lambda(t - t_{d-1}|\mathcal{I}_{d-1}), \quad t_{d-1} < t < t_d. \tag{4}$$

The integral of $\sigma^2_{\tilde{X},t}$ over a fixed time interval provides an approximation to the integral of $\sigma^2_{X,t}$ over the same time interval, and the approximation error disappears as $\delta \to 0$.

Let there be N price duration times during a day, then the general duration based estimator of integrated variance, IV, is given by

$$\widetilde{IV} = \int_{0}^{t_N} \sigma^2_{\tilde{X},t} dt = \sum_{d=1}^{N} \delta^2 \int_{t_{d-1}}^{t_d} \lambda(t - t_{d-1}|\mathcal{I}_{d-1}) dt$$
$$= -\delta^2 \sum_{d=1}^{N} \ln(1 - F(x_d|\mathcal{I}_{d-1})). \qquad (5)$$

The above estimator ignores price variation between the last price event of the day at time t_N and the end of the day, t_{eod}, which is expected to be of minor importance when δ is relatively small. A natural bias corrected general duration based integrated variance estimator is therefore

$$\widetilde{IV}_+ = -\delta^2 \sum_{d=1}^{N} \ln(1 - F(x_d|\mathcal{I}_{d-1})) + \delta^2 \int_{t_N}^{t_{eod}} \lambda(t - t_N|\mathcal{I}_N) dt. \qquad (6)$$

In practice, we do not know the true intensity function. We must therefore either estimate the functions $\lambda(.|.)$ or we can replace the integrals in (5) by their expectations. As these expectations are always one, the non-parametric, duration based variance estimator, $NPDV$, is simply

$$NPDV = N\delta^2. \qquad (7)$$

This equals the quadratic variation of the approximating step process over a single day, which we may hope is a good estimate of the quadratic variation of the diffusion process over the same time interval. An equation like (7), for the special case of constant volatility, can be found in the early investigation of duration based methods by Cho and Frees (1988). Relying on this setup and for N large it is immediately clear that the downward bias introduced by ignoring end of day effects is equal to $0.5\delta^2$, as in expectation we omit (counting) half an event at the end of the day. The bias corrected non-parametric estimator is therefore given by

$$NPDV_+ = (N + 0.5)\delta^2. \qquad (8)$$

A parametric implementation of (5) requires selection of appropriate hazard functions $\lambda(.|.)$. As first suggested by Engle and Russell (1998), we assume the durations $x_d = t_d - t_{d-1}$ have conditional expectations ψ_d determined by $\mathcal{I}_{d-1}$ and that scaled durations are independent variables. More precisely,

$$x_d = \psi_d \varepsilon_d, \text{ with } \psi_d = E[x_d|\mathcal{I}_{d-1}], \qquad (9)$$

and the scaled durations ε_d are i.i.d., positive random variables which are stochastically independent of the expected durations ψ_d.

Autoregressive specifications for ψ_d are standard choices, such as the autoregressive conditional duration (ACD) model of Engle and Russell (1998), the logarithmic ACD model of Bauwens and Giot (2000), the augmented ACD model of Fernandes and Grammig (2006) and others reviewed by Pacurar (2008). These specifications do not accommodate the long-range dependence present in our durations data. As a practical alternative to the fractionally integrated ACD model of Jasiak (1999), we develop the heterogenous autoregressive conditional duration (HACD) model in the spirit of the HAR model for volatility introduced by Corsi (2009). Short, medium and long range effects are arbitrarily associated with 1, 5 and 20 durations, and our HACD specification is then

$$\psi_d = \omega + \alpha x_{d-1} + \beta_1 \psi_{d-1} + \beta_2(\psi_{d-5} + \ldots + \psi_{d-1}) + \beta_3(\psi_{d-20} + \ldots + \psi_{d-1}). \tag{10}$$

A flexible shape for the hazard function can be obtained by assuming the scaled durations have a Burr distribution, as in Grammig and Maurer (2000) and Bauwens, Giot, Grammig, and Veredas (2004). The general Burr density and cumulative density functions, as parameterised by Lancaster (1997) and Hautsch (2004), are given by

$$f(y|\xi,\eta,\gamma) = \frac{\gamma}{\xi}(\frac{y}{\xi})^{\gamma-1}[1+\eta(y/\xi)^{\gamma}]^{-(1+(1/\eta))}, \quad y > 0, \tag{11}$$

and

$$F(y|\xi,\eta,\gamma) = 1 - [1+\eta(y/\xi)^{\gamma}]^{-1/\eta}, \quad y > 0, \tag{12}$$

with three positive parameters (ξ, η, γ). The Weibull special case is obtained when $\eta \to 0$ and its special case of an exponential distribution is given by also requiring $\gamma = 1$. The mean μ of the general Burr distribution is

$$\mu = \xi c(\eta,\gamma), \text{ with } c(\eta,\gamma) = B(1+\gamma^{-1}, \eta^{-1} - \gamma^{-1})/\eta^{1+(1/\gamma)}, \tag{13}$$

with $B(.,.)$ denoting the Beta function. For each scaled duration the mean is 1 so that ξ is replaced by $1/c(\eta,\gamma)$. For each duration x_d (having conditional mean ψ_d) we replace ξ by $\psi_d/c(\eta,\gamma)$. From (5) our parametric, duration based variance estimator, PDV, is therefore

$$PDV = \frac{\delta^2}{\eta}\sum_{d=1}^{N} \ln\left(1 + \eta\left[c(\eta,\gamma)\frac{x_d}{\psi_d}\right]^{\gamma}\right). \tag{14}$$

When we implement (14), we take account of the intraday pattern in the durations data. The duration x_{d-1} in (10) is replaced by the scaled quantity $x^*_{d-1} = x_{d-1}/s_{d-1}$ and each expected duration $\psi_{d-\tau}$ is replaced by the scaled quantity $\psi^*_{d-\tau} = \psi_{d-\tau}/s_{d-\tau}$, with $s_{d-\tau}$ the estimated average time between events at the time-of-day corresponding to duration $d-\tau$; each term $s_{d-\tau}$ is obtained from a Nadaraya-Watson kernel regression of price durations against time-of-day using one month of durations data. Then ψ_d is replaced by s_d/ψ^*_d, so the scaled duration x_d/ψ_d in (14) is simply x^*_d/ψ^*_d. End of day bias correction is obtained by adding $0.5\delta^2$ as above.

The theoretical framework above is for the logarithms of prices. It is much easier to set the threshold to be a dollar quantity related to the magnitude of the bid/ask spread. We then replace the log-price X_t in (2) by the price $P_t = \exp(X_t)$. As a small change δ in the price is equivalent to a change δ/P_t in the log-price, we redefine the estimators (including end of day bias correction) to be

$$NPDV_{+} = \delta^2 \sum_{d=1}^{N} 1/P_{d-1}^2 + 0.5\delta^2/P_N \tag{15}$$

and

$$PDV_{+} = \frac{\delta^2}{\eta} \sum_{d=1}^{N} \ln\left(1 + \eta \left[c(\eta,\gamma)\frac{x_d}{\psi_d}\right]^{\gamma}\right)/P_{d-1}^2 + 0.5\delta^2/P_N. \tag{16}$$

While the non-parametric estimator can easily be constructed with a reasonable number of events N, for example during a day, the additional parametric form assumption of the parametric estimator also guarantees a volatility estimator for small N and yields for example a local (intraday) volatility estimator for the case when $N = 1$.

2.2 Market microstructure noise

We first consider how the bid/ask spread, which is arguably the most important market microstructure noise component for transaction price datasets, affects our duration based volatility estimators. In particular, assume that at general times t we observe a noisy price

$$Y_t = X_t + 0.5\mathbb{1}_t\varsigma, \tag{17}$$

where ς denotes the size of the bid/ask spread. X_t is the unobserved true price and $\mathbb{1}_t$ is an indicator variable which equals 1 when Y_t represents an ask price and -1 when Y_t represents a bid price. We assume that ς is constant throughout the day and that Y_t takes prices on the bid or the ask side with equal probability 0.5. A price event occurs when

$$|Y_{t_d} - Y_{t_{d-1}}| = \left|(X_{t_d} - X_{t_{d-1}}) + 0.5(\mathbb{1}_{t_d} - \mathbb{1}_{t_{d-1}})\varsigma\right| \geq \delta, \tag{18}$$

and can be triggered by either the unobserved efficient price change component $(X_{t_d} - X_{t_{d-1}})$ or the bid/ask spread component $0.5(\mathbb{1}_{t_d} - \mathbb{1}_{t_{d-1}})$. The bid/ask spread component can take on three values, -1, 0 and 1, which together with an upward (downward) move of the diffusion component constitutes three possible scenarios:

1) A value of 0 corresponds to the case when both the first price and the last price of the price duration lie on the same side of the limit order book, i.e. bid-bid or ask-ask. In both cases the diffusion component alone has to change by δ to trigger a price event which is equivalent to the case in which we observe no noise.

2) A value of 1 (-1), i.e. bid-ask (ask-bid), together with an upward (downward) moving

diffusion component implies that the diffusion component only has to increase (decrease) by $\delta - \varsigma$ (assuming $\delta > \varsigma$)[2] to trigger a price event, which is on average less than in the no noise case (when $\delta \to 0$). Hence, we observe more of these price events within a day than in the no noise case which contributes to an upward biased variance estimator.

3) A value of -1 (1), i.e. ask-bid (bid-ask), together with an upward (downward) moving diffusion component implies that the diffusion component now has to increase (decrease) by $\delta + \varsigma$ to trigger a price event, which is on average more than in the no noise case. Hence, we observe less of these price events within a day, than in the no noise case which contributes to a downward biased variance estimator.

Scenario 2) is more likely to occur than scenario 3) and hence the bid/ask spread component creates on balance a positively biased duration based volatility estimator. For an explanation let us consider only the upward move case: For a given δ it is more likely that a price duration is closed with an ask price (scenario 2) than a bid price (scenario 3), as once the efficient price has entered into the $\varsigma/2$ distance window below the δ threshold any *transaction* price on the ask side (but not the bid side) will immediately trigger a price event, while triggering the event by a bid price would require the efficient price to pass the corresponding $\varsigma/2$ distance window above the δ threshold.

A larger spread level ς will lead to a wider ς window around the δ price change threshold and hence further increase the positive bias, while the selection of a large enough threshold δ for a given spread level will reduce the bias.

Note that the explanation above makes the implicit assumption that bid and ask transaction prices can occur anywhere within the ς window, which is guaranteed not only under the assumption of bid and ask prices being observed in continuous time, but also under the assumption of irregularly spaced observed bid and ask transaction prices. In the case of irregularly spaced observed transaction prices a further time discretization noise component needs to be addressed. We delegate this consideration to the simulation study in Section 4.

Let us further consider jumps with a jump size of κ and consider the case when a jump occurs

$$|Y_t - Y_{t-1}| = |(X_t - X_{t-1}) + 0.5(\mathbb{1}_t - \mathbb{1}_{t-1})\varsigma + \kappa| \,. \tag{19}$$

As we expect $\kappa \gg \delta$, a price jump would most likely trigger an immediate price event. Yet its impact on the integrated variance estimator is mitigated as κ would be truncated by δ. In addition, as the occurrence of large jumps are rare, we expect them to have very limited influence on the duration based variance estimator.

In the simulation study in Section 4, we further evaluate the performance of our duration based variance estimators under different market microstructure noise scenarios. To obtain

[2] In practice δ will always be chosen to be larger than ς. We discuss the case $\delta < \varsigma$ in the context of the simulation study in Section 4.

some representative input parameters for this study we first carry out a descriptive analysis of our high-frequency data.

3 Data

In the empirical analysis we use 20 of the 30 stocks of the Dow Jones Industrial Average (DJIA) index. The tick-by-tick trades and quotes data spanning 11 years (2769 tading days) from January 2002 to December 2012 are obtained from the New York Stock Exchange (NYSE) TAQ database and are time-stamped to a second. The stocks selected have their primary listing at NYSE without interruption during the sample period.[3] The raw data is cleaned using the method of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009). For our analysis we merge the individual trades and quotes files using a refined Lee and Ready algorithm as outlined in Nolte (2008) to identify trades with corresponding bid and ask quotes, which yields associated buy and sell indicators as well as bid/ask spreads.

The list of stocks and descriptive statistics for the whole sample period are presented in Table 1. Table 1 shows means and medians for bid/ask spreads and inter-trade durations, as well as means for the price levels and volatilities for all stocks, sorted in the ascending order of their mean spread level in the first column. The mean values of bid/ask spreads range from 1.4 to 3.5 cents, and from 3.55 to 7.01 seconds for trade durations. The corresponding medians range from 1 to 2 cents, and 2 to 3 seconds, respectively, implying right-skewed distributions for both variables. Table 1 also presents means and medians for a simple measure of a jump frequency. A jump is recorded when the absolute value of a price change exceeds five times the average bid/ask spread for a given day. Both mean and median values indicate that there are about 1 to 2 of these jump events on average per day. We also observe that the average level of volatility across the whole sample period lies between 15% and 31%, while the average price level ranges from \$25.59 to \$108.00. We clearly observe that the average bid/ask spread is increasing with the average price level. In our empirical analysis we divide our stocks into 4 groups on the basis of their bid/ask spread levels and select 4 reference stocks: HD, MCD, AXP, and IBM.

To obtain an idea of the time variation of the key variables, we plot (log) bid/ask spread, (log) trade duration, and (log) annualized volatility calculated using Equation (15) for AXP from 2002 to 2012 in Figure 1. We observe that periods of higher volatility coincide with periods of wider bid/ask spreads and lower trade durations. We observe very much the same pattern for all other stocks.

In Section 4, we carry out a comprehensive simulation study to analyze the properties of the duration based variance estimators. We will consider as benchmark the simulation scenario with 25% annualized volatility and 6 seconds average trade duration, which corre-

[3]From the list of 30 DJIA stocks as of December 2012, CSCO, INTC, and MSFT are excluded as their primary listing is at NASDAQ; BAC, CVX, HPQ, PFE, TRV, UNH, and VZ are excluded because of incomplete NYSE data samples.

Table 1: Descriptive statistics

Stock	bid/ask spread		trade duration		number of jumps		price	volatility
	mean	median	mean	median	mean	median	mean	mean
T	0.014	0.01	6.06	3.00	1.21	1.00	28.88	0.22
GE	0.014	0.01	4.58	2.00	0.98	1.00	27.95	0.24
DIS	0.015	0.01	6.01	3.00	1.63	1.00	29.60	0.24
HD	0.016	0.01	5.48	3.00	1.57	1.00	34.30	0.24
AA	0.016	0.01	6.82	3.00	1.32	1.00	25.59	0.31
KO	0.017	0.01	5.96	3.00	1.84	1.00	51.62	0.16
JPM	0.017	0.01	4.11	2.00	2.02	1.00	38.68	0.28
MRK	0.017	0.01	5.78	3.00	2.11	1.00	40.37	0.20
MCD	0.018	0.01	6.36	3.00	1.91	1.00	52.18	0.19
WMT	0.018	0.01	4.92	2.00	1.88	1.00	52.19	0.17
XOM	0.019	0.01	3.55	2.00	2.34	1.00	68.62	0.19
JNJ	0.018	0.01	5.40	3.00	2.11	1.00	61.36	0.15
DD	0.019	0.01	6.84	3.00	1.82	1.00	42.74	0.22
AXP	0.020	0.01	5.90	3.00	2.10	1.00	44.46	0.25
PG	0.020	0.01	5.41	3.00	2.31	1.00	66.14	0.15
BA	0.026	0.02	6.54	3.00	2.50	2.00	63.88	0.22
UTX	0.026	0.02	6.96	3.00	2.73	2.00	69.98	0.19
CAT	0.028	0.02	6.14	3.00	2.02	1.00	69.99	0.23
MMM	0.029	0.02	7.01	3.00	2.47	2.00	84.10	0.17
IBM	0.035	0.02	5.18	3.00	2.35	2.00	108.00	0.17

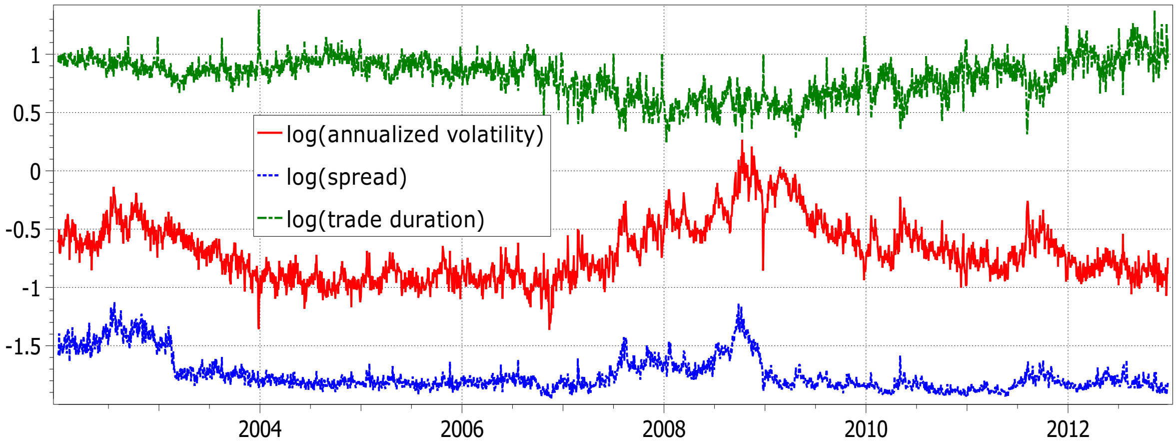
Notes: This table presents descriptive statistics for the bid/ask spread (in USD), the time between consecutive transactions (in seconds), the number of large price jumps per day, the transaction price, and the annualized volatility. A "large jump" is recorded when the absolute value of a price change exceeds 5 times the average bid/ask spread of the day. "Volatility" is calculated using Equation (15) and then annualized.

spond roughly to the average volatility and trade duration levels in Table 1. To assess the effect of bid/ask spread, we will consider scenarios with spread from 1 to 4 ticks. To assess the effect of time-discretization, we will consider scenarios with shorter trade durations of 3, 1 and 0.5 seconds. To assess the effect of jumps, we set the jump intensity to be 1 per day as a benchmark. We also examine the case where there are 100 small jumps per day for comparison. In both cases, the jump variance accounts for 20% of the total daily integrated variance.

4 Simulation study

We separate the MMS noise into time-discretization (Δ), bid/ask spread (ς), and price-discretization components. We investigate the separate and combined effects of the noise components as well as jumps on the non-parametric duration based volatility estimator, $NPDV$ (or NP for convenience), in a Monte Carlo study with 10000 replications. Specifically, we assess the performance of the NP estimator under different levels of: 1) time-

Figure 1: Bid/ask spread, trade duration and volatility for American Express (AXP)



Notes: Time series of trade duration, volatility, and bid/ask spread from 2002 to 2012. Bid/ask spread is the average spread in USD per day (logarithm) and trade duration is the average duration per day (in seconds, logarithm). The annualized volatility (logarithm) is calculated using Equation (15).

discretization, 2) bid/ask spread, 3) jump size and intensity.

The performance of the duration based integrated variance estimator depends on the selection of a preferred threshold value. Following the discussion of the two main sources of noise, bid/ask spread and time-discretization, we will discuss in Section 4.3 the tradeoff between efficiency and bias in the context of choosing a preferred threshold value δ^*.

Finally, in Section 4.6 we compare through simulation the accuracy and efficiency of the duration based variance estimator with other RV estimators, including the TSRV and RK estimators[4], which are also included in Section 6 for volatility forecasting comparison.

4.1 Time-discretization

Let us consider a discrete-time setting with a fixed time period, e.g. a trading day, Δ the discretization time interval and M equidistant intraday periods generated by (noisy) log-prices $Y_{i\Delta}$, $i = 0, \ldots, M$, which consist of a discretized efficient price process $X_{i\Delta}$ and a noise component process.

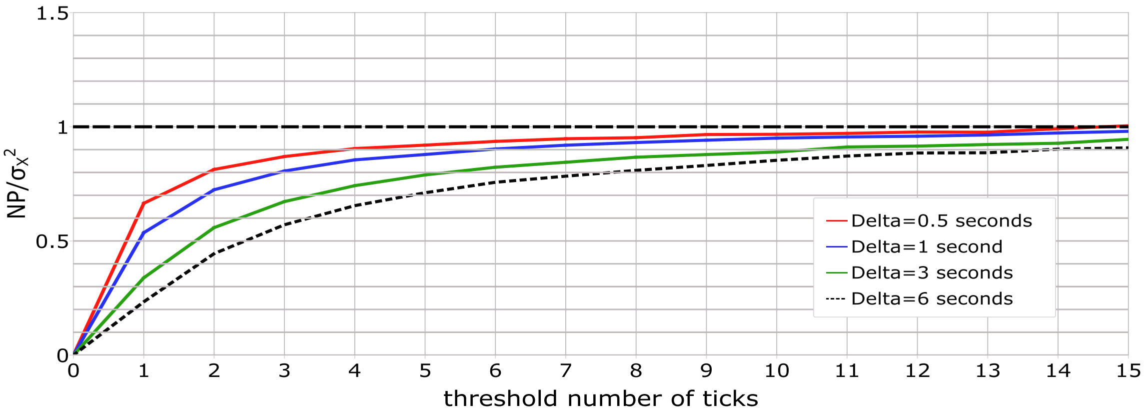
Discretizing X_t in Equation (1), yields a time-discretized diffusion component

$$X_t - X_{t-\Delta} = \sigma_X \sqrt{\Delta} Z_t. \tag{20}$$

[4]Both the cubic kernel and the Parzen kernel are used for the construction of the RK estimator. For the estimation of the optimal bandwidth H^*, we use 10 minutes sub-sampled RV to approximate the square-root of integrated quarticity and the 30 seconds sub-sampled RV to approximate the noise variance as suggested by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009). $c^* = 3.68$ for the cubic kernel and $c^* = 3.51$ for the Parzen kernel as stated in Table II by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008). The variance and auto-covariances are calculated using 1 minute returns, as suggested by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008). For the TSRV estimator, the fast scale is 30 seconds and the slow scale is 5 minutes.

Z_t is a standard normally distributed random variable and σ_X is the daily integrated volatility, which is assumed to be constant. In the finest time-discretized diffusion process, we set Δ to be half a second and there are 46800 half-seconds in a 6.5-hour daily trading session, so $i = 0, \ldots, 46800$. Upon this foundation process, we sample price points according to random Bernoulli distributions with probabilities 1/2, 1/6, and 1/12, resulting in three other time-discretized diffusion processes with average Δ of 1, 3, and 6 seconds respectively. Ratios of the NP variance estimates over the true integrated variance are plotted in Figure 2. We investigate how the average trade duration, Δ, and the threshold value, δ, affect the time-discretization noise, keeping the annualized[5] integrated volatility at 25% and no price variation outside trading sessions.

Figure 2: The time-discretization noise



Notes: Ratio of the NP variance estimates over σ_X^2. Δ's are 6, 3, 1, and 0.5 seconds from the bottom to the top. $\sigma_X = 0.25$ per year. Thresholds δ are from 0 to 15 ticks. $P_0 = 50$, tick size $= 0.01$.

Time-discretization decreases the number of events observed, due to the absence of price points that may have defined price events. As Δ decreases, the number of price points increases and N approaches its true value (in the case when prices are observed continuously). Thus, given δ, a smaller Δ leads to more accurate estimates of the integrated variance represented by the unit line in Figure 2, while increasing δ for a given Δ reduces the bias introduced by time-discretization.

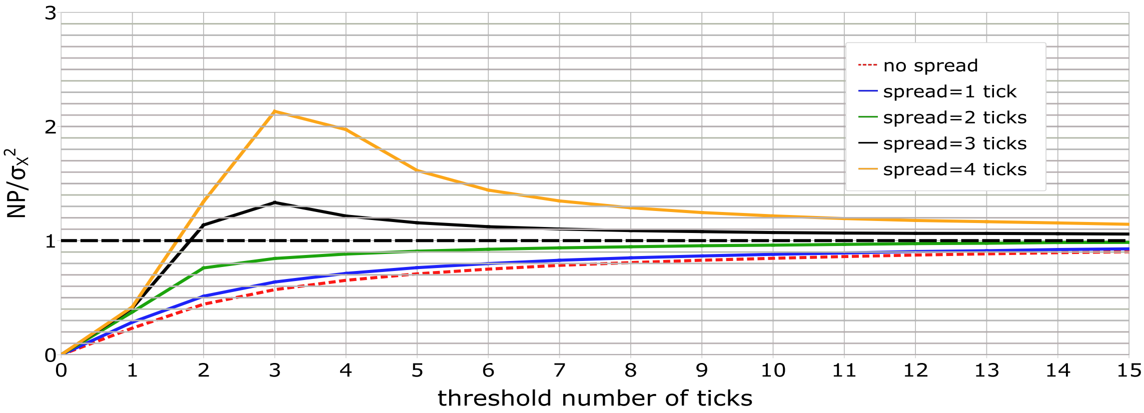
4.2 Bid/ask spread and time-discretization

As shown in Section 2.2, introduction of a bid/ask spread and corresponding bid and ask transaction prices biases the duration based variance estimates upwards, and the bias increases with the size of the spread ς, and decreases with the threshold value δ when $\delta > \varsigma$. We now consider discrete time, with an average Δ of 6 seconds, bid and ask transaction prices generated by $Y_t = X_t + 0.5 \mathbb{1}_t \varsigma$, with σ_X corresponding to 25% annualized volatility.

[5] Using 252 trading days per year.

The transaction price takes either the bid or the ask side with probability 0.5 and the variables $\mathbb{1}_t$ are i.i.d.

Figure 3: Combined effect of spread and time-discretization: bias



Notes: Ratio of the NP variance estimates over σ_X^2, with the range of threshold δs from 0 to 15 ticks. Bid/ask spreads ς from bottom to the top are 0 to 4 ticks. $\sigma_X = 0.25$ per year. Δ is 6 seconds on average. $P_0 = 50$, tick size=0.01.

Figure 3 shows ratios of the NP variance estimates over the true integrated variance. A deviation from the unit line indicates a bias. The hump-shaped curves occur as a result of the bid/ask spread component bias when the spread is relatively large. When $\delta < \varsigma$, one bid/ask bounce is enough to trigger a price event and N is inflated in comparison to the case when $\varsigma \to 0$ (dotted line). N does not decrease much as δ increases as long as $\delta < \varsigma$, causing the NP estimate, $N\delta^2$, to increase rapidly, until $\delta = \varsigma$. When δ further increases so $\delta > \varsigma$, the influence of bid/ask bounces is mitigated by the price changes from the efficient price component as a price event is now increasingly caused by the cumulative efficient price changes rather than by the bid/ask spread component. The bid/ask spread has the largest influence around the point where $\delta = \varsigma$.

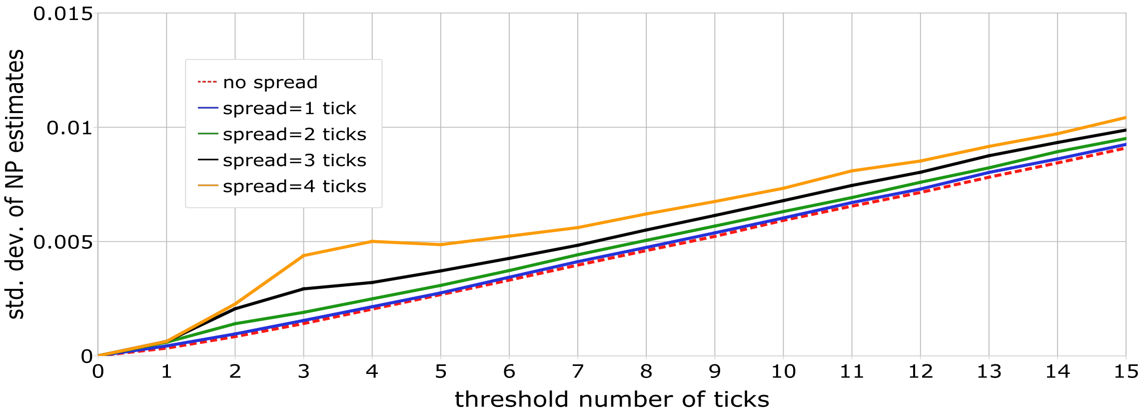
As δ increases past ς, the NP estimates start to stabilize, since both the time-discretization and the bid/ask spread biases are reduced by larger threshold values of δ. We observe two scenarios: 1) for smaller bid/ask spread levels (here 1 and 2 ticks) the negative bias contribution of the time-discretization is partially off-set by the positive contribution of the bid/spread components and the curves in Figure 3 for these cases tend to the unit line from below; 2) for larger bid/ask spread levels (here 3 and 4 ticks) the negative bias contribution of the time-discretization is, as discussed above, clearly dominated by the positive contribution of the bid/ask spread component and the curves in Figure 3 for these cases tend after the initial hump to approach the unit line from above.

4.3 Bias versus efficiency: the preferred threshold value

In reality we have no influence on the size of the bid/ask spreads nor the length of the trade durations, yet we must choose a threshold level δ for the implementation of our estimators.

From Sections 4.1 and 4.2 we know that the bias of the NP estimator decreases for a large enough threshold value, regardless of the bid/ask spread level. But, increasing the threshold level will inevitably result in a decreasing number of price events over the course of a day, rendering the NP estimates more dispersed and hence less efficient. Figure 4 shows this effect, as the standard deviation of the NP variance estimates is seen to increase over the range of δ from 0 to 15 ticks.

Figure 4: Standard deviations of the NP variance estimator



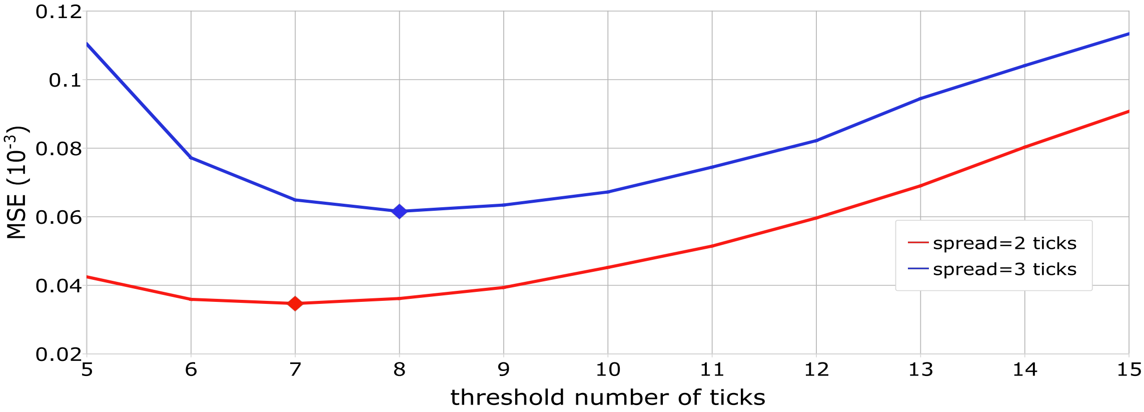
Notes: Standard deviations of the NP variance estimates over the range of threshold δs from 0 to 15 ticks. Bid/ask spreads ς from bottom to the top are 0 to 4 ticks. $\sigma_X = 0.25$ per year. Δ is 6 seconds on average. $P_0 = 50$, tick size=0.01.

To illustrate this trade-off we present in Figure 5 mean squared error (MSE) statistics for the NP estimator over the range of δ from 5 to 15 ticks, for 2-tick and 3-tick bid/ask spread levels. These are on average realistic bid/ask spread levels as shown in Table 1. For the 2-tick bid/ask spread case, the minimum MSE lies at $\delta^* = 7$ ticks, while for the 3-tick spread case, the minimum is given for $\delta^* = 8$ ticks. As these MSE minimum implying δ threshold value increase with the size of the bid/ask spread, we suggest for practical implementations to choose a preferred threshold δ^* equal to 2.5 to 3.5 times the bid/ask spread. A threshold in the range of 3 to 6 times the bid-ask spread is recommended in Andersen, Dobrev, and Schaumburg (2008) for a different duration based estimator. Further guidance about the choice of δ^* on the basis of bias-type curves, similar to those in Figure 3, for real data is presented in Section 5.2.

4.4 Price-discretization

In reality transaction prices are recorded as multiples of a minimum tick size, usually 1 cent. To account for this additional price-discretization component of market microstructure noise in our simulation study we now consider a setup in which, in addition to the above, bid and ask prices and consequently transaction prices are recorded discretely as multiples of 0.01 (one tick). First we obtain mid-quote prices by rounding the efficient price to the nearest 0.005 when $\varsigma/0.01$ is an odd number and to the nearest 0.01 when $\varsigma/0.01$ is an

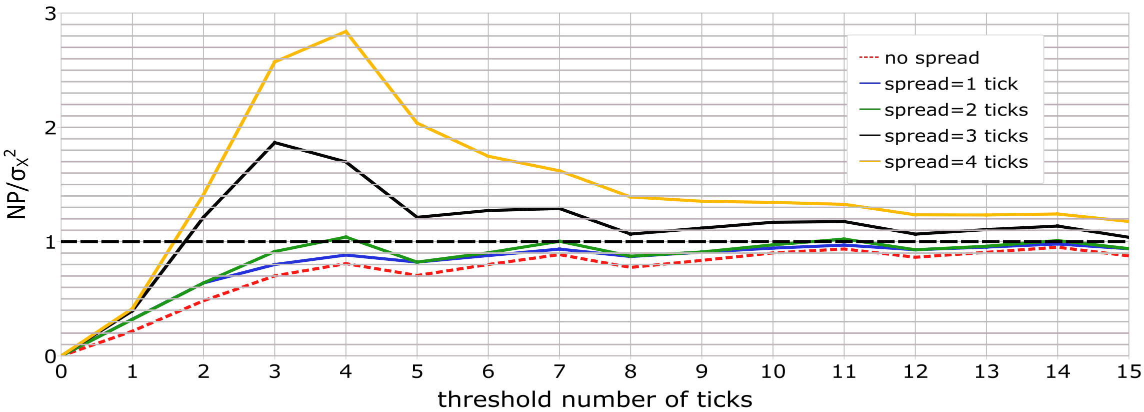
Figure 5: Plot of MSE as a function of the threshold value



Notes: MSE of the NP variance estimates over the range of δ from 5 to 15 ticks. Bid/ask spreads ς are 2 and 3 ticks. $\sigma_X = 0.25$ per year. Δ is 6 seconds on average. $P_0 = 50$, tick size=0.01.

even number. The resulting ask and bid prices are then given by "mid-quote+$\varsigma/2$" and "mid-quote$-\varsigma/2$", respectively. As before, trades arrive on average every 6 seconds and transaction prices take either the bid or the ask price according to a Bernoulli distribution with equal probability. Figure 6 shows that price-discretization produces a ragged patterns within our curves. The general effects of bid/ask spreads and time-discretization are, however, unchanged and the estimates still tend to the unit line as δ increases beyond ς.

Figure 6: Including price-discretization noise

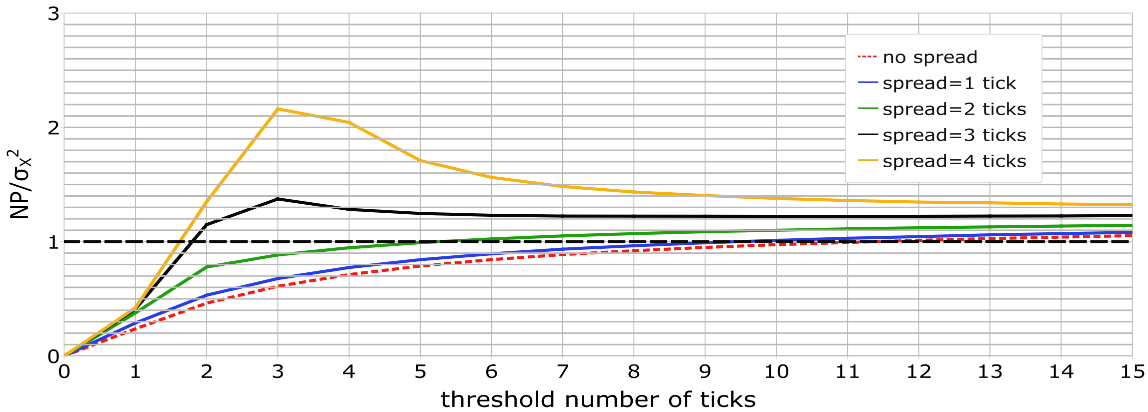


Notes: Ratio of the NP variance estimates over σ_X^2. Prices are multiples of one tick. Bid/ask spreads ς from bottom to the top are 0 to 4 ticks. Δ is 6 seconds on average. Thresholds δ are from 0 to 15 ticks. $\sigma_X = 0.25$ per year. $P_0 = 50$, tick size=0.01.

4.5 Jumps

To investigate how potential jumps affect our duration based integrated variance estimators, we consider the simulation setup of Section 4.2 and allow for price jumps. The size of jumps is set to be normally distributed with mean zero and a total variance of 20% of the true daily integrated variance. Jumps are simulated to arrive according to a random Poisson distribution. The jump intensity determines the standard deviation of the jump size and we consider two potential scenarios: 1) one large jump on average and 2) 100 small jumps on average during a day.

Figure 7: 100 small jumps a day



Notes: Ratio of the NP variance estimates over σ_X^2. The discretization interval is 6 seconds on average. There are on average 100 small jumps a day, with a total variance of 20% of the integrated variance. Bid/ask spreads from bottom to the top are 0 to 4 ticks. The discretization interval is 6 seconds on average. Thresholds δ are from 0 to 15 ticks. $\sigma = 0.25$ per year. $P_0 = 50$, tick size=0.01.

As discussed in Section 2.2, due to a truncation at δ, rare large jumps are expected to have little influence on the duration based variance estimates and indeed in scenario 1) there is no visible impact[6] as N is large and an increase of one potential additional price event, triggered by an expected single large jump, results only in a tiny upward bias of the NP estimator in the order of $1/N$. In scenario 2) the standard deviation of the jump size is 3.5 ticks. Here, on the contrary, we do observe in Figure 7 that small jumps increase the integrated variance estimates by around 16.3% in comparison to the no jump case. In this case estimates are inflated considerably as small jumps are mixed with the diffusion price changes and effectively increase the number of price events by a non-trivial amount. In reality we expect there to be less than one large jump per day, to which the duration based estimator in its current form is quite robust, and at most even only a small number of detectable smaller jumps per day. In fact many studies focussing on the detection of large jumps find on average less than a jump per week (c.f. Andersen, Bollerslev, and Dobrev (2007)). Lee and Hannig (2010) investigate the occurrence of big and small jumps in stock indices and individual stocks and find roughly one big jump every 3rd day and 0.6 small jumps per day for individual stocks with even fewer jumps detected in stock indices. Nonetheless, if the number of jumps is known (or can be estimated) a bias correction for

[6]We omit the graph for brevity.

jumps can readily be obtained.

4.6 Simulation comparison of different estimators

In Table 2, we present the simulation results under three reasonable scenarios where we compare the duration based variance estimators (with threshold values set as a range of multiples of spread, ς), with the TSRV, RK, and the subsampled 5-minute RV estimators. In scenario 1, $\Delta = 4$ seconds, $\varsigma = 1.5$ ticks; in scenario 2, $\Delta = 6$ seconds, $\varsigma = 2$ ticks; and in scenario 3, $\Delta = 10$ seconds, $\varsigma = 3$ ticks. The duration based estimator tends to be more efficient, showing lower STD, but also more biased, especially compared to the RK estimators. Overall, given a proper threshold value, such as 2.5 to 3.5 times the spread, the duration based estimator gives the lowest RMSE.

Table 2: Simulation comparison with other estimators

	Scenario 1			Scenario 2			Scenario 3		
δ	Bias	STD	RMSE	Bias	STD	RMSE	Bias	STD	RMSE
1ς	-0.0143	0.0012	0.0143	-0.0147	0.0013	0.0147	-0.0039	0.0021	0.0044
1.5ς	-0.0142	0.0012	0.0143	-0.0092	0.0020	0.0094	-0.0017	0.0035	0.0039
2ς	-0.0096	0.0019	0.0098	-0.0069	0.0024	0.0074	-0.0018	0.0038	0.0042
2.5ς	-0.0073	0.0025	0.0077	-0.0054	0.0030	0.0062	-0.0008	0.0052	0.0053
3ς	-0.0059	0.0031	0.0066	-0.0046	0.0040	0.0061	-0.0006	0.0056	0.0056
3.5ς	-0.0057	0.0030	0.0064	-0.0037	0.0044	0.0058	-0.0006	0.0071	0.0071
4ς	-0.0048	0.0037	0.0061	-0.0030	0.0050	0.0058	-0.0001	0.0079	0.0079
4.5ς	-0.0039	0.0044	0.0058	-0.0025	0.0056	0.0061	0.0003	0.0087	0.0087
5ς	-0.0032	0.0050	0.0059	-0.0021	0.0063	0.0066	0.0009	0.0095	0.0095
RK_{cubic}	-0.0004	0.0088	0.0088	-0.0005	0.0088	0.0088	-0.0000	0.0093	0.0093
RK_{Parzen}	-0.0004	0.0109	0.0109	-0.0001	0.0108	0.0108	-0.0004	0.0114	0.0114
$TSRV$	-0.0026	0.0080	0.0085	-0.0025	0.0081	0.0084	-0.0015	0.0081	0.0082
$5min$	0.0003	0.0081	0.0081	0.0009	0.0081	0.0081	0.0029	0.0081	0.0086

Notes: Scenario 1: $\Delta = 4$ seconds, $\varsigma = 1.5$ ticks; Scenario 2: $\Delta = 6$ seconds, $\varsigma = 2$ ticks; Scenario 3: $\Delta = 10$ seconds, $\varsigma = 3$ ticks. δ's (in ticks) are set as a range of multiples of spread, ς. $\sigma_X = 0.25$ per year. $P_0 = 50$, tick size=0.01.

5 Empirical Analysis

5.1 Parametric duration based variance estimator

For the implementation of the parametric duration based variance estimator, PDV, we consider three distributional assumptions for ε_d in equation (9): Exponential, Weibull, and Burr distributions. Price durations are obtained for a range of threshold values δ and are scaled each month, as described after equation (14), using a daily seasonality function obtained from a Nadaraya-Watson kernel regression. Maximum likelihood estimations (MLE) of the duration models represented by equations (9) and (10) under the three distributional assumptions are performed on a monthly basis.

We perform Likelihood ratio (LR), Ljung-Box (LB), and Density forecast (DF) tests to assess the goodness-of-fit of the models. The LR test compares the overall model fit between two nested models on the basis of their likelihood values. The LB test has the null hypothesis of i.i.d. distributed ε_d. The DF test of Diebold et al. (1998) tests the null hypothesis that the assumed distribution for ε_d is actually the true distribution and relies on a probability integral transformation of ε_d, namely the c.d.f. $F(\varepsilon_d)$, which under the null is i.i.d. $U(0,1)$ distributed. Provided that the HACD specification in equation (10) accommodates long-range dependence of the price durations data appropriately, and the assumed distribution for ε_d reflects the true distribution of the scaled duration, neither the LB nor the DF test should be rejected.

Table 3: Illustrative parameter values and tests results: AXP, year 2008, with threshold value equal to 12 ticks

Month	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
ω	0.053	0.173	0.035	0.077	0.161	0.167	0.020	0.085	0.008	0.031	0.025	0.067
	(0.015)	(0.059)	(0.011)	(0.044)	(0.083)	(0.043)	(0.008)	(0.028)	(0.003)	(0.007)	(0.010)	(0.028)
α	0.223	0.261	0.179	0.234	0.197	0.177	0.137	0.229	0.242	0.294	0.208	0.234
	(0.027)	(0.042)	(0.027)	(0.053)	(0.067)	(0.028)	(0.019)	(0.044)	(0.026)	(0.029)	(0.033)	(0.044)
β_1	0.727	0.462	0.835	0.643	0.683	0.652	0.866	0.346	0.675	0.436	0.644	0.676
	(0.078)	(0.154)	(0.093)	(0.161)	(0.317)	(0.075)	(0.074)	(0.167)	(0.059)	(0.103)	(0.135)	(0.139)
β_2	-0.014	0.032	-0.015	0.008	0.004	0.017	-0.031	0.086	0.002	0.050	0.015	0.006
	(0.021)	(0.030)	(0.017)	(0.031)	(0.051)	(0.015)	(0.014)	(0.037)	(0.011)	(0.018)	(0.024)	(0.023)
β_3	0.004	-0.002	0.002	0.001	-0.003	-0.004	0.007	-0.004	0.004	0.000	0.003	0.000
	(0.003)	(0.003)	(0.001)	(0.005)	(0.004)	(0.003)	(0.002)	(0.004)	(0.002)	(0.001)	(0.002)	(0.003)
γ	1.396	1.344	1.449	1.377	1.238	1.432	1.536	1.391	1.274	1.316	1.554	1.532
	(0.035)	(0.053)	(0.042)	(0.063)	(0.048)	(0.048)	(0.038)	(0.058)	(0.030)	(0.027)	(0.052)	(0.068)
η	0.518	0.421	0.480	0.475	0.187	0.424	0.533	0.383	0.478	0.481	0.571	0.552
	(0.050)	(0.074)	(0.056)	(0.089)	(0.059)	(0.062)	(0.051)	(0.077)	(0.043)	(0.038)	(0.067)	(0.087)
LL	-0.834	-0.907	-0.842	-0.900	-0.917	-0.908	-0.889	-0.896	-0.755	-0.810	-0.857	-0.770
LB50	0.051	0.951	0.643	0.200	0.097	0.057	0.349	0.291	0.044	0.016	0.678	0.761
DF	0.000	0.749	0.193	0.515	0.307	0.623	0.000	0.030	0.586	0.034	0.002	0.432

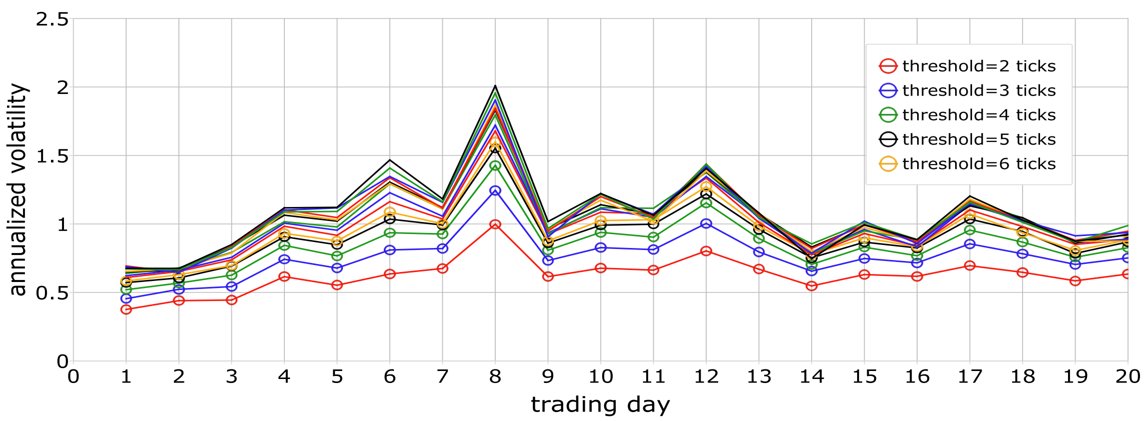
Notes: The first 14 rows are the parameter estimates and robust standard errors in parentheses for the Burr-HACD model in equations (9) and (10). "LL" are the average log-likelihood values and the last two rows are the p-values for LB statistics (at 50 lags) and DF tests.

All tests are performed, for each of the 132 months from January 2002 to December 2012, over a selected range of δ threshold values (between 2 to up to 20 ticks) for four reference stocks: HD, MCD, AXP and IBM. In the interest of brevity, all tests results are relegated to the Appendix. The conclusion is unequivocal: the Burr-HACD model fits the price durations data best. As an illustration, Table 3 presents the parameter values for the Burr-HACD model for AXP in 2008, with δ equal to 12 ticks, together with LB and DF tests results. As expected, we observe that, although there is some variation over the months, generally price durations are very persistent with an average β_1 equal to 0.64 and an average α equal to 0.22. The parameters η and γ have values that are significantly different from 0 and 1, respectively, which shows that the Burr specification provides a better fit than the Weibull or Exponential ones. The LB tests' p-values at lag 50 for the generalized model residuals indicate that the null hypotheses can only be rejected in 2 out of 12 cases on a 5% significance level and shows that generally the HACD specification dynamics provides a very good fit. The density forecasting tests' p-values reveal that the null hypotheses can be rejected in 5 out of 12 cases and indicates that there is scope to further improve, especially through the choice of a more flexible density function for ε_d, upon the Burr-HACD specification. The selection of more flexible densities than the Burr density usually comes with the cost of loosing some computational tractability and we refrain from considering them in this paper. Taken together, the fit provided by the Burr-ACD specification is very good and seems sensible, also in the light of Section 6 that focuses on out-of-sample forecasting comparisons.

5.2 The preferred threshold value

As discussed in Section 4.3, the selection of δ^* needs to take into account the tradeoff between improving efficiency and reducing bias: a larger δ reduces bias while a smaller δ improves efficiency. In the simulation study we know the true value of the integrated variance, and their MSE statistics for sensible simulation setups suggest that a threshold value δ^* should preferably be chosen to lie within the range of 2.5 to 3.5 times the bid/ask spread. In this section we provide a number of selective empirical results that support the conclusions of the simulation study and provide further guidance on how to select a preferred threshold δ^*. The results presented in this section focus on the reference stock AXP.[7]

Figure 8: Daily NP estimates for AXP: October 2008



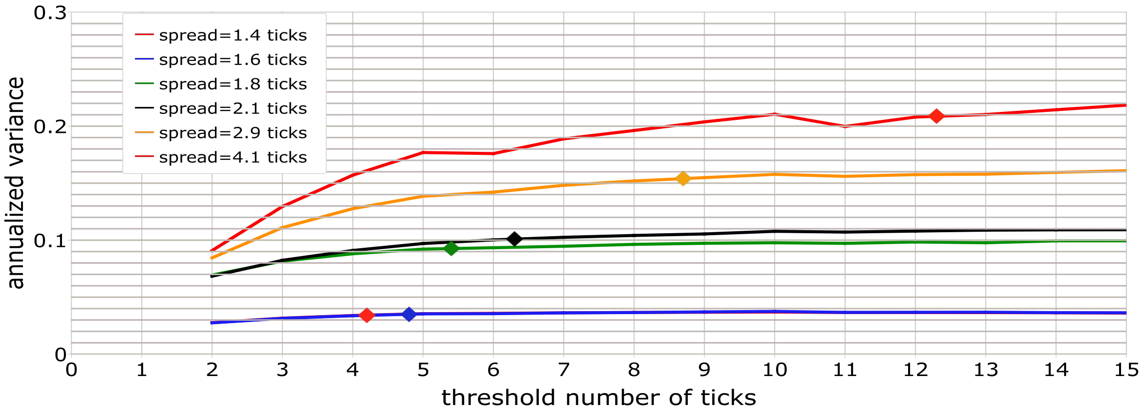
Notes: Daily NP estimates for the first 20 trading days of October 2008 for stock AXP, over the range of threshold values from 2 to 15 ticks (ordered generally from bottom to top).

We start by considering NP variance estimates in October 2008, when volatility is

[7]Results for the other stocks are available from the authors upon request.

peaked during the recent financial crisis. This month is governed by high uncertainty and average bid/ask spread levels of 4.6 ticks in this month are amongst the highest in our sample period. Figure 8 plots the NP variance estimates for the first 20 trading days of October 2008 for stock AXP, over the range of threshold values from 2 ticks to 15 ticks. We observe that, even during this high bid/ask spread level regime, duration based variance estimates first increase with the chosen threshold value and then stabilize, which is a stabilizing pattern that is similar to the one shown in Figure 3 for the simulation setting.

Figure 9: AXP: relationship between variance, threshold and bid/ask spread level



Notes: The average spreads of groups 1 to 6 for AXP are 1.4, 1.6, 1.8, 2.1, 2.9, and 4.1 ticks. One tick equals one cent.

The results of the simulation study suggest that estimates are less biased once stabilization has been achieved and pinpointing the lower bound of this stabilizing region would provide a good trade-off between bias and efficiency and a good choice for the preferred threshold value δ^*. To obtain a better picture of this stabilizing behavior, and its relationship to the level of the bid/ask spread in reality, we consider the full data sample for AXP. We divide the 132 months into 6 groups based on their average spread levels and obtain for each group daily NP variance estimates (annualized) for δ between 2 and 15 ticks and show their averages across days in Figure 9. The 6 groups represent in ascending average spread level order the lower 1/3, the middle 1/3 and the upper 1/3, subdivided into 4 ascending groups (1/12 each), of the data. Table 4 shows the distribution of the 6 groups across the 132 month in the data sample. It should be noted that many of the high bid/ask spread level months, besides the ones during the financial crises, are in the early years of the data sample when trading was less liquid, and consequently many of the low bid/ask spread level months are concentrated at the end of the data sample.

Figure 9 shows the stabilizing behavior of the duration based variance estimates very clearly and upon visual inspection, we observe that the threshold value at the point where the estimates start to stabilize, δ^*, is roughly three times the average bid/ask spread which is in line with the guidance obtained from the simulation study. We will use the "three-times-bid/ask-spread" rule henceforth as guidance to select δ^* for the computation of the

Table 4: Bid/ask spread level groups, AXP

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Jan.	5	6	2	2	2	1	5	3	1	1	1
Feb.	6	5	2	2	2	1	4	2	1	1	1
Mar.	5	2	2	1	1	2	4	1	1	1	1
Apr.	5	3	2	2	2	1	3	2	1	1	2
May	5	3	2	2	2	2	4	3	2	1	2
Jun.	6	4	1	1	2	2	4	1	1	1	2
Jul.	6	4	2	1	2	4	5	2	2	2	2
Aug.	6	3	1	1	2	5	5	2	1	3	2
Sep.	6	3	1	1	2	4	6	1	1	3	1
Oct.	6	3	2	2	1	4	6	1	1	3	2
Nov.	6	2	2	2	1	5	5	2	1	2	1
Dec.	6	2	2	1	1	4	4	1	1	1	1

PDV and NP estimators in the subsequent forecasting study.

Table 19 in the Appendix presents goodness-of-fit results (LB and DF tests) of the Burr-HACD model for all 20 stocks, with the price durations obtained by setting the threshold value to be δ^*. It confirms that, when the threshold value is set to be three times the average bid/ask-spread, the Burr-HACD fits the price durations data well.

6 Volatility forecasts evaluation

To assess the quality of our duration based variance estimators we conduct a comprehensive forecasting study. We compare our duration based variance estimators with variance estimators from two other important classes, RV-type and option-implied variance estimators. Our target volatility measures are the realized one-day, one-week, and one-month ahead 5-minute realized volatilities. From the class of duration based variance estimators we consider the PDV estimator based on the Burr-HACD specification discussed above, with parameters estimated monthly, variance estimates computed on the basis of previous month parameter values and δ^* equal to the three times the bid/ask spread of the previous month to avoid any forward information bias. We also consider NP variance estimators with δ^* equal to three times the bid/ask spread of the previous day, NP_d, and δ^* equal to three times the bid/ask spread of the previous month, NP_m. From the class of RV-type variance estimators we consider a realised kernel, RK, a two-scale realized variance, $TSRV$, a bi-power realized variance, BV, and a sub-sampled 5 minute realized variance, RV, estimator.[8] From the class of option-implied variance estimators we consider an at-the-money implied volatility, ATM, a model free implied volatility (with delta-implied-volatility-curves fitted using a quadratic specification), $MFIV_2$, and a model free implied volatility (with delta-implied-volatility-curves fitted using a cubic specification), $MFIV_3$,

[8]The cubic kernel is used here. The construction of the RK and TSRV estimators is the same as described in Section 4.6.

estimator.[9] In the interest of brevity, we present results only for the best performing estimators in the RV-type and option-implied classes[10]: RK and ATM. We find that the ATM option-implied volatility estimator gives more accurate forecasts of future volatility than both model free option-implied volatility estimators. This result is consistent with the finding of Martin, Reidy, and Wright (2009) showing that when three individual stocks are considered ATM estimators outperform model free estimators. For stock indices, as for example considered by Jiang and Tian (2005), model free estimators are usually found to show superior forecasting performance. For stock indices, in contrast to individual stocks, there is normally a more liquid and larger set of index options available, which allows for a more accurate approximation of the delta-implied-volatility-curves necessary for the construction of model free estimators. In the analysis below all estimators are used in an annualized form.

6.1 Individual Forecast

We employ a HAR-type forecasting equation,

$$RV_{n:n+h} = c + b_1 Z_{n-1} + b_2 Z_{n-5:n-1} + b_3 Z_{n-22:n-1} + \epsilon_{n:n+h}, \tag{21}$$

where Z_n represents the day-n volatility estimate from one of the five estimators discussed above and RV_n the day-n 5-minute realized volatility. Both $Z_{n-h:n-1}$ and $RV_{n:n+h}$ are in their logarithm forms: $Z_{n-h:n-1} = 0.5\log(\sum_{s=n-h}^{n-1} Z_s^2)$, similarly for $RV_{n:n+h}$, h=1, 5, or 22. For one day ($h = 1$) ahead forecasts the in-sample estimation period for the HAR model ranges from 1 February 2002 to 29 January 2010 (2013 trading days) and the first out-of-sample forecast is obtained for the 1 February 2010. For one week ($h = 5$) and one month ($h = 22$) horizons forecasts are constructed accordingly and a total of 735, 731 and 714 out-of-sample observations are obtained for $h = 1, 5$ and 22, respectively. All forecasts are constructed using a one day ahead rolling window.

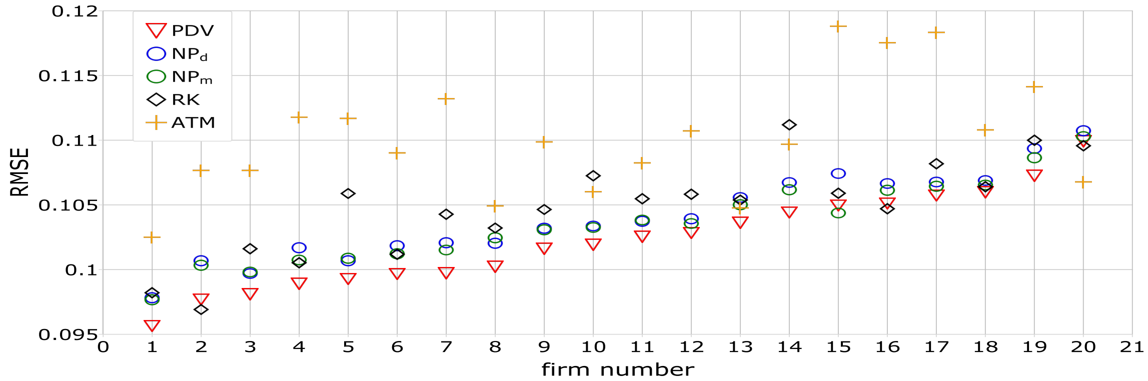
Figures 10, 11, and 12 show root-mean-squared-errors, RMSE, of the forecasts from the five different estimators, for one day, one week, and one month horizons. The 20 firms on the x-axis are sorted in ascending order of their RMSEs obtained from the PDV estimator

[9]The options data, which cover the same time period as the high-frequency trades and quotes data, are obtained from the OptionMetrics database. We directly employ the Black-Scholes implied volatility (IV), including the at-the-money implied volatility, provided by OptionMetrics. We retain options with time-to-maturities between 7 and 42 calendar days, and with positive bid-quotes and positive bid-ask spreads. Nearest-to-maturity options are usually chosen, but if they provide less than four (five) IV's for fitting the quadratic (cubic) curve, we switch to the second nearest-to-maturity day. For the construction of the model-free implied volatility estimator, we follow Taylor, Yadav, and Zhang (2010) and estimate the IV curve as a function of the Black-Scholes delta. We construct two versions of MFIV by fitting quadratic and cubic functions to the delta-IV curves. To prevent delta/IV points from clustering on one side of the curve, we require at least four (five) delta/IV observations a day, and at least one delta below 0.3, at least one above 0.7, and at least one between 0.3 and 0.7. In addition, we exclude extreme delta's larger than 0.99 or smaller than 0.01.

[10]Results for the other estimators can be obtained from the authors upon request.

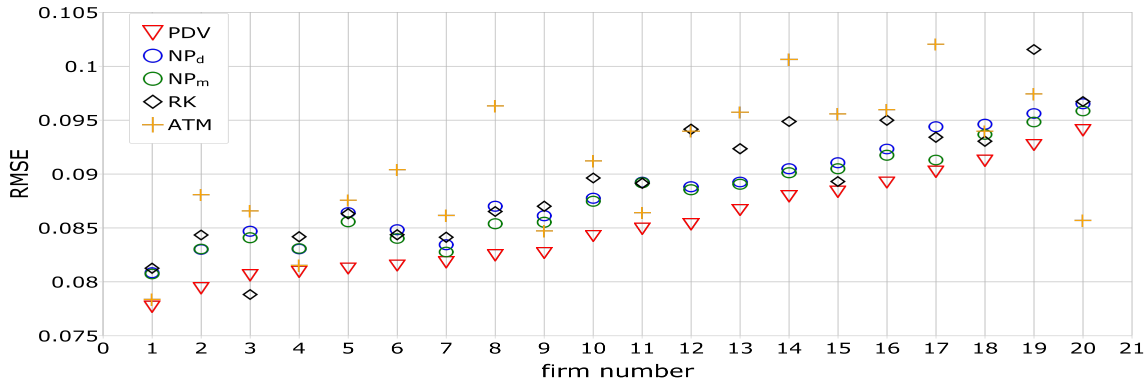
case.

Figure 10: RMSEs, individual forecast, one-day ahead



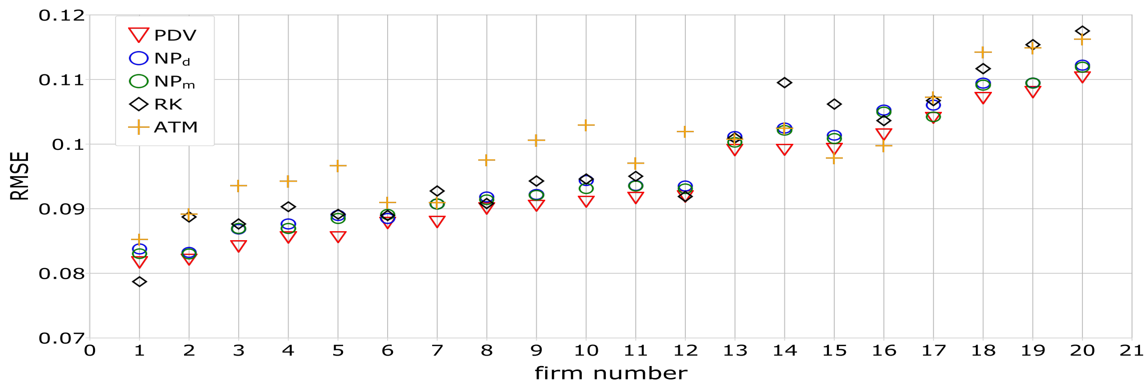
Notes: The firms on the x-axis are sorted in ascending order of the RMSEs of the *PDV* estimates.

Figure 11: RMSEs, individual forecast, one-week ahead



Notes: The firms on the x-axis are sorted in ascending order of the RMSEs of the *PDV* estimates.

Figure 12: RMSEs, individual forecast, one-month ahead



Notes: The firms on the x-axis are sorted in the ascending order of RMSEs of the *PDV* estimates.

Over all three forecasting horizons PDV generally produces the lowest RMSEs. To assess whether two competing forecasts perform significantly differently we perform a modified Diebold-Mariano (DM) test, using a squared error loss function. The DM test tests the null hypothesis of equal predictive ability of two competing forecasts by assessing the significance of their average loss differentials. For the 5% significance level, results are presented in Table 5. Each figure counts the number of significantly negative/positive loss differentials out of the 20 firms for the corresponding (estimators) pair in the first row. The figures in the "−" rows count the number of firms that favor the first estimator and the "+" rows favor the second estimator.

Table 5: DM test summary results, individual forecasts, 10 pairs, 3 horizons, 20 firms

pair	PDV-NP_d	PDV-NP_m	NP_m-NP_d	PDV-RK	NP_d-RK	NP_m-RK	PDV-ATM	NP_d-ATM	NP_m-ATM	RK-ATM
horizon					ONE DAY					
-	16	15	9	10	5	4	18	14	14	14
+	0	0	0	0	1	1	0	0	0	0
horizon					ONE WEEK					
-	19	18	7	10	1	1	10	2	4	4
+	0	0	0	0	1	1	0	0	0	1
horizon					ONE MONTH					
-	16	16	2	5	1	1	2	0	0	0
+	0	0	0	0	0	0	0	0	0	0

Notes: Each figure counts the number of significantly (5%) negative/positive loss differentials (out of the 20 firms) for the corresponding pair in the first row.

The first three columns compare forecasts based on the three estimators from the duration based variance class. Over all three horizons PDV estimators lead to significantly more accurate forecasts than the two corresponding nonparametric duration based variance estimators. Note, that a NP estimator for a given day uses only information from this specific day, while the PDV estimator for a given day uses additionally also information from past price durations for example through the dynamic Burr-HACD specification. Hence, the PDV estimator is based on a richer information set whose exploitation yields more precise variance estimators and consequently more accurate forecasts. The information advantage of this richer information set (together with a careful Burr-HACD model selection) also seems to dominate any additional model estimation noise from estimating the Burr-HACD specification. Moreover, NP_m performs better than NP_d which could stem form a smoothing effect of NP_m as it uses monthly average bid/ask spread levels to construct δ^* while NP_d relies on daily average bid/ask spread levels that are more volatile.

Columns 4-6 compare forecasts of duration based variance estimators with those from RKs. PDV leads to more precise forecasts than RK over all three horizons. The two nonparametric estimators perform better than RK over the one-day horizon, on par with RK over the one-week horizon, and are marginally better (winning by 1) over the one-month horizon.

Columns 7-9 compare forecasts of duration based variance estimators with those from

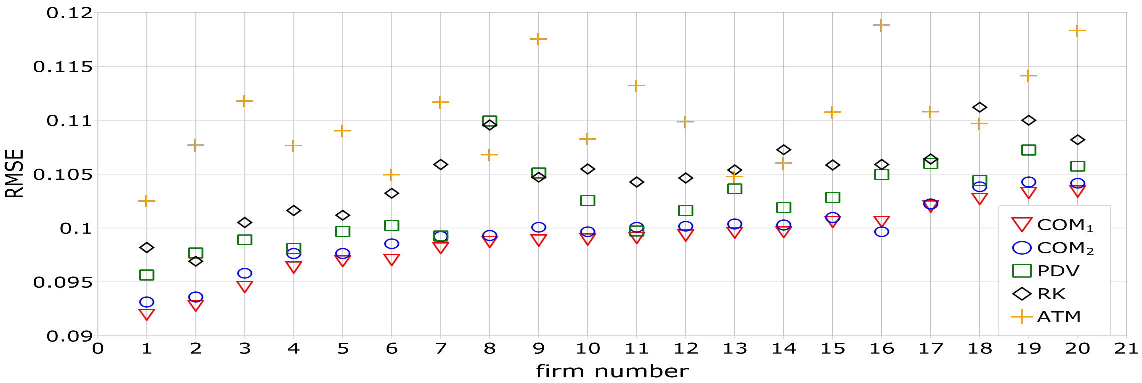
ATM option-implied variance estimators. Also here, the duration based variance estimators generally produce more accurate forecasts than ATM over all three horizons.

The last column compares forecasts from ATMs with those from RKs. RK shows better performance than ATM over the one-day and one-week horizons but performs on par with ATM over the one-month horizon. ATM estimators perform better as the forecasting horizon extends to one month. The forecasting comparison between the ATM option-implied volatility estimators and the RV estimators constructed from historical high-frequency returns is well documented in the existing literature. For example, Blair, Poon, and Taylor (2001), Pong, Shackleton, Taylor, and Xu (2004), and Taylor, Yadav, and Zhang (2010) find that, when the forecast horizon matches the maturity of the corresponding options, which is usually chosen to be around one month, the option-implied volatility estimator shows a better forecasting performance than estimators constructed from historical return data. Martin, Reidy, and Wright (2009) and Busch, Christensen, and Nielsen (2011) compare noise- and jump-robust RV measures, including RK, $TSRV$, and BV with ATM and draw similar conclusions. Over the one-month horizon, we find that RK and ATM perform on par and hence we do not find ATM to be a superior predictor of the one-month ahead future volatility.

Overall, among individual volatility estimators, PDV gives the most accurate forecasts over all three horizons. The easy-to-construct non-parametric duration based variance estimators also outperform the established RK and ATM variance estimators in most cases. We conclude that the duration-variance estimators are capable to extract information for integrated variance estimation and forecasting better than RV-type and option-implied variance estimators.

6.2 Combination Forecasts

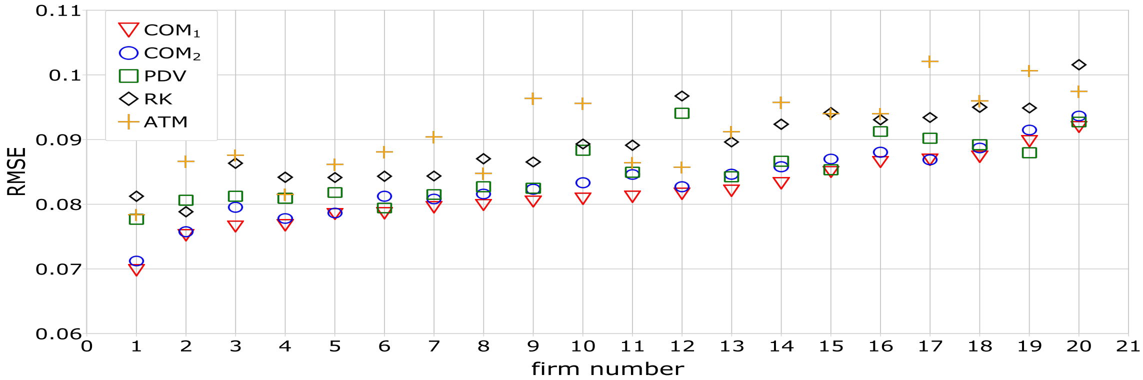
Figure 13: RMSEs, combination forecast, one-day ahead



Notes: The firms on the x-axis are sorted in ascending order of the RMSEs of the COM_1 estimates.

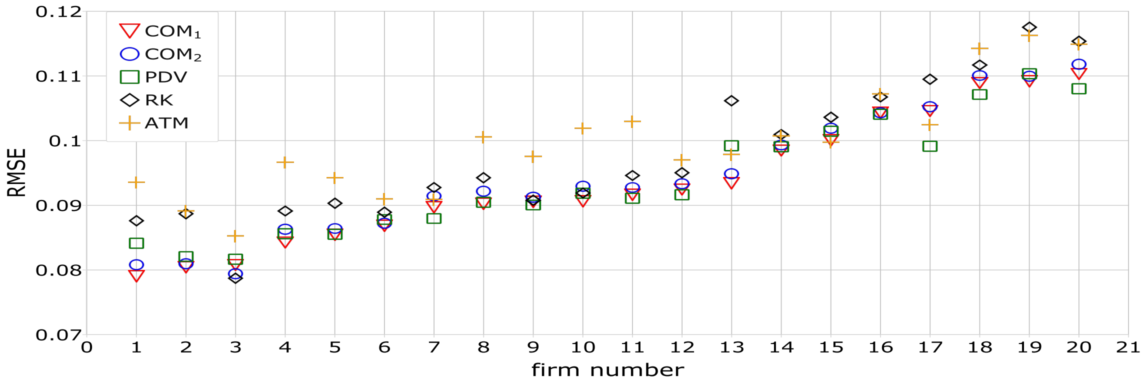
To address the question whether different integrated variance estimators extract differ-

Figure 14: RMSEs, combination forecast, one-week ahead



Notes: The firms on the x-axis are sorted in ascending order of the RMSEs of the COM_1 estimates.

Figure 15: RMSEs, combination forecast, one-month ahead



Notes: The firms on the x-axis are sorted in ascending order of the RMSEs of the COM_1 estimates.

ent and potentially complementary information from historical data that when combined leads to improved forecasting accuracy, we perform a combination forecasting study, in which estimators from the three classes above are used within an encompassing forecasting setup:

$$RV_{n:n+h} = c + \sum_{v=1}^{3} [b_{v,1}\ b_{v,2}\ b_{v,3}][Z_{v,n-1}\ Z_{v,n-5:n-1}\ Z_{v,n-22:n-1}]' + \epsilon_{n:n+h}. \quad (22)$$

We consider two combinations: COM_1, with $[Z_1\ Z_2\ Z_3] = [PDV\ RK\ ATM]$ and COM_2, with $[Z_1\ Z_2\ Z_3] = [NP_m\ RK\ ATM]$. These two combination forecasts are then compared with forecasts based on the three individual PDV, RK, and ATM estimators. Their RMSEs over the three forecasting horizons are plotted in Figures 13, 14, and 15. The 20 firms on the x-axis are sorted in ascending order of their RMSEs obtained from the COM_1 estimator.

Figures 13, 14, and 15 show that in general the combination forecasts are more accurate than any of the individual estimators based forecasts. Yet over longer horizons of one week/one month, PDV outperforms the combination forecasts in one/two cases. The corresponding DM tests are performed to assess whether these differences are significant on the 5% level.

Table 6: DM test summary results, combination forecasts, 7 pairs, 3 horizons, 20 firms

pair	COM_1-PDV	COM_1-RK	COM_1-ATM	COM_2-PDV	COM_2-RK	COM_2-ATM	COM_1-COM_2
horizon				ONE DAY			
-	16	20	20	12	20	20	16
+	0	0	0	0	0	0	1
horizon				ONE WEEK			
-	12	18	19	4	16	14	15
+	0	0	0	1	0	0	0
horizon				ONE MONTH			
-	1	3	6	0	3	5	3
+	0	0	0	1	0	0	0

Notes: Each figure counts the number of significantly (5%) negative/positive loss differentials (out of the 20 firms) for the corresponding pair in the first row.

In Table 6, the first 6 columns show that the combination forecasts are more accurate than any individual estimator based forecast. The last column compares the combination forecast using PDV with the one using NP_m: over all three horizons and as established from the comparison above of the single estimators based forecasts, the PDV-based combination forecast is more accurate.

Taken together, combining information from different sources improves forecast accuracy, and the parametric duration based variance estimator seems to extract relevant information better than the nonparametric one.

7 Conclusions

Duration based variance estimators are calculated by using the times of price change events; an event occurs when the magnitude of the price change since the previous event first equals or exceeds some threshold value. These estimators have been neglected in previous research, despite their potentially superior efficiency compared with realized variance estimators. The potential for superior efficiency occurs because duration based estimators make use of the complete path of prices, while standard RV estimators discard almost all prices for liquid securities such as the DJIA stocks studied in this paper. Market microstructure noise obscures theoretical comparisons and, furthermore, requires careful consideration to be given to the selection of the threshold value.

We use both Monte Carlo methods and real price data to recommend that an appropriate choice of the threshold is three times a measure of the average bid/ask spread. For this choice, duration based estimators have relatively small bias and relatively high efficiency (i.e. low mean squared error). We propose both parametric and nonparametric duration based estimators and find that they both forecast future volatility more accurately than either RV-type estimators or implied-volatility estimators for three forecast horizons (one day, one week and one month), when the forecasts are out-of-sample predictions for heterogeneous autoregressive models. Diebold-Mariano tests show that many of the forecast improvements are significant at the 5% level; for example, comparing the parametric duration estimator with the realized kernel estimator gives 10 significant results for the 20 stocks studied at a one-day horizon, 10 significant for the one-week horizon and 5 significant for the one-month horizon, with all significant differences favouring the duration estimator.

Calculation of the nonparametric duration estimator from a complete record of transaction prices is a trivial task. The parametric estimator is more accurate but does require the estimation of a parametric model for price events, which requires specifying intensity functions for durations whose conditional expectations are functions of previous durations. We recommend considering duration based estimators of integrated variation whenever transaction prices are available because of their potential to provide more accurate estimates and forecasts.

References

AIT-SAHALIA, Y., P. A. MYKLAND, AND L. ZHANG (2011): "Ultra High Frequency Volatility Estimation with Dependent Microstructure Noise," *Journal of Econometrics*, 160, 160–175.

ANDERSEN, T., D. DOBREV, AND E. SCHAUMBURG (2008): "Duration-Based Volatility Estimation," Working Paper, Northwestern University.

ANDERSEN, T. G., T. BOLLERSLEV, F. X. DIEBOLD, AND H. EBENS (2001): "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43–76.

ANDERSEN, T. G., T. BOLLERSLEV, AND D. DOBREV (2007): "No-arbitrage semimartingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications," *Journal of Econometrics*, 138(1), 125 – 180.

BALI, T. G., AND D. WEINBAUM (2007): "A Conditional Extreme Value Volatility Estimator Based on High-Frequency Returns," *Journal of Economic Dynamics & Control*, 31, 361–397.

BARNDORFF-NIELSEN, O. E., P. R. HANSEN, A. LUNDE, AND N. SHEPHARD (2008): "Designing Realized Kernels to Measure the Ex post Variation of Equity Prices In the Presence of Noise," *Econometrica*, 76, 1481–1536.

BARNDORFF-NIELSEN, O. E., P. R. HANSEN, A. LUNDE, AND N. SHEPHARD (2009): "Realized Kernels in Practice: Trades and Quotes," *Econometrics Journal*, 12, C1–C32.

BARNDORFF-NIELSEN, O. E., AND N. SHEPHARD (2002): "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society*, 64, 253–280.

BAUWENS, L., AND P. GIOT (2000): "The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks," *Annals of Economics and Statistics*, pp. 117–149.

BAUWENS, L., P. GIOT, J. GRAMMIG, AND D. VEREDAS (2004): "A Comparison of Financial Duration Models Via Density Forecasts," *International Journal of Forecasting*, 20, 589–609.

BECKER, R., A. E. CLEMENTS, AND S. I. WHITE (2007): "Does Implied Volatility Provide Any Information Beyond That Captured in Model-based Volatility Forecasts?," *Journal of Banking & Finance*, 31, 2535–2549.

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

Busch, T., B. J. Christensen, and M. O. Nielsen (2011): "The Role of Implied Volatility in Forecasting Future Realized Volatility and Jumps in Foreign Exchange, Stock and Bond Markets," *Journal of Econometrics*, 160, 48–57.

Cho, D., and E. W. Frees (1988): "Estimating the Volatility of Discrete Stock Prices," *Journal of finance*, 43, 451–466.

Corsi, F. (2009): "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7, 174–196.

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

Fernandes, M., and J. Grammig (2006): "A Family of Autoregressive Conditional Duration Models," *Journal of Econometrics*, 130, 1–23.

Gerhard, F., and N. Hautsch (2002): "Volatility Estimation on the Basis of Price Intensities," *Journal of Empirical Finance*, 9, 57–89.

Giot, P., and S. Laurent (2007): "The Information Content of Implied Volatility in Light of the Jump/Continuous Decomposition of Realized Volatility," *Journal of Futures Markets*, 27, 337–359.

Grammig, J., and K. Maurer (2000): "Non-monotonic Hazard Functions and the Autoregressive Conditional Duration Model," *The Econometrics Journal*, 3, 16–38.

Hautsch, N. (2004): *Modelling Irregularly Spaced Financial Data: Theory and Practice of Dynamic Duration Models*. Springer Science and Business Media.

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009): "Microstructure Noise in the Continuous Case: The Pre-averaging Approach," *Stochastic Processes and their Applications*, 119, 2249–2276.

Jasiak, J. (1999): "Persistence in Intertrade Durations," *Finance*, 19, 166–195.

Jiang, G. J., and Y. S. Tian (2005): "The Model-Free Implied Volatility and Its Information Content," *The Review of Financial Studies*, 18, 1305–1342.

Lancaster, T. (1997): *The Econometric Analysis of Transition Data*. Cambridge University Press.

Lee, S. S., and J. Hannig (2010): "Detecting jumps from Lévy jump diffusion processes," *Journal of Financial Economics*, 96(2), 271–290.

Martin, G. M., A. Reidy, and J. Wright (2009): "Does the Option Market Produce Superior Forecasts of Noise-Corrected Volatility Measures?," *Journal of Applied Econometrics*, 24, 77–104.

Nolte, I. (2008): "Modeling a Multivariate Transaction Process," *Journal of Financial Econometrics*, 6, 143–170.

Nolte, I., and V. Voev (2012): "Least Squares Inference on Integrated Volatility and the Relationship between Efficient Prices and Noise," *Journal of Business & Economic Statistics*, 30, 94–108.

Pacurar, M. (2008): "Autoregressive Conditional Duration Models in Finance: A Survey of the Theoretical and Empirical Literature," *Journal of Economic Surveys*, 22, 711–751.

Podolskij, M., and M. Vetter (2009): "Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps," *Bernoulli*, 15(3), 634–658.

Pong, S., M. B. Shackleton, S. J. Taylor, and X. Xu (2004): "Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA Models," *Journal of Banking & Finance*, 28, 2541–2563.

Taylor, S. J., P. K. Yadav, and Y. Zhang (2010): "The Information Content of Implied Volatilities and Model-free Volatility Expectations: Evidence from Options Written on Individual Stocks," *Journal of Banking & Finance*, 34, 871–881.

Tse, Y., and T. T. Yang (2012): "Estimation of High-Frequency Volatility: An Autoregressive Conditional Duration Approach," *Journal of Business & Economic Statistics*, pp. 533–545.

Zhang, L., P. A. Mykland, and Y. Ait-Sahalia (2005): "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411.

A Appendix

For the choice of a suitable density function for the scaled price durations we first consider LR tests for the four reference stocks: HD, MCD, AXP and IBM. The results in Tables 7, 10, 13 and 16 show that the Burr density is preferred over the Weibull and Exponential densities most of the time over a wide range of price change threshold values δ.

Corresponding LB test results for LB statistics with lags 30 and 50 are presented in Tables 8, 11, 14 and 17. For the majority of the months the null hypothesis of i.i.d. distributed generalized residuals cannot be rejected on the 1% and 5% significance levels, which indicates that the price duration dynamics are well captured by the HACD specification.

The associated density forecast (DF) test results in Tables 9, 12, 15 and 18 show that the Burr density clearly outperforms the other two distributional assumptions, by giving the highest percentages of months in which the null is not rejected on either the 1% or 5% significance level. From the three densities considered the Burr density provides the best fit for the scaled price durations.

Overall, the test results for the four reference stocks indicate that the HACD-Burr combination fits the price duration data best.

Finally, we present in Table 19 the LB and DF tests results for all 20 stocks, when the price change threshold δ is selected using the "3-times-spread" rule. We observe that the HACD-Burr model fits the price durations data well.

Table 7: LR test results, HD

δ(ticks)	2	3	4	5	6	7	8	9	10
Exp.	-0.92	-0.91	-0.91	-0.92	-0.92	-0.92	-0.92	-0.91	-0.92
Weibull	-0.92	-0.91	-0.91	-0.91	-0.91	-0.91	-0.91	-0.90	-0.90
Burr	-0.90	-0.89	-0.89	-0.89	-0.89	-0.89	-0.89	-0.89	-0.89
Wei. vs. Burr	1.00	1.00	1.00	1.00	1.00	0.95	0.93	0.74	0.68
Exp. vs. Burr	1.00	1.00	1.00	1.00	1.00	0.99	0.98	0.90	0.87
Exp. vs. Wei.	0.73	0.67	0.61	0.63	0.60	0.61	0.54	0.51	0.53

Notes: The first three rows are the average LL values from the MLE of the model in equations (9) and (10), and the last three rows LR test results presented as percentages of the months in which the null is rejected on the 5% significance level. The assumed density under the null is stated first in the 1st column.

Table 8: LB test results for 30 and 50 lags, HD

δ(ticks)	2	3	4	5	6	7	8	9	10
				30 lags 1% significance level					
Exp.	0.98	0.95	0.97	0.98	0.97	0.98	0.94	0.92	0.89
Weibull	0.97	0.94	0.95	0.98	0.95	0.98	0.93	0.92	0.92
Burr	0.87	0.86	0.90	0.92	0.95	0.95	0.94	0.86	0.89
				30 lags 5% significance level					
Exp.	0.86	0.89	0.91	0.90	0.93	0.93	0.89	0.87	0.85
Weibull	0.82	0.85	0.88	0.86	0.89	0.92	0.87	0.87	0.86
Burr	0.66	0.70	0.78	0.76	0.80	0.83	0.81	0.77	0.80
				50 lags 1% significance level					
Exp.	0.94	0.96	0.96	0.96	0.98	0.99	0.96	0.92	0.89
Weibull	0.93	0.95	0.96	0.96	0.96	0.98	0.95	0.93	0.92
Burr	0.87	0.91	0.93	0.90	0.96	0.99	0.96	0.87	0.90
				50 lags 5% significance level					
Exp.	0.82	0.86	0.91	0.89	0.92	0.95	0.90	0.86	0.86
Weibull	0.79	0.83	0.90	0.86	0.90	0.95	0.88	0.86	0.88
Burr	0.67	0.73	0.81	0.80	0.88	0.89	0.83	0.80	0.86

Notes: The upper part of the table are LB test results for 30 lags, and the lower part are the results for 50 lags. Significance levels of 1% and 5% are considered. Each figure is the proportion of months in which the null is not rejected.

Table 9: Density Forecast test results, HD

δ(ticks)	2	3	4	5	6	7	8	9	10
				1% significance level					
Exp.	0.00	0.00	0.01	0.03	0.11	0.34	0.31	0.52	0.53
Weibull	0.00	0.02	0.02	0.08	0.21	0.36	0.49	0.60	0.67
Burr	0.21	0.57	0.69	0.80	0.86	0.95	0.92	0.88	0.89
				5% significance level					
Exp.	0.00	0.00	0.00	0.01	0.03	0.20	0.23	0.32	0.44
Weibull	0.00	0.00	0.01	0.04	0.11	0.25	0.30	0.45	0.53
Burr	0.14	0.43	0.56	0.67	0.76	0.85	0.80	0.81	0.83

Notes: DF tests results for significance levels of 1% and 5% are presented. Each figure is the proportion of months in which the null that the assumed density is the true density is not rejected.

Table 10: LR test results, MCD

δ(ticks)	2	3	4	5	6	7	8	9	10
Exp.	-0.92	-0.91	-0.90	-0.90	-0.90	-0.90	-0.90	-0.89	-0.90
Weibull	-0.92	-0.91	-0.90	-0.90	-0.90	-0.89	-0.89	-0.89	-0.89
Burr	-0.90	-0.89	-0.88	-0.88	-0.88	-0.87	-0.87	-0.87	-0.87
Wei. vs. Burr	1.00	1.00	0.99	0.98	0.95	0.88	0.84	0.78	0.73
Exp. vs. Burr	1.00	1.00	0.99	0.99	0.98	0.95	0.93	0.91	0.88
Exp. vs. Wei.	0.86	0.59	0.50	0.50	0.43	0.42	0.34	0.35	0.32

Notes: The first three rows are the average LL values from the MLE of the model in equations (9) and (10), and the last three rows LR test results presented as percentages of the months in which the null is rejected on the 5% significance level. The assumed density under the null is stated first in the 1st column.

Table 11: LB test results for 30 and 50 lags, MCD

δ(ticks)	2	3	4	5	6	7	8	9	10
30 lags 1% significance level									
Exp.	0.93	0.96	0.98	0.95	0.98	0.99	0.93	0.93	0.89
Weibull	0.92	0.96	0.96	0.95	0.98	0.98	0.93	0.94	0.92
Burr	0.90	0.87	0.89	0.86	0.94	0.96	0.91	0.90	0.86
30 lags 5% significance level									
Exp.	0.83	0.86	0.88	0.86	0.87	0.96	0.88	0.90	0.83
Weibull	0.82	0.83	0.86	0.83	0.86	0.95	0.84	0.89	0.85
Burr	0.73	0.67	0.74	0.76	0.82	0.89	0.77	0.80	0.77
50 lags 1% significance level									
Exp.	0.90	0.92	0.97	0.95	0.98	0.99	0.90	0.92	0.88
Weibull	0.90	0.92	0.97	0.95	0.98	0.98	0.89	0.93	0.90
Burr	0.89	0.89	0.92	0.92	0.96	0.98	0.89	0.89	0.86
50 lags 5% significance level									
Exp.	0.85	0.87	0.88	0.89	0.96	0.98	0.86	0.86	0.86
Weibull	0.84	0.86	0.87	0.88	0.92	0.97	0.85	0.86	0.88
Burr	0.76	0.73	0.77	0.80	0.86	0.91	0.79	0.81	0.80

Notes: The upper part of the table are LB test results for 30 lags, and the lower part are the results for 50 lags. Significance levels of 1% and 5% are considered. Each figure is the proportion of months in which the null is not rejected.

Table 12: DF test results, MCD

δ(ticks)	2	3	4	5	6	7	8	9	10
				1% significance level					
Exp.	0.00	0.04	0.11	0.15	0.27	0.39	0.51	0.51	0.48
Weibull	0.01	0.07	0.18	0.21	0.34	0.43	0.57	0.63	0.61
Burr	0.24	0.55	0.75	0.80	0.83	0.92	0.88	0.85	0.84
				5% significance level					
Exp.	0.00	0.00	0.07	0.08	0.13	0.27	0.43	0.38	0.36
Weibull	0.01	0.03	0.10	0.13	0.23	0.32	0.40	0.47	0.48
Burr	0.14	0.45	0.61	0.70	0.72	0.83	0.83	0.80	0.76

Notes: DF tests results for significance levels of 1% and 5% are presented. Each figure is the proportion of months in which the null that the assumed density is the true density is not rejected.

Table 13: LR test results, AXP

δ(ticks)	2	3	4	5	6	7	8	9	10	11	12
Exp.	-0.93	-0.92	-0.91	-0.90	-0.91	-0.90	-0.90	-0.90	-0.90	-0.90	-0.90
Weibull	-0.93	-0.91	-0.90	-0.90	-0.90	-0.90	-0.90	-0.89	-0.89	-0.89	-0.89
Burr	-0.91	-0.89	-0.88	-0.88	-0.88	-0.87	-0.87	-0.87	-0.87	-0.87	-0.87
Wei. vs. Burr	1.00	1.00	1.00	1.00	0.99	0.99	0.96	0.95	0.89	0.77	0.65
Exp. vs. Burr	1.00	1.00	1.00	1.00	0.99	0.99	0.99	0.97	0.96	0.90	0.86
Exp. vs. Wei.	0.61	0.61	0.68	0.63	0.60	0.58	0.52	0.57	0.51	0.52	0.56

Notes: The first three rows are the average LL values from the MLE of the model in equations (9) and (10), and the last three rows LR test results presented as percentages of the months in which the null is rejected on the 5% significance level. The assumed density under the null is stated first in the 1st column.

Table 14: LB test results for 30 and 50 lags, AXP

δ(ticks)	2	3	4	5	6	7	8	9	10	11	12
					30 lags 1% significance level						
Exp.	0.93	0.93	0.95	0.98	0.98	0.97	0.96	0.97	0.95	0.87	0.92
Weibull	0.91	0.93	0.95	0.97	0.95	0.96	0.96	0.96	0.95	0.88	0.92
Burr	0.86	0.86	0.82	0.92	0.90	0.92	0.89	0.89	0.92	0.89	0.90
					30 lags 5% significance level						
Exp.	0.79	0.89	0.86	0.90	0.92	0.91	0.83	0.91	0.92	0.83	0.90
Weibull	0.73	0.88	0.85	0.88	0.89	0.90	0.83	0.92	0.91	0.82	0.90
Burr	0.60	0.69	0.67	0.75	0.73	0.82	0.77	0.81	0.83	0.77	0.82
					50 lags 1% significance level						
Exp.	0.89	0.95	0.98	0.97	0.98	0.96	0.94	0.98	0.95	0.87	0.91
Weibull	0.89	0.95	0.97	0.95	0.96	0.97	0.95	0.98	0.95	0.89	0.92
Burr	0.85	0.92	0.88	0.89	0.92	0.94	0.93	0.96	0.92	0.90	0.90
					50 lags 5% significance level						
Exp.	0.74	0.89	0.86	0.90	0.92	0.95	0.89	0.96	0.92	0.83	0.89
Weibull	0.73	0.88	0.83	0.88	0.89	0.95	0.89	0.94	0.91	0.85	0.89
Burr	0.65	0.75	0.77	0.79	0.80	0.88	0.85	0.83	0.86	0.80	0.85

Notes: The upper part of the table are LB test results for 30 lags, and the lower part are the results for 50 lags. Significance levels of 1% and 5% are considered. Each figure is the proportion of months in which the null is not rejected.

Table 15: DF test results, AXP

δ(ticks)	2	3	4	5	6	7	8	9	10	11	12
					1% significance level						
Exp.	0.00	0.00	0.00	0.02	0.08	0.13	0.27	0.35	0.45	0.46	0.52
Weibull	0.00	0.00	0.00	0.04	0.12	0.16	0.34	0.45	0.54	0.55	0.64
Burr	0.14	0.45	0.57	0.70	0.74	0.82	0.83	0.86	0.86	0.85	0.86
					5% significance level						
Exp.	0.00	0.00	0.00	0.01	0.02	0.06	0.16	0.20	0.30	0.30	0.40
Weibull	0.00	0.00	0.00	0.02	0.05	0.08	0.22	0.27	0.36	0.48	0.51
Burr	0.11	0.35	0.42	0.51	0.66	0.64	0.74	0.76	0.78	0.74	0.80

Notes: DF tests results for significance levels of 1% and 5% are presented. Each figure is the proportion of months in which the null that the assumed density is the true density is not rejected.

Table 16: LR test results, IBM

δ(ticks)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Exp.	-0.87	-0.86	-0.85	-0.85	-0.84	-0.84	-0.84	-0.84	-0.84	-0.84	-0.84	-0.82	-0.84	-0.84	-0.84	-0.82	-0.84	-0.83	-0.84
Wei.	-0.87	-0.85	-0.85	-0.84	-0.84	-0.84	-0.84	-0.83	-0.83	-0.83	-0.83	-0.82	-0.83	-0.83	-0.83	-0.80	-0.82	-0.83	-0.83
Burr	-0.85	-0.84	-0.83	-0.82	-0.82	-0.82	-0.82	-0.81	-0.81	-0.81	-0.81	-0.80	-0.81	-0.81	-0.81	-0.79	-0.80	-0.81	-0.81
Wei. vs. Burr	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.94	0.96	0.94	0.85	0.83	0.79	0.78	0.70
Exp. vs. Burr	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.90	0.98	0.98	0.91	0.89	0.86	0.89	0.83
Exp. vs. Wei.	0.84	0.82	0.77	0.69	0.65	0.66	0.64	0.67	0.65	0.63	0.61	0.55	0.61	0.64	0.63	0.61	0.58	0.60	0.62

Notes: The first three rows are the average LL values from the MLE of the model in equations (9) and (10), and the last three rows LR test results presented as percentages of the months in which the null is rejected on the 5% significance level. The assumed density under the null is stated first in the 1st column.

Table 17: LB test results for 30 and 50 lags, IBM

δ(ticks)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
30 lags 1% significance level																			
Exp.	0.76	0.88	0.92	0.94	0.98	0.98	0.96	0.97	0.96	0.98	0.97	0.91	0.98	0.99	0.95	0.93	0.91	0.95	0.91
Wei.	0.72	0.87	0.89	0.90	0.97	0.98	0.94	0.95	0.98	0.99	0.97	0.93	0.97	0.95	0.87	0.89	0.91	0.89	0.86
Burr	0.68	0.80	0.83	0.81	0.89	0.92	0.89	0.89	0.92	0.92	0.92	0.94	0.94	0.91	0.95	0.92	0.91	0.91	0.88
30 lags 5% significance level																			
Exp.	0.6	0.77	0.81	0.83	0.9	0.92	0.89	0.87	0.9	0.92	0.87	0.87	0.92	0.92	0.9	0.9	0.83	0.89	0.89
Wei.	0.55	0.76	0.78	0.81	0.89	0.91	0.89	0.86	0.92	0.91	0.84	0.88	0.91	0.89	0.83	0.86	0.83	0.82	0.83
Burr	0.43	0.59	0.65	0.67	0.72	0.75	0.75	0.75	0.75	0.82	0.72	0.82	0.82	0.83	0.83	0.8	0.8	0.78	0.8
50 lags 1% significance level																			
Exp.	0.72	0.89	0.89	0.96	0.95	0.98	0.98	0.98	0.98	0.99	0.97	0.89	0.98	0.99	0.95	0.92	0.92	0.93	0.9
Wei.	0.7	0.86	0.88	0.94	0.96	0.98	0.98	0.98	0.99	1	0.98	0.91	0.96	0.96	0.89	0.88	0.92	0.88	0.86
Burr	0.69	0.83	0.84	0.91	0.92	0.95	0.94	0.93	0.96	0.97	0.92	0.92	0.94	0.93	0.95	0.91	0.95	0.9	0.87
50 lags 5% significance level																			
Exp.	0.58	0.71	0.8	0.86	0.91	0.92	0.91	0.9	0.93	0.96	0.86	0.87	0.95	0.95	0.9	0.9	0.89	0.89	0.87
Wei.	0.53	0.69	0.77	0.83	0.88	0.9	0.89	0.89	0.95	0.95	0.86	0.89	0.93	0.89	0.82	0.85	0.87	0.83	0.83
Burr	0.49	0.6	0.65	0.67	0.81	0.8	0.79	0.8	0.82	0.86	0.79	0.83	0.84	0.85	0.81	0.82	0.85	0.83	0.83

Notes: The upper part of the table are LB test results for 30 lags, and the lower part are the results for 50 lags. Significance levels of 1% and 5% are considered. Each figure is the proportion of months in which the null is not rejected.

Table 18: DF test results, IBM

δ(ticks)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1% significance level																			
Exp.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.5	0.5
Wei.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.5	0.5
Burr	0.0	0.0	0.1	0.3	0.5	0.5	0.6	0.6	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.8	0.9	0.9	0.8
5% significance level																			
Exp.	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.02	0.03	0.07	0.11	0.13	0.17	0.24	0.25	0.25	0.36	0.33
Wei.	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.02	0.02	0.04	0.12	0.16	0.15	0.21	0.21	0.27	0.33	0.39	0.37
Burr	0.00	0.00	0.07	0.21	0.30	0.40	0.39	0.49	0.57	0.67	0.70	0.76	0.78	0.75	0.77	0.77	0.80	0.77	0.75

Notes: DF tests results for significance levels of 1% and 5% are presented. Each figure is the proportion of months in which the null that the assumed density is the true density is not rejected.

Table 19: All tests results for 20 stocks

	LB30(1%)	LB30(5%)	LB50(1%)	LB50(5%)	DF(1%)	DF(5%)
HD	0.93	0.84	0.95	0.83	0.80	0.68
MCD	0.91	0.75	0.94	0.80	0.82	0.73
AXP	0.88	0.69	0.91	0.77	0.73	0.55
IBM	0.94	0.80	0.95	0.83	0.73	0.57
AA	0.90	0.75	0.92	0.80	0.80	0.70
BA	0.87	0.73	0.92	0.82	0.87	0.79
CAT	0.95	0.84	0.95	0.86	0.67	0.51
DD	0.91	0.82	0.96	0.86	0.82	0.67
DIS	0.92	0.78	0.98	0.84	0.92	0.78
GE	0.96	0.80	0.93	0.85	0.82	0.62
JNJ	0.91	0.72	0.91	0.77	0.80	0.68
JPM	0.89	0.70	0.89	0.77	0.58	0.42
KO	0.90	0.73	0.94	0.81	0.83	0.73
MMM	0.96	0.83	0.97	0.89	0.79	0.69
MRK	0.90	0.77	0.92	0.86	0.77	0.61
PG	0.92	0.73	0.94	0.80	0.77	0.63
T	0.92	0.81	0.93	0.84	0.81	0.70
UTX	0.92	0.81	0.96	0.88	0.86	0.69
WMT	0.95	0.78	0.92	0.80	0.79	0.61
XOM	0.91	0.77	0.94	0.83	0.44	0.28
Avg.	0.92	0.77	0.94	0.82	0.77	0.63

Notes: LB and DF test results from the MLE of the HACD-Burr model in equations (9) and (10). The price durations are obtained with δ^* using the "3-times-spread" rule. Each figure in the table is the proportion of months in which the null is not rejected.