

Schulte-Tillmann, Björn; Segnon, Mawuli; Wiedemann, Timo

Article — Published Version

A Comparison of Realized Measures of Integrated Volatility: Price Duration- vs. Return-Based Approaches

Journal of Forecasting

Suggested Citation: Schulte-Tillmann, Björn; Segnon, Mawuli; Wiedemann, Timo (2026) : A Comparison of Realized Measures of Integrated Volatility: Price Duration- vs. Return-Based Approaches, Journal of Forecasting, ISSN 1099-131X, Vol. 45, Iss. 4, pp. 1579-1600, <https://doi.org/10.1002/for.70095>

This Version is available at:

<https://hdl.handle.net/10419/342890>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>

RESEARCH ARTICLE OPEN ACCESS

A Comparison of Realized Measures of Integrated Volatility: Price Duration- vs. Return-Based Approaches

Björn Schulte-Tillmann¹ | Mawuli Segnon²  | Timo Wiedemann³ 

¹Continentale Krankenversicherung a.G., Dortmund, Germany | ²Department of Economics, University of Münster, Münster, Germany | ³Department of Finance, Chair of Derivatives and Financial Engineering, University of Münster, Germany

Correspondence: Mawuli Segnon (segnon@uni-muenster.de)

Received: 6 May 2024 | **Revised:** 23 October 2025 | **Accepted:** 20 December 2025

Keywords: high-frequency data | price duration | realized measures of integrated volatility | value-at-risk

ABSTRACT

We study the accuracy of a variety of parametric price duration-based realized variance estimators constructed via various financial duration models and compare their forecasting performance with the performance of various nonparametric return-based realized variance estimators. Our financial duration models consist of an autoregressive conditional duration (ACD), its logarithmic version (Log-ACD), and the fractionally integrated ACD (FIACD), as well as the Markov switching multifractal duration (MSMD) model and the factorial hidden Markov duration (FHMD) process. In an empirical study using high-frequency data on 10 stocks traded on the New York stock exchange (NYSE), our in-sample and out-of-sample results show that the parametric price duration-based realized variance (RV) estimators, especially the ACD-based RV estimator, seem to be robust to price jumps and microstructure noise and perform better than the non-parametric return-based RV estimators. Furthermore, we also find that the price duration- and return-based RV models perform relatively well and produce more accurate and valid value-at-risk forecasts than the GARCH(1,1).

1 | Introduction

Although 60 years of research has been dedicated to the study of asset price variability, it has lost none of its appeal for academia and practitioners due to its ongoing importance for risk management, derivative pricing, and asset allocation. Initially, a lot of attention has been devoted to the development of parametric approaches for volatility modeling. The incorporation of the stylized facts of financial return time series data, such as volatility clustering, leverage, fat tails, jumps, and long memory, has led to an abundance of GARCH-type and stochastic volatility (SV) models[see Zaharieva et al. (2020) and the references therein for a recent overview.

Owing to the advent of high-frequency data, during the last two decades, the focus for the measurement of asset price variability

has been shifted from parametric to nonparametric methods. The first model-free approaches for constructing daily variance estimators utilizing high-frequency intraday return data date back to the seminal works of Andersen et al. (2001a), Andersen et al. (2001b) and Barndorff-Nielsen and Shephard (2002a,b). They establish a realized variance (RV) estimator to obtain a consistent estimate of an asset's integrated variance. However, as remarked by Hansen and Lunde (2006) as well as by Bandi and Russell (2008), exploiting the entire record of observations induces a severe bias in the proposed RV estimator, when the price process is distorted by market microstructure noise. In fact, it is well-documented that asset prices recorded at high frequencies are characterized by discontinuities and contaminated by microstructure noise due to various market frictions such as information asymmetries; see Glosten and Milgrom (1985), transaction costs; see Garman (1976) and their presence lead to

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Journal of Forecasting* published by John Wiley & Sons Ltd.

biases and inconsistency of the RV estimators; see Hansen and Lunde (2006), Hautsch and Podolskij (2013). Subsequently, a variety of approaches have been proposed to handle these issues; see Liu et al. (2015) and Mukherjee et al. (2020) for a more comprehensive review on RV estimators.

The availability of high-frequency data has also unleashed the development of financial duration models for testing some market microstructure theories, understanding the abundance of issues related to trading and price adjustment processes and estimating instantaneous volatility; see Engle and Russell (1998) for the autoregressive conditional duration (ACD) model, Bauwens and Giot (2000) for the logarithmic ACD (Log-ACD), Jasiak (1998) for the fractionally integrated ACD (FIACD), Chen et al. (2013) and Žikeš et al. (2017) for the Markov switching multifractal duration (MSMD) model, and Segnon and Schulte-Tillmann (2025) for the factorial hidden Markov duration (FHMD) model.¹ Recent applications of ACD models in estimating instantaneous volatility include Tse and Yang (2012), Liu and Tse (2015), as well as Hong et al. (2023) and indicate a superiority of the price intensity-based estimators² over the return-based RV ones.

Despite the promising results of the price duration-based approaches, it lacks a systematic study of the impact of different modeling frameworks, the sampling methods, and the distributional assumptions on the forecasting performance of parametric price duration-based RV estimators. To this end, we consider a portfolio that consists of the ACD(1,1) model, its logarithmic version, Log-ACD(1,1) and the FIACD(1,1) model. Apart from the ACD-type models, we further incorporate the MSMD and the FHMD model and adopt a Burr distribution for the innovations in all our financial duration models. The reasons for our choice of the financial duration models and the distributional assumption for the innovation are twofold: (i) the empirical findings in Bauwens et al. (2004) show that the ACD(1,1) and Log-ACD(1,1) model with a Burr or Generalized Gamma distribution for the innovation outperform more complex financial duration models in terms of modeling the dynamics of the data and forecasting. (ii) It is well-documented that financial duration data are characterized by long memory and nonlinearities, features that can be reproduced by the FIACD, MSMD, and the FHMD model. In an empirical study using time series data on price durations of 10 actively traded stocks on the New York Stock Exchange (NYSE), we compare the forecasting performance of our price duration-based RV estimators with the performance of returns-based RV estimators. Our in-sample and out-of-sample results show that the price duration-based RV estimators are robust to microstructure noise and perform better than the return-based RV estimators in most cases. However, we note that the performance of the price duration-based RV estimators across various modeling frameworks depends on the distributional assumption used for the innovation. Furthermore, we also find that the ACD(1,1) model produces more accurate and valid value-at-risk forecasts than the standard GARCH(1,1).

The rest of this paper is organized as follows. Section 2 presents the theoretical background and setup for the estimation of integrated variance. Section 3 discusses the empirical application. Finally, Section 4 concludes the study.

2 | Setup

We briefly describe the setup for both return- and price duration-based realized measures of the integrated variance. We consider a standard arbitrage-free framework, in which the log-price of a financial asset at continuous time t , say $Y = (Y_t)_{t \geq 0}$, follows a Brownian semi-martingale process without jumps. Y is defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, \infty)}, P)$ and adapted to the filtration $\{\mathcal{F}_t\}_{t \in [0, \infty)}$, that is, a family of σ -fields with $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for all $s \leq t \leq \infty$.

Let Y be characterized by the following stochastic differential:

$$Y_t = Y_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \quad (1)$$

where $\mu = (\mu_t)_{t \geq 0}$ is locally bounded and predictable, $\sigma = (\sigma_t)_{t \geq 0}$ is càdlàg, and $W = (W_t)_{t \geq 0}$ a standard Brownian motion. As pointed out by Hull and White (1987) when there are no jumps, quadratic variation equals the integrated variance. So, the true variation of process Y_t in our framework³ is given by

$$IV_t = \int_{t-1}^t \sigma_s^2 ds, t = 1, \dots, T \quad (2)$$

where IV is the integrated variance at day t . It is well-known that IV is crucial in asset- and derivatives-pricing and risk management, and since IV is latent, different methods have been proposed to estimate integrated variance at intraday level. Moreover, it well-documented that high-frequency data are contaminated by market microstructure noise. This suggests that the observed log-price X can be formalized as follows:

$$X = Y + \epsilon, \quad (3)$$

where Y is the unobserved log-price and ϵ denotes market microstructure noise.

We assume that high-frequency data are available through $[0, t]$, which is the sampling period and is thought of as being a trading day. Let Δ be the sampling frequency. The intraday return at the equidistant interval $[t_{(i-1)\Delta}, t_{i\Delta}]$ is $r_{t_i} = X_{t_{i\Delta}} - X_{t_{(i-1)\Delta}}$.

To complete the setup, we consider the following stochastic process, where at the beginning of a trading period, t_0 , we mark every point in time, at which the absolute cumulative price change exceeds a certain threshold, ι , by $t_1 < t_2 < \dots < t_N$. As a result, we obtain the strictly increasing sequence of hitting times $\{t_j\}_{j=0, \dots, N}$. Moreover, we denote the associated total number of hitting events (or price events) up to time t by N_t . The following conditional price intensity function

$$\lambda(t|N_t, t_1, \dots, t_{N_t}) = \lim_{\Delta t \rightarrow 0} \frac{Pr(N_{t+\Delta t} > N_t | N_t, t_1, \dots, t_{N_t})}{\Delta t} \quad (4)$$

is well-known to completely describe the stochastic process underlying the price events. We use the conditional price intensity

function to infer the integrated variance in the price duration modeling frameworks, our objective of interest.

2.1 | Return-Based RV

The first nonparametric return-based realized measure of integrated volatility, called RV, has been proposed by Andersen et al. (2001b). They showed that under certain conditions, RV is an unbiased and efficient estimator of the quadratic variation of intraday returns. To estimate the integrated variance in this paper, we consider the RV ($RV^{(\Delta)}$),⁴ its subsampled version ($RV_{ss}^{(\Delta)}$), the time-scale ($TSRV^{(\Delta)}$), and multiple time scale RV estimator (MSRV); see Zhang et al. (2005) and Zhang (2006). Moreover, we also utilize the realized kernel estimator (RK), see Barndorff-Nielsen et al. (2008); the noise-robust RV estimator ($RV_{pa}^{(\Delta)}$), see Christensen et al. (2014) and the realized bipower variation estimator (BPV), see Barndorff-Nielsen and Shephard (2006). We refer the reader to the [Supporting Information](#) for details on the return-based RV estimators used in this paper.

2.2 | Price Duration-Based RV

We revisit the estimation of integrated volatility based on the conditional price intensity function which represents the probability of a price event at time $t > t_{i-1}$ given the price event has not occurred before time t and the information set available up to arrival time t_{i-1} , denoted by $\mathcal{F}_{i-1}$ as follows:

$$\lambda(d|\mathcal{F}_{i-1}) = \frac{f(d|\mathcal{F}_{i-1})}{1 - F(d|\mathcal{F}_{i-1})}, \quad (5)$$

where $d = t - t_{i-1}$ and $f(\cdot)$ and $F(\cdot)$ denote the density and cumulative distribution function, respectively. According to Engle and Russell (1998), the conditional intensity function (or equivalently the conditional hazard function) can be linked to the conditional instantaneous variance at time t_i by

$$\sigma^2(t|\mathcal{F}_{i-1}) = \left(\frac{t}{P_{i-1}} \right)^2 \lambda(d|\mathcal{F}_{i-1}), \quad (6)$$

where P_{i-1} denotes the asset's price at time t_{i-1} . Building on this work, Tse and Yang (2012) and Hong et al. (2023) provide the foundation for non-parametric as well as parametric price duration-based estimators. By integrating over the interval (t_{i-1}, t_i) , we obtain an estimator of the integrated variance. However, typically, we are interested in the integrated variance over a specific time interval $[t_0, t_N]$, which represents a trading day, an hour, and so on. Hence, we aggregate the single integrals as follows:

$$\begin{aligned} \widehat{IV}_t &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \sigma^2(t|\mathcal{F}_{i-1}) dt \\ &= \sum_{i=1}^N \left(\frac{t}{P_{i-1}} \right)^2 \int_{t_{i-1}}^{t_i} \lambda(t - t_{i-1}|\mathcal{F}_{i-1}) dt \\ &= -t^2 \sum_{i=1}^N \frac{\ln(1 - F(d_i|\mathcal{F}_{i-1}))}{P_{i-1}^2}. \end{aligned} \quad (7)$$

Given Equation (7), we see that the estimation of the integrated variance necessitates an appropriate choice of the conditional intensity function. Following Engle and Russell (1998), we model financial price durations, d_i , as

TABLE 1 | Overview about RV measures.

	Realized measures	Acronym	Source
Return-based	Basic realized variance	$RV^{(\Delta)}$	Inter alia Andersen et al. (2001a)
	Subsampled realized variance	$RV_{ss}^{(\Delta)}$	Zhang et al. (2005)
	Realized kernel	$RK^{(\Delta)}$	Barndorff-Nielsen et al. (2008)
	Two-scale realized variance	$TSRV^{(\Delta)}$	Zhang et al. (2005)
	Multiscale realized variance	$MSRV^{(\Delta)}$	Zhang (2006)
	Pre-averaged realized variance	$RV_{pa}^{(\Delta)}$	Inter alia Jacod et al. (2009)
	Realized bipower variation	$BPV^{(\Delta)}$	Barndorff-Nielsen and Shephard (2006)
Price-Intensity-based	Factorial hidden Markov duration	PDV_{FHMD}	Segnon and Schulte-Tillmann (2025)
	Markov-switching multifractal duration	PDV_{MSMD}	Chen et al. (2013) and Zikes et al. (2017)
	Autoregressive conditional duration	PDV_{ACD}	Engle and Russell (1998)
	Fractionally integrated ACD	PDV_{FIACD}	Jasiak (1998)
	Logarithmic-ACD	$PDV_{Log-ACD}$	Bauwens and Giot (2000)
	Nonparametric duration-based variance	$NPDV$	Hong et al. (2023)

Note: For return-based RV measures, we attach each estimator's sampling frequency (Δ) in the superscript. In the subscript, we provide further details on the estimator in abbreviated form (ss - subsampled; parzen, cubic, th2 or bartlett - names of weighting functions; db - debiased estimator; pa - pre-averaged). For duration-based RV measures, we point out which data samples we use, that is, "5min" or "Spread" data, in the superscript. In the subscript, we denote the employed duration model in combination with the choice of the distribution for the innovation term. A model name without a specifically mentioned distribution refers to the baseline case of exponentially distributed errors.

$$d_i = \psi_i \epsilon_i, i \in \mathbb{Z}, \quad (8)$$

where ψ_i denotes the conditional duration and $\{\epsilon_i\}$ is a sequence of i.i.d. unit-mean innovations with nonnegative support. Furthermore, we assume that the innovation $\{\epsilon_i\}$ follows a Burr distribution that encompasses the exponential and the Weibull distribution as special cases; see Hautsch (2012). We consider the ACD, the Log-ACD, the FIACD, the MSMD, and the FHMD frameworks; see Table 1⁵ for an overview of our modeling approaches for the conditional duration, ψ_i . We note that all the models are estimated via the maximum likelihood approach.

Hong et al. (2023) also propose a nonparametric duration-based variance estimator (NPDV) which is given by

$$\text{NPDV}_t = \sum_{i=1}^{N_t} \left(\frac{t}{P_{i-1}} \right)^2, \quad (9)$$

where $N \equiv N_t$ denotes in praxis a reasonable number of price events over interval $[t_0, t_N]$ that represents a trading day for example.

3 | Empirical Analysis

3.1 | Data Description and Adjustment

We apply the parametric as well as the nonparametric approaches for the estimation of the daily variance on 10 actively traded stocks. We investigate the returns and the price durations of Apple (AAPL), Bank of America (BAC), The Walt Disney Company (DIS), Evergy Inc (EVRG), Facebook (FB), General Electric (GE), International Business Machines (IBM), Pfizer (PFE), Tesla, Inc. (TSLA), and Walmart (WMT). For each stock, except FB and TSLA, we extract the data for the time period from January 3, 2007 to December 31, 2019, spanning 3272 trading days, from the Trade and Quote (TAQ) database. Owing to later initial public offerings, we base our analysis on data beginning on May 18, 2012 for FB and on June 29, 2010 for TSLA, covering 1917 and 2394 trading days, respectively. The TAQ database is inevitably prone to market microstructure and requires a certain amount of cleaning. Our data cleaning steps are described in Appendix A. Generally, we concentrate on trades that occurred between 9:30 and 16:00 and resample the trade data to a second by second frequency taking volume weighted price averages within one second.

We define the price duration for each stock as the minimal time required to observe a cumulative change in the price not less than the threshold ι .⁶ We follow two approaches to find an appropriate threshold. Firstly, in line with Tse and Yang (2012), we aim to set the value for the daily threshold, such that we obtain an average price duration of five minutes. The choice of 5 min sampling intervals and its impact on duration-based realized volatility have been documented by Andersen et al. (2008). Their empirical findings show that this sampling interval choice reduces the impact of microstructure noise on duration-based RV estimators considerably and guarantees their robustness.

To this end, we define a fine grid of relative thresholds $\iota = \iota^* / \bar{P}$, where $\bar{P}$ is the average daily stock price and select the threshold

value that leads to an average duration that comes closest to our target. Secondly, akin to Hong et al. (2023), we define the threshold parameter as the average daily bid/ask spread times a factor lying in the range of 3 to 10. In both approaches, we discard all overnight durations following Engle and Russell (1998).

The intensity of trading activities on financial markets changes over the course of a trading day. Typically, the average duration is short at market open and close and comparatively longer in between. At the beginning of a day, trading activities are very high due to new events that have occurred during the night (macroeconomic or firm-related news that have become public after the previous market close). At the end of a day, traders tend to close their positions to reduce their exposure to overnight stock market risk.⁷ These systematic variations are responsible for the pronounced diurnal seasonality observed in high-frequency financial duration data. However, since these variations are predetermined by market characteristics, we remove the intraday seasonality component, so that we are able to implement our parametric estimation approaches to the stochastic part of the data. Some studies, like Rodríguez-Poo et al. (2008), directly incorporate the deterministic part in their modeling approach. But as Engle (2000) notes, this procedure does not yield a remarkable efficiency gain relative to a two-stage adjustment process. Since Tse and Dong (2014) find that the diurnal adjustment does not have a huge impact on the daily volatility estimation, we follow the simple-to-implement dummy-adjustment approach of Ghysels et al. (2004)⁸:

We split the 6.5-h trading time per day into 13 intervals of 30 min. Then, we construct an indicator variable

$$y_{ki} = \begin{cases} 1 & \text{if } i \in k, \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

for each 30-minutes interval $k = 1, \dots, 13$ and perform a logarithmic dummy variable regression

$$\ln(D_i) = \sum_{k=1}^{13} \alpha_k y_{ki} + u_i, \quad (11)$$

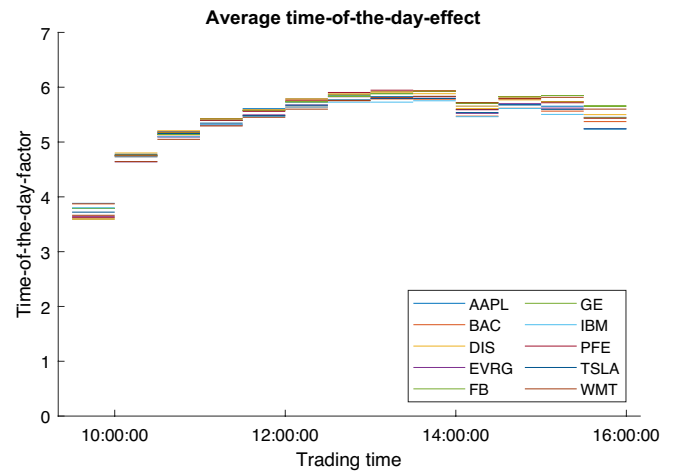


FIGURE 1 | Plots of the mean diurnal seasonality factor for all 10 stocks using 5-min data.

TABLE 2 | Descriptive statistics of the raw and adjusted duration data.

		Raw duration data					Adjusted duration data			
		No. of obs.	Mean	Overdisp.	Skewness	Kurtosis	Mean	Overdisp.	Skewness	Kurtosis
5-min data	AAPL	251,348	299.731	1.346	3.768	30.547	1.613	1.072	3.416	31.136
	BAC	253,726	296.695	1.406	4.057	33.295	1.689	1.142	3.718	38.271
	DIS	251,566	299.364	1.359	3.509	24.091	1.714	1.175	3.823	31.864
	EVRG	251,864	299.244	1.419	4.161	36.818	1.858	1.210	4.210	50.773
	FB	146,714	299.486	1.378	3.787	30.656	1.597	1.059	3.372	33.972
	GE	253,982	297.236	1.412	4.013	33.176	1.755	1.180	3.556	28.246
	IBM	251,565	299.895	1.385	3.814	30.253	1.722	1.156	3.578	28.198
	PFE	253,489	298.128	1.368	3.688	26.967	1.744	1.159	3.321	23.325
	TSLA	182,464	299.669	1.534	4.363	37.765	1.800	1.296	5.779	94.962
	WMT	251,848	299.485	1.412	3.906	32.217	1.716	1.164	3.704	34.967
Spread data	AAPL	168,165	430.600	1.957	6.549	71.632	1.709	1.309	8.047	206.391
	BAC	585,670	127.545	2.771	10.489	209.479	1.746	1.394	13.949	1151.886
	DIS	253,402	294.969	1.668	5.455	61.693	1.797	1.338	7.845	233.522
	EVRG	181,249	409.463	1.748	5.554	55.344	1.862	1.357	7.469	179.940
	FB	396,104	111.695	1.841	9.759	259.408	1.699	1.244	5.898	118.267
	GE	378,093	198.547	2.114	8.096	138.145	1.828	1.372	5.816	97.187
	IBM	252,155	296.280	1.727	5.569	60.275	1.777	1.233	4.153	39.834
	PFE	95,662	747.073	1.721	4.597	35.803	1.783	1.335	8.736	266.290
	TSLA	261,735	207.892	1.990	8.367	147.362	1.858	1.415	8.248	220.006
	WMT	128,303	526.859	1.796	4.946	42.316	1.823	1.616	19.660	910.529

Note: For each stock, we report the descriptive statistics for the (raw and adjusted) 5-min data in the top panel and for the (raw and adjusted) spread-based data in the bottom panel.

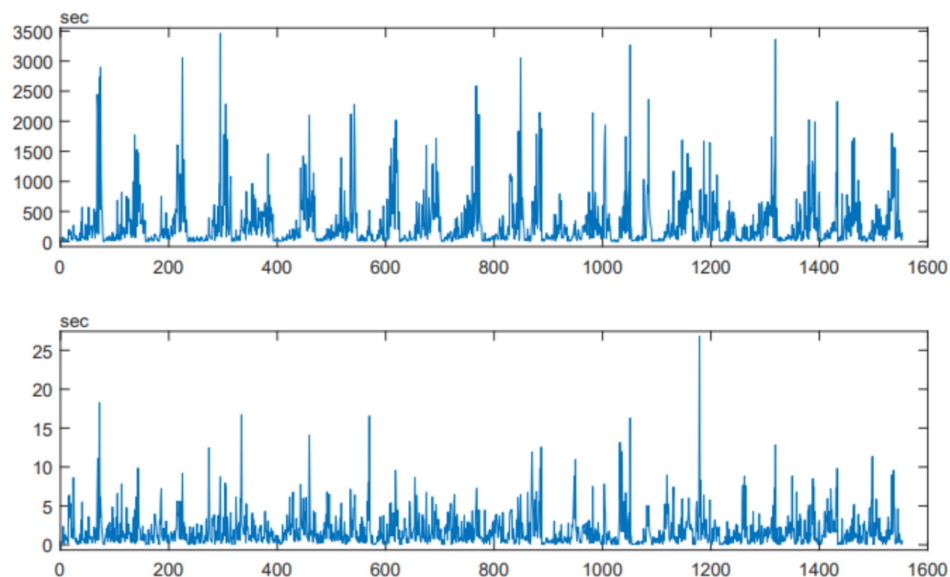


FIGURE 2 | Plot of raw and adjusted IBM price durations for September 2019.

where D_i represents the raw duration data and u_i the error term. Finally, we eliminate the *time-of-the-day-effect* and obtain the adjusted durations by

$$d_i = D_i \exp(-\hat{\alpha}' y_i), \quad (12)$$

where $\hat{\alpha} = (\hat{\alpha}_1, \dots, \hat{\alpha}_{13})'$ and $y_i = (y_{1i}, \dots, y_{13i})'$. In line with the aforementioned literature, our durations exhibit a hump-shaped pattern over a trading day. As Figure 1 illustrates, the increase of the trading activity at the end of a day is not as distinct as at the beginning of the day.⁹ Table 2 reveals that all stocks exhibit overdispersion, right skewness and heavy tails, irrespective of the method used to obtain the raw duration data (5 min or spread-based). However, the data sets differ in many aspects. The mean duration ranges from almost 2 (FB) to more than 12 min (PFE) for the spread-based data, whereas for the other data set, the average time for a price event is close to 5 min as desired. We also find that the degree of overdispersion, skewness, and kurtosis is always remarkably higher for the spread-based data. The data adjustment process leads to a small attenuation of the

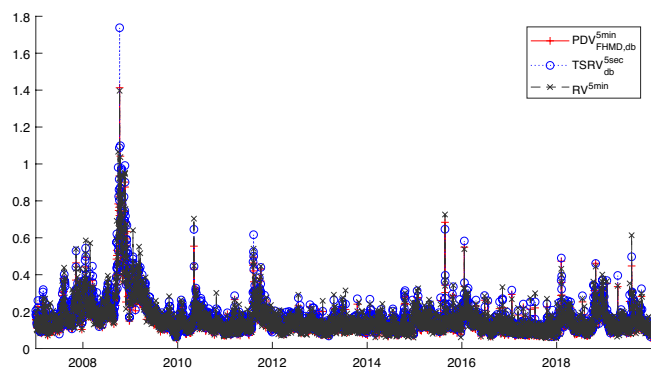


FIGURE 3 | Annualized volatility of IBM.

degree of overdispersion but to an amplification of the other two characteristics for both data sets in most of the cases. Hence, we can ascertain that the method used to determine the threshold parameter will have a great impact on the received price duration data. Finally, Figure 2 depicts the raw IBM price durations and the adjusted IBM price durations obtained via the logarithmic dummy variable regression approach, clearly indicating that the adjustment process successfully removes the diurnal pattern present in the top panel of Figure 2.

3.2 | In-Sample Analysis

In our in-sample analysis, we estimate the daily variance of all 10 stocks over the whole sample period using all previously presented estimators. Figure 3 shows the annualized volatility of IBM for an illustrative set of RV estimators ($\text{PDV}_{\text{FHMD, db}}^{\text{5min}}$, $\text{TSRV}_{\text{db}}^{\text{5sec}}$ and RV^{5min}). For all estimates, volatility peaks during the financial crisis, and also beyond that, the trajectories follow each other quite closely but differ in their level of annualized volatility.

The general tendency that estimates keep close track of each other is also confirmed for the remaining estimates by the correlation heat map in Figure 4, in which we display the correlation coefficients between all RV estimates for IBM. We note that especially the parametric duration-based estimates originating from different duration models are highly correlated with each other. In comparison to former studies; see Patton and Sheppard (2009), which rely on older high-frequency data from 1996 to 2008, we find a higher overall level of correlation. Results for the other stocks are qualitatively (and quantitatively) the same and not shown here.

In the next step, we evaluate the estimation accuracy of all presented RV estimators. In contrast to Tse and Yang (2012) and Hong et al. (2023), who assess the accuracy of their newly

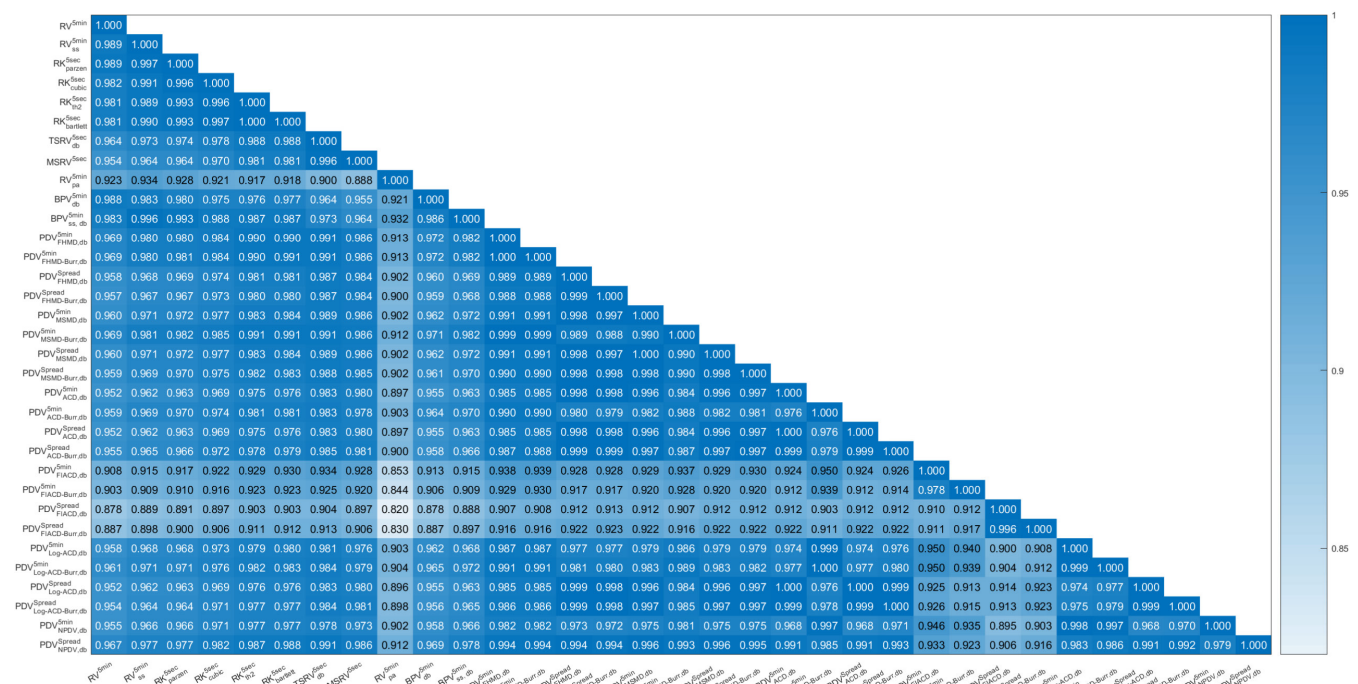


FIGURE 4 | Correlation heatmap for realized variance estimates (IBM).

TABLE 3 | In-sample performance of realized measures (proxy: 5-min RV).

	AAPL		BAC		DIS		EVRG		FB	
	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE
RV ^{5min}	5.064	0.279	14.013	0.225	2.900	0.211	4.292	0.237	3.787	0.231
RV ^{5min} _{ss}	5.317	0.279	15.106	0.219	3.487	0.208	4.263	0.220	3.461	0.229
RK ^{5sec} _{parzen}	5.296	0.265	15.750	0.224	3.860	0.201	4.261	0.223	3.424	0.218
RK ^{5sec} _{cubic}	5.479	0.258	16.104	0.219	3.842	0.193	4.185	0.211	3.386	0.213
RK ^{5sec} _{th2}	6.211	0.245	15.701	0.197	3.767	0.183	4.332	0.193	3.388	0.204
RK ^{5sec} _{bartlett}	5.899	0.245	15.634	0.197	3.741	0.182	4.310	0.192	3.391	0.203
TSRV ^{5sec} _{db}	5.061	0.230	15.181	0.181	3.357	0.171	5.417	0.189	3.808	0.190
MSRV ^{5sec}	4.578	0.675	14.071	0.417	3.101	0.510	4.221	0.235	3.490	0.601
RV ^{5min} _{pa}	4.754	0.627	14.791	0.538	3.017	0.464	4.794	0.684	3.529	0.506
BPV ^{5min} _{db}	5.248	0.313	13.100	0.259	2.689	0.234	4.260	0.280	3.431	0.258
BPV ^{5min} _{ss,db}	5.281	0.309	14.796	0.242	3.187	0.229	4.057	0.258	3.297	0.253
PDV ^{5min} _{FHMD,db}	4.366	0.284	14.134	0.206	2.965	0.231	3.918	0.388	3.271	0.242
PDV ^{5min} _{FHMD–Burr,db}	4.380	0.282	13.951	0.204	2.987	0.228	3.915	0.385	3.141	0.240
PDV ^{Spread} _{FHMD,db}	4.196	0.241	12.893	0.184	2.605	0.185	3.726	0.262	2.993	0.228
PDV ^{Spread} _{FHMD–Burr,db}	4.148	0.234	12.853	0.181	2.608	0.182	3.757	0.257	2.981	0.222
PDV ^{5min} _{MSMD,db}	4.494	0.283	13.969	0.207	3.108	0.229	3.920	0.390	3.199	0.241
PDV ^{5min} _{MSMD–Burr,db}	4.464	0.279	14.176	0.203	3.009	0.227	3.916	0.386	3.220	0.237
PDV ^{Spread} _{MSMD,db}	4.216	0.250	12.893	0.189	2.616	0.191	3.748	0.272	3.004	0.236
PDV ^{Spread} _{MSMD–Burr,db}	4.203	0.241	13.075	0.187	2.630	0.187	3.801	0.264	2.987	0.230
PDV ^{5min} _{ACD,db}	4.377	0.268	13.603	0.189	3.049	0.224	4.033	0.386	3.047	0.235
PDV ^{5min} _{ACD–Burr,db}	4.371	0.269	13.473	0.190	2.961	0.223	4.007	0.387	3.089	0.235
PDV ^{Spread} _{ACD,db}	4.086	0.236	12.871	0.182	2.561	0.181	3.705	0.261	2.979	0.229
PDV ^{Spread} _{ACD–Burr,db}	4.088	0.237	12.762	0.182	2.561	0.185	3.706	0.261	2.982	0.229
PDV ^{5min} _{FIACD,db}	4.546	1.004	14.572	0.815	3.074	0.818	3.829	1.205	3.527	0.846
PDV ^{5min} _{FIACD–Burr,db}	4.880	1.081	14.365	0.795	3.140	0.841	3.775	1.391	3.554	0.911
PDV ^{Spread} _{FIACD,db}	4.461	12.446	13.023	3.449	2.905	2.304	3.893	5.646	3.107	0.388
PDV ^{Spread} _{FIACD–Burr,db}	4.463	11.267	13.004	3.305	2.875	2.021	3.880	5.177	3.111	0.386
PDV ^{5min} _{Log–ACD,db}	4.383	0.268	13.576	0.189	2.993	0.223	4.038	0.384	3.052	0.234
PDV ^{5min} _{Log–ACD–Burr,db}	4.376	0.269	13.455	0.191	2.929	0.223	4.015	0.385	3.063	0.234
PDV ^{Spread} _{Log–ACD,db}	4.089	0.239	12.831	0.181	2.574	0.182	3.705	0.258	2.983	0.227
PDV ^{Spread} _{Log–ACD–Burr,db}	4.076	0.244	12.673	0.183	2.562	0.184	3.701	0.258	2.992	0.231
NPDV ^{5min} _{db}	4.405	0.272	14.039	0.192	3.050	0.225	3.986	0.402	3.095	0.233
NPDV ^{Spread} _{db}	4.298	0.278	13.150	0.203	2.656	0.212	3.771	0.301	3.028	0.253

Continued on next page.

TABLE 3 | Continued.

	GE		IBM		PFE		TSLA		WMT	
	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE
RV ^{5min}	6.168	0.207	2.470	0.199	2.060	0.186	9.221	0.273	2.303	0.244
RV ^{5min} _{ss}	6.204	0.207	2.442	0.198	2.204	0.181	8.846	0.268	2.856	0.239
RK ^{5sec} _{parzen}	6.457	0.207	2.432	0.191	2.222	0.179	8.929	0.264	2.696	0.230
RK ^{5sec} _{cubic}	6.596	0.203	2.371	0.181	2.337	0.172	9.034	0.250	2.975	0.217
RK ^{5sec} _{th2}	6.468	0.185	2.574	0.170	2.461	0.158	8.817	0.229	2.926	0.206
RK ^{5sec} _{bartlett}	6.428	0.185	2.527	0.169	2.420	0.157	8.868	0.229	2.879	0.205
TSRV ^{5sec} _{db}	6.095	0.174	2.809	0.160	2.251	0.149	9.086	0.209	2.314	0.198
MSRV ^{5sec}	5.907	0.387	2.418	0.465	2.121	0.362	9.489	0.412	2.187	0.589
RV ^{5min} _{pa}	5.864	0.557	2.357	0.497	1.985	0.514	10.156	0.701	2.507	0.618
BPV ^{5min} _{db}	5.981	0.243	2.538	0.223	1.986	0.214	8.739	0.311	2.305	0.268
BPV ^{5min} _{ss,db}	6.005	0.234	2.448	0.217	2.166	0.207	8.470	0.300	2.740	0.266
PDV ^{5min} _{FHMD,db}	5.647	0.203	2.333	0.209	2.021	0.196	8.578	0.298	2.241	0.269
PDV ^{5min} _{FHMD-Burr,db}	5.641	0.201	2.332	0.207	2.022	0.193	8.698	0.297	2.236	0.266
PDV ^{Spread} _{FHMD,db}	5.280	0.153	2.180	0.176	1.778	0.142	8.682	0.246	1.994	0.193
PDV ^{Spread} _{FHMD-Burr,db}	5.263	0.151	2.224	0.173	1.781	0.141	8.662	0.241	1.990	0.191
PDV ^{5min} _{MSMD,db}	5.795	0.203	2.374	0.208	2.142	0.195	8.584	0.299	2.372	0.266
PDV ^{5min} _{MSMD-Burr,db}	5.720	0.199	2.418	0.205	2.038	0.192	8.584	0.294	2.273	0.266
PDV ^{Spread} _{MSMD,db}	5.318	0.160	2.207	0.185	1.825	0.150	8.705	0.255	2.013	0.200
PDV ^{Spread} _{MSMD-Burr,db}	5.305	0.157	2.225	0.178	1.815	0.147	8.777	0.251	2.013	0.193
PDV ^{5min} _{ACD,db}	5.618	0.180	2.384	0.203	2.088	0.174	8.544	0.297	2.254	0.257
PDV ^{5min} _{ACD-Burr,db}	5.548	0.181	2.387	0.204	2.085	0.177	8.433	0.297	2.269	0.257
PDV ^{Spread} _{ACD,db}	5.255	0.154	2.106	0.174	1.734	0.147	8.690	0.250	1.970	0.185
PDV ^{Spread} _{ACD-Burr,db}	5.248	0.154	2.126	0.178	1.746	0.144	8.655	0.250	1.973	0.187
PDV ^{5min} _{FIACD,db}	6.219	0.527	2.399	0.737	2.166	0.558	8.845	0.677	2.275	0.623
PDV ^{5min} _{FIACD-Burr,db}	6.053	0.619	2.448	0.782	2.084	0.647	10.979	0.860	2.154	0.741
PDV ^{Spread} _{FIACD,db}	5.570	1.748	2.249	3.157	2.445	14.044	9.795	1.578	2.155	8.194
PDV ^{Spread} _{FIACD-Burr,db}	5.575	1.660	2.233	2.750	2.430	13.399	9.713	1.211	2.138	7.711
PDV ^{5min} _{Log-ACD,db}	5.611	0.180	2.385	0.204	2.084	0.175	8.555	0.297	2.261	0.256
PDV ^{5min} _{Log-ACD-Burr,db}	5.543	0.181	2.385	0.204	2.080	0.177	8.447	0.298	2.274	0.256
PDV ^{Spread} _{Log-ACD,db}	5.269	0.154	2.107	0.175	1.764	0.149	8.699	0.245	1.974	0.189
PDV ^{Spread} _{Log-ACD-Burr,db}	5.269	0.155	2.120	0.179	1.755	0.149	8.676	0.247	1.970	0.185
NPDV ^{5min} _{db}	5.587	0.182	2.462	0.206	2.077	0.180	8.369	0.298	2.262	0.259
NPDV ^{Spread} _{db}	5.354	0.170	2.242	0.196	1.876	0.188	8.768	0.281	2.036	0.237

Note: We report the average RMSE and QLike losses. The RMSE means are multiplied with factor 10,000 for the sake of legibility. For some estimators, we perform a small-sample adjustment, which we abbreviate by “db” in the subscript. The sampling frequency (e.g., 5 min) of every RV estimator is attached to its superscript. A gray-shaded cell indicates that the associated estimate belongs to the 75 % model confidence set. The entries in the first column are as follows: The RV (RV^{Δ}), the subsampled RV (RV_{ss}^{Δ}), time scale RV ($TSRV^{\Delta}$), and multiple time scale RV (MSRV), the realized kernel estimator (RK), the noise-robust RV estimator (RV_{pa}), the realized bipower variation estimator (BPV), parametric variance estimator (PDV), and nonparametric variance estimator (NPDV). Δ denotes the sampling interval.

TABLE 4 | Model confidence set inclusion frequency.

	1-day ahead		1-week ahead		1-month ahead		Total		Σ
	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE	
RV ^{5min}	3	0	1	2	10	10	14	12	26
RV ^{5min} _{ss}	4	2	1	2	10	9	15	13	28
RK ^{5sec} _{parzen}	4	3	2	2	10	9	16	14	30
RK ^{5sec} _{cubic}	4	2	1	2	10	9	15	13	28
RK ^{5sec} _{th2}	4	1	1	2	10	9	15	12	27
RK ^{5sec} _{bartlett}	4	1	1	2	10	9	15	12	27
TSRV ^{5sec} _{db}	3	0	1	2	10	9	14	11	25
MSRV ^{5sec}	3	0	1	1	9	8	13	9	22
RV ^{5min} _{pa}	4	0	3	3	10	9	17	12	29
BPV ^{5min} _{db}	3	1	3	2	10	9	16	12	28
BPV ^{5min} _{ss,db}	4	2	2	3	10	9	16	14	30
PDV ^{5min} _{FHMD,db}	3	1	1	2	10	9	14	12	26
PDV ^{5min} _{FHMD-Burr,db}	4	1	2	2	10	9	16	12	28
PDV ^{Spread} _{FHMD,db}	3	0	1	2	10	10	14	12	26
PDV ^{Spread} _{FHMD-Burr,db}	4	0	3	3	10	10	17	13	30
PDV ^{5min} _{MSMD,db}	3	1	1	2	10	9	14	12	26
PDV ^{5min} _{MSMD-Burr,db}	4	1	1	2	10	9	15	12	27
PDV ^{Spread} _{MSMD,db}	3	0	1	2	10	8	14	10	24
PDV ^{Spread} _{MSMD-Burr,db}	4	0	1	2	10	10	15	12	27
PDV ^{5min} _{ACD,db}	5	1	1	1	10	10	16	12	28
PDV ^{5min} _{ACD-Burr,db}	5	1	1	2	10	10	16	13	29
PDV ^{Spread} _{ACD,db}	10	10	10	8	10	10	30	28	58
PDV ^{Spread} _{ACD-Burr,db}	6	5	5	3	10	10	21	18	39
PDV ^{5min} _{FIACD,db}	2	0	1	1	10	9	13	10	23
PDV ^{5min} _{FIACD-Burr,db}	2	0	1	1	10	9	13	10	23
PDV ^{Spread} _{FIACD,db}	4	1	3	5	10	9	17	15	32
PDV ^{Spread} _{FIACD-Burr,db}	4	1	3	5	10	9	17	15	32
PDV ^{5min} _{Log-ACD,db}	5	1	1	1	10	10	16	12	28
PDV ^{5min} _{Log-ACD-Burr,db}	4	1	1	1	10	10	15	12	27
PDV ^{Spread} _{Log-ACD,db}	9	6	7	7	10	10	26	23	49
PDV ^{Spread} _{Log-ACD-Burr,db}	6	4	7	4	10	10	23	18	41
NPDV ^{5min} _{db}	4	1	2	2	10	10	15	13	28
NPDV ^{Spread} _{db}	3	0	1	1	10	9	14	10	24

Note: For each RV estimator, we report the number of inclusions in the 75% model confidence set across all 10 stocks for both loss functions (RMSE and QLIKE) and for three forecast horizons (1 day, 1-week ahead, 1-month ahead). The entries in the first column are as follows: The realized variance (RV^(Δ)), the subsampled RV (RV ^{Δ} _{ss}), time scale RV (TSRV ^{Δ}), and multiple time scale RV (MSRV), the realized kernel estimator (RK), the noise-robust realized variance estimator (RV_{pa}), the realized bipower variation estimator (BPV), parametric variance estimator (PDV), and nonparametric variance estimator (NPDV). Δ denotes the sampling interval.

TABLE 5 | Out-of-sample relative forecasting performance for WMT.

	1-day ahead		1-week ahead		1-month ahead	
	RMSE	QLIKE	RMSE	QLIKE	RMSE	QLIKE
RV^{5min}	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
RV_{ss}^{5min}	0.9982	0.9786	0.9991	0.9912	0.9998	0.9967
RK_{parzen}^{5sec}	0.9969	0.9748	0.9982	0.9858	0.9995	0.9938
RK_{cubic}^{5sec}	0.9980	0.9767	0.9981	0.9854	0.9991	0.9903
RK_{th2}^{5sec}	0.9971	0.9780	0.9981	0.9865	0.9987	0.9874
$RK_{bartlett}^{5sec}$	0.9968	0.9745	0.9980	0.9849	0.9987	0.9871
$TSRV_{db}^{5sec}$	1.0018	1.0088	1.0000	0.9899	0.9993	0.9856
$MSRV^{5sec}$	1.0024	1.0105	1.0003	0.9924	0.9994	0.9854
RV_{pa}^{5min}	0.9998	1.0056	1.0013	0.9950	1.0007	1.0001
BPV_{db}^{5min}	0.9974	0.9834	0.9997	0.9979	0.9999	0.9962
$BPV_{ss,db}^{5min}$	0.9969	0.9634	0.9988	0.9895	0.9998	0.9963
$PDV_{FHMD,db}^{5min}$	0.9966	0.9707	0.9980	0.9809	0.9988	0.9826
$PDV_{FHMD-Burr,db}^{5min}$	0.9964	0.9703	0.9979	0.9801	0.9988	0.9830
$PDV_{FHMD,db}^{Spread}$	0.9901	0.9431	0.9939	0.9502	0.9987	0.9837
$PDV_{FHMD-Burr,db}^{Spread}$	0.9888	0.9185	0.9931	0.9450	0.9986	0.9830
$PDV_{MSMD,db}^{5min}$	0.9968	0.9708	0.9981	0.9808	0.9987	0.9825
$PDV_{MSMD-Burr,db}^{5min}$	0.9990	0.9790	0.9987	0.9815	0.9990	0.9834
$PDV_{MSMD,db}^{Spread}$	0.9882	0.9233	0.9940	0.9540	0.9991	0.9862
$PDV_{MSMD-Burr,db}^{Spread}$	0.9885	0.9267	0.9938	0.9505	0.9990	0.9841
$PDV_{ACD,db}^{5min}$	0.9958	0.9698	0.9975	0.9781	0.9985	0.9842
$PDV_{ACD-Burr,db}^{5min}$	0.9961	0.9679	0.9976	0.9796	0.9986	0.9844
$PDV_{ACD,db}^{Spread}$	0.9791	0.8760	0.9896	0.9185	0.9977	0.9739
$PDV_{ACD-Burr,db}^{Spread}$	0.9824	0.8846	0.9913	0.9295	0.9981	0.9773
$PDV_{FIACD,db}^{5min}$	1.0114	1.0877	1.0020	1.0075	0.9983	0.9767
$PDV_{FIACD-Burr,db}^{5min}$	1.0156	1.0920	1.0056	1.0147	0.9994	0.9860
$PDV_{FIACD,db}^{Spread}$	1.0197	1.1537	0.9988	0.9737	1.0004	1.0037
$PDV_{FIACD-Burr,db}^{Spread}$	1.0136	1.1121	0.9963	0.9524	0.9998	0.9947
$PDV_{Log-ACD,db}^{5min}$	0.9959	0.9694	0.9975	0.9785	0.9986	0.9845
$PDV_{Log-ACD-Burr,db}^{5min}$	0.9962	0.9673	0.9977	0.9800	0.9987	0.9846
$PDV_{Log-ACD,db}^{Spread}$	0.9807	0.8962	0.9896	0.9166	0.9980	0.9770
$PDV_{Log-ACD-Burr,db}^{Spread}$	0.9828	0.8879	0.9911	0.9294	0.9982	0.9787
$NPDV_{db}^{5min}$	0.9970	0.9729	0.9983	0.9837	0.9988	0.9859
$NPDV_{db}^{Spread}$	0.9917	0.9580	0.9964	0.9784	0.9989	0.9863

Note: We report ratios of the average RMSE and QLike losses of all RV measures relative to the RV^{5min} measure. A gray-shaded cell indicates that the associated estimate belongs to the 75% model confidence set. The entries in the first column are as follows: The realized variance ($RV^{(\Delta)}$), the subsampled RV (RV_{ss}^{Δ}), time scale RV ($TSRV^{\Delta}$), and multiple time scale RV ($MSRV$), the realized kernel estimator (RK), the noise-robust realized variance estimator (RV_{pa}), the realized bipower variation estimator (BPV), parametric variance estimator (PDV), and nonparametric variance estimator (NPDV). Δ denotes the sampling interval.

introduced price duration-based estimators in the context of a simulation study, we compare their estimation accuracy in a real-world application. We overcome the fact that the object of interest is not observable, even not ex post, by applying the data-based ranking methodology of Patton (2011a). The approach requires an unbiased but not necessarily precise proxy for the integrated variance. The RV estimator based on 5-min data is an example for a suitable proxy. However, since the proxy and the estimators share the same data basis, the estimation errors are likely to be correlated. Therefore, Patton (2011a) proposes to employ a 1-day lead of the proxy as an instrument to break the correlation.

We measure the estimation accuracy using two well-established loss functions, the root mean squared error (RMSE),

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (\theta_t - M_{i,t})^2}, \quad (13)$$

and the quasi-likelihood (QLike) that is defined by

$$\text{QLike} = \frac{1}{T} \sum_{t=1}^T \left(\frac{\theta_t}{M_{i,t}} - \ln \left(\frac{\theta_t}{M_{i,t}} \right) - 1 \right), \quad (14)$$

where θ_t displays the instrument at day t and $M_{i,t}$ denotes the i -th RV estimator at day t . This setup enables us to obtain a consistent estimate of the difference in the accuracy between two competing estimators, so that an iterative testing procedure, such as the model confidence set (MCS) approach of Hansen et al. (2011), can be applied. In fact, we assess the forecast performance of our proposed return- and duration-based RV estimators and infer whether the differences observed between their loss functions are statistically significant at a given confidence level using the MCS test. The MCS test allows us to obtain from an initial set of competing return- and duration-based RV estimators, without a predefined benchmark model, a set of superior RV estimators at a given forecasting horizon with a given confidence level. Based on both expected loss functions, return- and duration-based RV estimators are ranked, and the worst performing RV estimator at each step is eliminated. The sequential elimination is performed until

the null hypothesis of equal loss differentials for all RV estimators cannot be rejected.

In our in-sample analysis, we estimate the parameters underlying the parametric approaches on a monthly basis and use the 1-day lead of the 5-min RV estimator as an instrument. In Table 3, we report the values for the RMSE and QLike criteria for all stocks and mark estimators that belong to the 75%-model confidence set. Depending on the considered loss function, the results differ substantially. In general, the sets of surviving models under the RMSE loss function are less sparse compared to the model confidence sets under the QLike criterion. The final sets are also found to be different in terms of their composition. Thus, for the RMSE loss function, parametric approaches dominate the model confidence sets, which can be seen by the fact that the estimates of daily volatility based on the ACD(-Burr) model using spread data are always part of the model confidence sets for all stocks. Additionally, the FHMD-Burr as well as the Log-ACD-Burr model also relying on spread data are included in 8 out of 10 cases, whereas the most successful return-based approaches (BPV^{5min}_{ss,db} and BPV^{5min}_{ss,db}) only appear in three model confidence sets. The good performance of the parametric price duration-based estimators is also reflected in the smallest average losses across all stocks. With 4.58×10^{-4} , the ACD-Burr (spread data) has

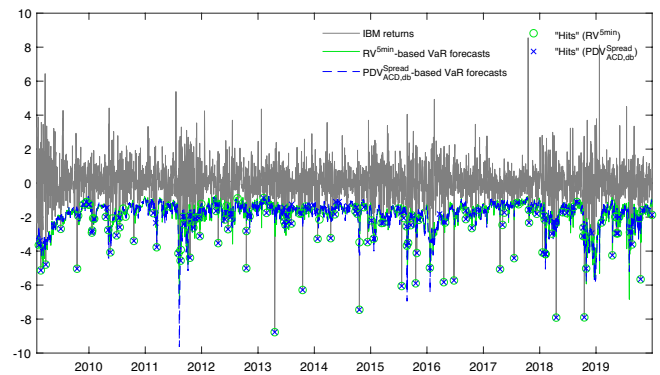


FIGURE 6 | 5%-VaR forecasts for IBM returns using RV^{5min} and $PDV^{Spread}_{ACD,db}$.

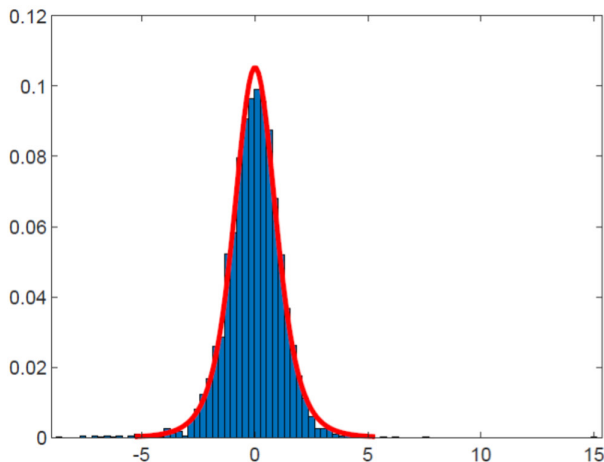
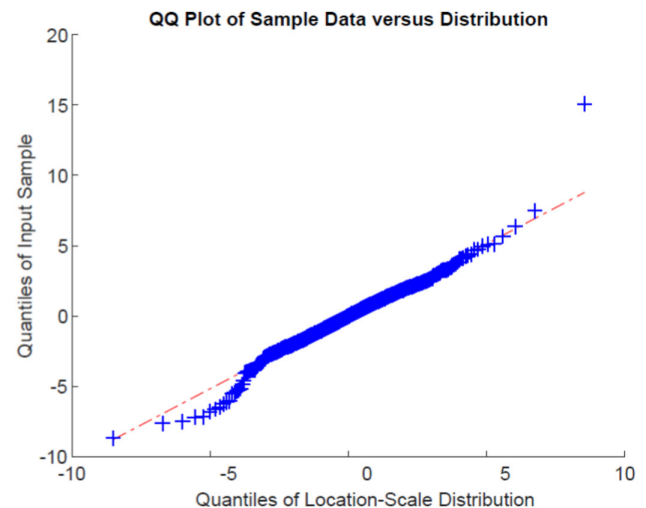


FIGURE 5 | Histogram of IBM standardized returns and Student-t Q-Q IBM standardized returns plot.



the lowest mean loss across all stocks, closely followed by the FHMD–Burr model (spread data) with an average loss of 4.63×10^{-4} .

However, under the QLike criterion, the composition of the model confidence sets changes. The two-scale RV estimator ($\text{TSRV}_{\text{db}}^{\text{5sec}}$) dominates in this situation, since it is part of the surviving set for seven stocks and also exhibits the smallest average loss across all stocks with 0.19. The FHMD–Burr model on spread data appears the second most frequently in the model

confidence sets and also has the fourth lowest average loss with 0.20. Interestingly, the performance of the two-scale RV estimator in terms of the QLike criterion is in stark contrast to its performance under the RMSE loss function, where it is only present in two model confidence sets and has the fourth highest average RMSE-loss across all stocks.

Moreover, we can ascertain that the performance of the duration-based estimators is better, when using spread data compared to 5-min data. This can be explained by the fact that the bid/ask

TABLE 6 | Out-of-sample Likelihood ratio test results for the VaR predictions of all 10 stocks at the $\alpha = 5\%$ confidence level.

	IBM	BAC	GE	AAPL	DIS	EVRG	FB	PFE	TSLA	WMT
GARCH										
EFV	0.042	0.041	0.045	0.038	0.041	0.040	0.048	0.039	0.051	0.035
LR-UC	0.043	0.021	0.228	0.003	0.027	0.016	0.729	0.007	0.804	0.000
LR-IND	0.072	0.000	0.000	0.065	0.059	0.018	0.052	0.012	0.079	0.359
LR-CC	0.025	0.000	0.000	0.002	0.015	0.003	0.142	0.001	0.207	0.000
RV										
EFV	0.040	0.04	0.039	0.039	0.044	0.048	0.037	0.043	0.045	0.040
LR-UC	0.010	0.001	0.007	0.007	0.140	0.625	0.016	0.065	0.352	0.013
LR-IND	0.096	0.236	0.088	0.199	0.466	0.155	0.918	0.088	0.559	0.770
LR-CC	0.009	0.001	0.006	0.012	0.258	0.323	0.053	0.043	0.546	0.043
TSRV										
EFV	0.036	0.034	0.039	0.038	0.037	0.047	0.034	0.042	0.049	0.038
LR-UC	0.000	0.000	0.007	0.003	0.002	0.505	0.005	0.043	0.861	0.003
LR-IND	0.092	0.316	0.708	0.333	0.940	0.131	0.785	0.162	0.117	0.333
LR-CC	0.001	0.000	0.026	0.008	0.007	0.256	0.018	0.048	0.287	0.008
BPV										
EFV	0.036	0.035	0.040	0.036	0.040	0.045	0.035	0.040	0.045	0.040
LR-UC	0.001	0.000	0.016	0.000	0.013	0.265	0.007	0.016	0.352	0.016
LR-IND	0.481	0.174	0.248	0.455	0.841	0.175	0.829	0.248	0.559	0.801
LR-CC	0.002	0.000	0.029	0.001	0.044	0.215	0.026	0.029	0.546	0.054
FHMD										
EFV	0.038	0.032	0.039	0.041	0.042	0.048	0.035	0.041	0.046	0.040
LR-UC	0.003	0.000	0.006	0.034	0.043	0.689	0.007	0.021	0.476	0.010
LR-IND	0.157	0.572	0.678	0.149	0.693	0.037	0.829	0.053	0.158	0.739
LR-CC	0.005	0.000	0.019	0.037	0.118	0.105	0.026	0.011	0.286	0.033
ACD										
EFV	0.039	0.034	0.040	0.041	0.042	0.049	0.035	0.041	0.048	0.041
LR-UC	0.007	0.000	0.016	0.027	0.043	0.823	0.007	0.021	0.697	0.021
LR-IND	0.088	0.902	0.475	0.286	0.693	0.046	0.829	0.126	0.211	0.832
LR-CC	0.006	0.000	0.044	0.049	0.118	0.134	0.026	0.022	0.424	0.068

Note: EFV denotes the ratio of VaR violations to the out-of-sample size. The entries are the p -values of the likelihood ratio test for the unconditional coverage (UC) hypothesis, independence (IND) hypothesis, and conditional coverage (CC) hypothesis. The out-of-sample evaluation time period stretches from December 2008 until December 2019.

spread adjustments seem to reflect the trading intensity and the volatility level in the market properly. Thus, the bid/ask spread-based price durations are more informative and robust than the 5-min-based price durations.

3.3 | Out-of-Sample Analysis

Next, we investigate whether accurate estimates translate into precise forecasts. In order to construct individual variance

forecasts for each estimator, we need to employ a time series model that reflects the dynamics of the IV estimates. Owing to the strong persistence of the RV measures, we base our predictions on the heterogeneous autoregressive (HAR) model of Corsi (2009) that accommodates the long-memory feature of RV:

$$\theta_{t+h} = \beta_{0,i,h} + \beta_{1,i,h}M_{i,t} + \beta_{2,i,h}\frac{1}{5}\sum_{l=0}^4M_{i,t-l} + \beta_{3,i,h}\frac{1}{22}\sum_{l=0}^{21}M_{i,t-l} + u_{i,t}. \quad (15)$$

TABLE 7 | Out-of-sample Likelihood ratio test results for the VaR predictions of all 10 stocks at the $\alpha = 1\%$ coverage rate.

	IBM	BAC	GE	AAPL	DIS	EVRG	FB	PFE	TSLA	WMT
GARCH										
EFV	0.005	0.007	0.004	0.003	0.008	0.004	0.009	0.003	0.008	0.004
LR-UC	0.002	0.052	0.000	0.000	0.193	0.001	0.589	0.000	0.370	0.000
LR-IND	0.725	0.000	0.787	0.850	0.570	0.041	0.089	0.829	0.108	0.787
LR-CC	0.008	0.000	0.001	0.000	0.365	0.001	0.204	0.000	0.184	0.001
RV										
EFV	0.007	0.005	0.003	0.007	0.005	0.009	0.005	0.004	0.008	0.007
LR-UC	0.084	0.004	0.000	0.052	0.004	0.492	0.038	0.001	0.370	0.130
LR-IND	0.607	0.705	0.829	0.107	0.705	0.206	0.791	0.041	0.108	0.008
LR-CC	0.197	0.016	0.000	0.041	0.016	0.355	0.113	0.001	0.184	0.009
TSRV										
EFV	0.006	0.004	0.002	0.006	0.004	0.011	0.004	0.003	0.009	0.005
LR-UC	0.017	0.000	0.000	0.017	0.001	0.512	0.016	0.000	0.516	0.002
LR-IND	0.665	0.766	0.871	0.665	0.746	0.401	0.820	0.850	0.125	0.050
LR-CC	0.052	0.002	0.000	0.052	0.004	0.567	0.053	0.000	0.250	0.001
BPV										
EFV	0.007	0.0044	0.003	0.006	0.004	0.010	0.005	0.003	0.008	0.006
LR-UC	0.084	0.0008	0.000	0.030	0.000	0.922	0.038	0.000	0.370	0.017
LR-IND	0.607	0.746	0.829	0.646	0.766	0.464	0.791	0.850	0.108	0.081
LR-CC	0.197	0.004	0.000	0.086	0.002	0.762	0.113	0.000	0.184	0.013
FHMD										
EFV	0.008	0.005	0.002	0.007	0.004	0.012	0.004	0.003	0.009	0.007
LR-UC	0.274	0.002	0.000	0.084	0.000	0.308	0.016	0.000	0.516	0.052
LR-IND	0.552	0.050	0.871	0.607	0.787	0.371	0.820	0.808	0.125	0.005
LR-CC	0.460	0.001	0.000	0.197	0.001	0.398	0.053	0.000	0.250	0.003
ACD										
EFV	0.008	0.004	0.003	0.007	0.004	0.013	0.004	0.003	0.009	0.007
LR-UC	0.374	0.000	0.000	0.084	0.000	0.169	0.016	0.000	0.516	0.084
LR-IND	0.533	0.766	0.850	0.607	0.787	0.342	0.820	0.850	0.600	0.006
LR-CC	0.555	0.002	0.000	0.197	0.001	0.247	0.053	0.000	0.705	0.005

Note: EFV denotes the ratio of VaR violations to the out-of-sample size. The entries are the p -values of the likelihood ratio test for the unconditional coverage (LR-UC) hypothesis, independence (LR-IND) hypothesis, and conditional coverage (LR-CC) hypothesis. The out-of-sample evaluation time period stretches from December 2008 until December 2019.

We estimate the HAR model based on a rolling window of the 500 most recent observations for three forecast horizons, $h = 1, 5, 22$, that represent forecasts one day/week/month ahead. In line with Clements and Preve (2021), we use a weighted least squares estimation technique¹⁰ to enhance the predictive accuracy. Again, we employ the 5-min RV estimates as the variance proxy and use the RMSE and QLike criteria to assess the forecast performance following Patton (2011b).

To rule out any look-ahead bias, we estimate the duration-based models day-by-day using the 21 most recent days of intraday observations. In Table 4, we summarize the results of the MCS test and report how often a particular model is part of the final surviving set. Across all stocks and for each forecasting horizon, the parametric duration estimator $PDV_{ACD, db}^{Spread}$ exhibits the highest precision in terms of the RMSE and QLike criteria in most cases; see, for example, Table 5 for WMT stock.¹¹ Especially, in terms of the QLike

TABLE 8 | The Expected Future Value (EFV) of the (unconditional) probability of a VaR hit for all 10 stocks.

Models	IBM	BAC	GE	AAPL	DIS	EVRG	FB	PFE	TSLA	WMT
5-day ahead forecasting horizons										
Coverage rate=5%										
GARCH	0.069	0.111	0.109	0.101	0.067	0.053	0.082	0.068	0.180	0.046
RV	0.072	0.070	0.075	0.072	0.060	0.054	0.050	0.076	0.081	0.074
TSRV	0.070	0.068	0.077	0.069	0.062	0.053	0.052	0.076	0.080	0.074
BPV	0.071	0.068	0.075	0.069	0.060	0.053	0.050	0.078	0.076	0.073
FHMD	0.069	0.066	0.076	0.071	0.061	0.051	0.051	0.078	0.079	0.074
ACD	0.069	0.068	0.077	0.071	0.063	0.052	0.051	0.077	0.078	0.076
Coverage rate=1%										
GARCH	0.020	0.027	0.009	0.039	0.026	0.053	0.041	0.036	0.026	0.031
RV	0.016	0.045	0.024	0.028	0.023	0.020	0.018	0.027	0.079	0.028
TSRV	0.016	0.044	0.026	0.026	0.022	0.017	0.019	0.027	0.079	0.027
BPV	0.016	0.047	0.024	0.026	0.023	0.019	0.020	0.028	0.075	0.028
FHMD	0.017	0.045	0.027	0.026	0.023	0.019	0.020	0.027	0.077	0.026
ACD	0.011	0.046	0.028	0.026	0.025	0.011	0.011	0.026	0.077	0.026
20-day ahead forecasting horizons										
Coverage rate=5%										
GARCH	0.123	0.105	0.114	0.097	0.094	0.086	0.109	0.093	0.069	0.092
RV	0.070	0.069	0.076	0.078	0.055	0.060	0.053	0.066	0.042	0.075
TSRV	0.071	0.068	0.075	0.080	0.055	0.058	0.053	0.067	0.044	0.075
BPV	0.069	0.068	0.076	0.078	0.054	0.062	0.056	0.064	0.041	0.076
FHMD	0.072	0.067	0.079	0.079	0.057	0.062	0.053	0.067	0.042	0.076
ACD	0.072	0.068	0.076	0.081	0.054	0.062	0.053	0.067	0.042	0.076
Coverage rate=1%										
GARCH	0.067	0.045	0.033	0.033	0.046	0.046	0.029	0.031	0.033	0.078
RV	0.023	0.031	0.018	0.014	0.021	0.017	0.023	0.018	0.030	0.019
TSRV	0.023	0.031	0.018	0.015	0.020	0.015	0.024	0.017	0.032	0.020
BPV	0.022	0.030	0.019	0.013	0.021	0.017	0.024	0.018	0.032	0.020
FHMD	0.024	0.029	0.017	0.014	0.021	0.018	0.021	0.018	0.030	0.021
ACD	0.024	0.029	0.012	0.014	0.021	0.019	0.022	0.014	0.030	0.020

Note: The entries are the expected future values (EFVs) of the (unconditional) probability of a VaR hit (estimated by the ratio of the VaR hits counted in relation to the sample size for each stock). Boldface numbers denote the cases where the null hypothesis of the likelihood ratio test for the unconditional coverage (LR-UC) hypothesis, independence (LR-IND) hypothesis, and conditional coverage (LR-CC) cannot be rejected. All the models produce independent value-at-risk forecasts. The out-of-sample evaluation time period stretches from December 2008 until December 2019.

criterion, we find large statistically significant improvements up to 12% compared to the benchmark model RV^{5min} , see Table 5. This is also reflected by the highest number of inclusions in the model confidence sets. In 58 out of 60 cases, the $PDV_{ACD,db}^{Spread}$ belongs to the final set, followed by the Log-ACD with 49 out of 60 cases. The best return-based method, the $BPV_{ss,db}^{5min}$, follows at a wide gap with only 30 inclusions. The other duration-based approaches also do not appear frequently in the model confidence set; for example, the $PDV_{FHMD-Burr,db}^{Spread}$ has a total of 30 inclusions. In general, we find that the forecasting performance of the parametric duration-based approaches hinges on the specification underlying the PDV estimator. Moreover, the accuracy of all parametric duration methods is once again higher when they are based on spread-data instead of 5-min data.

3.4 | Value-at-Risk Forecasting

Without loss of generality, financial returns can be formalized as

$$r_t = \mu + \epsilon_t, \quad \epsilon_t = \sigma_t u_t, \quad (16)$$

where $u_t \stackrel{iid}{\sim} t_\nu(0, 1)$ and σ_t^2 follow a standard GARCH(1,1):

$$\sigma_t^2 = \beta_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (17)$$

where $\beta_0 > 0$, $\alpha_1 > 0$ and $\beta_1 > 0$ and $\alpha_1 + \beta_1 < 1$. The nonnegativity constraints on β_0 , α_1 , and β_1 guarantee the positivity of σ_t^2 .

Based on this location-scale family setting, the resulting one-step-ahead $100\alpha\%$ VaR conditional on the information set available at time $t-1$ can be computed by

$$VaR_{t|t-1}^\alpha = \hat{\mu} + \hat{\sigma}_t F^{-1}(\alpha), \quad (18)$$

where $F^{-1}(\alpha)$ displays the α -quantile of the distribution of u_t and $\hat{\sigma}_t$ is the one-step ahead volatility forecast.

Some recent approaches employ intraday data in the form of different RV measures for VaR forecasting with promising results, like inter alia Giot and Laurent (2004), Clements et al. (2008) or Fuentes and Olmo (2013). Given the success of the parametric price duration-based realized measures of integrated variance in the in-sample as well as in the out-of-sample analysis, we examine whether they can help to produce more accurate and valid VaR forecasts.

TABLE 9 | Out-of-sample VaR 1-day ahead forecasts evaluated by tick loss function.

		GARCH	RV^{5min}	$TSRV_{db}^{5sec}$	$BPV_{ss,db}^{5min}$	$PDV_{FHMD-Burr,db}^{Spread}$	$PDV_{ACD,db}^{Spread}$
AAPL	5%	0.1956***	0.1820**	0.1815**	0.1817**	0.1813***	0.1801
	1%	0.0642***	0.0569	0.0569	0.0569	0.0568	0.0567
BAC	5%	0.3110***	0.2546	0.2514	0.2539	0.2565***	0.2544
	1%	0.1123***	0.0786	0.0767	0.0781	0.0768**	0.0756
DIS	5%	0.1646***	0.1502**	0.1494***	0.1484	0.1482**	0.1473
	1%	0.0580***	0.0459***	0.0454***	0.0449**	0.0440	0.0438
EVRG	5%	0.1259***	0.1212	0.1212**	0.1207	0.1205***	0.1197
	1%	0.0428***	0.0382	0.0386**	0.0379	0.0379**	0.0374
FB	5%	0.2147***	0.2027***	0.2015**	0.2020**	0.2001	0.1997
	1%	0.0796**	0.0714	0.0718**	0.0713	0.0706	0.0705
GE	5%	0.2016***	0.1811	0.1772	0.1802	0.1794	0.1792
	1%	0.0631***	0.0529	0.0517††	0.0523	0.0534	0.0532
IBM	5%	0.1576***	0.1431	0.1429**	0.1424	0.1422*	0.1416
	1%	0.0568***	0.0515	0.0517	0.0517	0.0517***	0.0510
PFE	5%	0.1353***	0.1288***	0.1273	0.1286***	0.1276**	0.1261
	1%	0.0404***	0.0366**	0.0353	0.0363**	0.0362***	0.0351
TSLA	5%	0.3525	0.3507	0.3534**	0.3507	0.3516	0.3508
	1%	0.1185	0.1148	0.1158	0.1145	0.1143	0.1142
WMT	5%	0.1281***	0.1230**	0.1226**	0.1226**	0.1225***	0.1212
	1%	0.0460***	0.0416**	0.0411	0.0408	0.0409***	0.0400

Note: We report the average tick loss function values for 5%- and 1% -VaR 1-day ahead forecast. Boldface numbers denote the lowest average tick loss values for every considered stock. We test for significant improvements (deteriorations) via the Diebold–Mariano test and report the significance levels 1%, 5%, and 10% by *** (††), ** (†), and * (†).

For incorporating RV measures in the VaR predictions, we need to have (i) a law of motion for the RV dynamics and (ii) a mapping between the conditional expectation of the RV measure and the conditional variance of the returns; see Brownlees and Gallo (2010). Most approaches rely on a two-step procedure that builds up on, for example, a HAR or ARFIMA(X) model; see Giot and Laurent (2004) and Clements et al. (2008) in the first instance and relate the conditional RV prediction with the conditional variance by means of a linear function in a consecutive step. A more novel approach of Maheu and McCurdy (2011) combines the two steps into one by using a bivariate model. Owing to its success in the closely related field of return density forecasting, we follow their approach and model the return and the RV simultaneously in the following way:¹²

$$\begin{aligned} r_t &= \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t u_t, \text{ where } u_t \sim t_v(0, 1), \\ \log(RV_t) &= \omega + \phi_1 \log(RV_{t-1}) + \phi_2 \log(RV_{t-5,5}) \\ &\quad + \phi_3 \log(RV_{t-22,22}) + v_t, \text{ where } v_t \sim \mathcal{N}(0, 1). \end{aligned} \quad (19)$$

The authors argue that—under certain empirically realistic conditions—the conditional expectation of an unbiased RV measure is equal to the conditional variance of returns, that is, $\mathbb{E}_{t-1}(RV_t) = \text{Var}_{t-1}(r_t) = \sigma_t^2$. By that, they establish a linkage between these two measures. Due to the log-normal

distribution of the RV measure in this specification, the connection becomes

$$\sigma_t^2 = \mathbb{E}_{t-1}(RV_t) = \exp(\mathbb{E}_{t-1}(\log(RV_t)) + 0.5 \text{Var}_{t-1}(\log(RV_t))). \quad (20)$$

Based on this framework, we employ duration-based as well as return-based RV measures for our VaR forecasting exercise. In our analysis, we concentrate on the most successful RV measures from Sections 3.2 and 3.3. Relying on the $\text{PDV}_{\text{FHMD-Burr,db}}^{\text{Spread}}$ and the $\text{PDV}_{\text{ACD,db}}^{\text{Spread}}$ as duration-based RV measures and on the $\text{RV}^{5\text{min}}$, the $\text{TSRV}_{\text{db}}^{5\text{sec}}$ as well as the $\text{BPV}_{\text{ss,db}}^{5\text{min}}$ as return-based benchmarks; we estimate the bivariate model via maximum-likelihood and project the conditional return variance 1-day ahead to obtain a VaR forecast. The distributional assumption for the standardized returns in the location-scale framework is motivated by our empirical results indicating that the Student- t is the appropriate; see Figure 5. Doing so for a rolling window of 500 observations, we receive a series of 2752 VaR predictions for the time span of December 2008 until December 2019.¹³

An exemplary illustration of our VaR predictions using the $\text{RV}^{5\text{min}}$ and $\text{PDV}_{\text{ACD,db}}^{\text{Spread}}$ measures as components for the bivariate model for IBM against its returns in Figure 6 reveals that our forecasts sometimes fail to quantify the maximum

TABLE 10 | Out-of-sample VaR 5-day ahead forecasts evaluated by tick loss function.

		GARCH	RV^{5min}	TSRV^{5sec}_{db}	BPV^{5min}_{ss,db}	PDV^{Spread}_{FHMD-Burr,db}	PDV^{Spread}_{ACD,db}
AAPL	5%	0.8538***	0.4554***	0.4523**	0.4545*	0.4478	0.4478
	1%	0.1625***	0.1432	0.1425*	0.1446	0.1394	0.1393
BAC	5%	1.0092**	0.7378 ††	0.7452††	0.7416††	0.7981	0.8004
	1%	0.1433 †††	0.2760	0.2762	0.2758	0.2863	0.2885
DIS	5%	0.7421***	0.3886	0.3869	0.3898	0.3911	0.3868
	1%	0.0588 †††	0.1299	0.1284†	0.1303	0.1307	0.1317
EVRG	5%	0.6385***	0.2990	0.2940	0.2989	0.2955	0.2940
	1%	0.0403 †††	0.0970	0.0940	0.0954	0.0935	0.0939
FB	5%	0.7717***	0.5192***	0.5137***	0.5179***	0.5041	0.5040
	1%	0.1705 ††	0.1875†	0.1864††	0.1885†	0.1824	0.1822
GE	5%	0.9332***	0.5550 ††	0.5624††	0.5654	0.5799	0.5806
	1%	0.1005 †††	0.1832††	0.1868	0.1884	0.1932	0.1927
IBM	5%	0.7529***	0.3801**	0.3756	0.3796*	0.3739	0.3739
	1%	0.1417***	0.1265*	0.1233	0.1272*	0.1228	0.1228
PFE	5%	0.7019***	0.3444	0.3408 †	0.3450	0.3454	0.3455
	1%	0.1421***	0.1254	0.1233	0.1253	0.1254	0.1246
TSLA	5%	1.1854***	0.8169	0.8143	0.8159	0.8148	0.8143
	1%	0.1615 †††	0.3745	0.3788†	0.3726	0.3696	0.3701
WMT	5%	0.6317***	0.2990**	0.2962	0.2970	0.2922 †	0.2939
	1%	0.1331***	0.1071	0.1062	0.1053	0.1036	0.1043

Note: We report the average tick loss function values for 5%- and 1% -VaR 5-day ahead forecast. Boldface numbers denote the lowest average tick loss values for every considered stock. We test for significant improvements (deteriorations) via the Diebold–Mariano test and report the significance levels 1%, 5%, and 10% by *** (†††), ** (††), and * (†).

loss that will not be exceeded with 95% probability and the returns occasionally exceed (or “hit”) the threshold set by the forecasted VaR. To make inference about the validity of our VaR predictions, we adopt the likelihood ratio (LR) test of Christoffersen (1998) that is designed on a transition matrix that governs a two-states Markov chain model that describes the behavior of the violations of the VaR predictions. In other words, a VaR model is valid if the violation process satisfies the martingale difference hypothesis, that is, (i) the expected frequency of VaR violations or hits is equal to the nominal coverage rate α and (ii) the VaR violations are distributed independently. The p -values of the LR for 1-day ahead VaR forecasts are reported in Tables 6 and 7. For all the VaR models, the unconditional probability of violations is significantly smaller than the coverage rate (5% or 1%). This implies that all our VaR models in most cases tend to overestimate the level of market risk. Based on the LR-IND and LR-CC results, it seems that price duration-based and returns-based RV estimators are more capable to produce violations that are distributed independently than the standard GARCH(1,1). For 5- and 20-day VaR forecasts, we report the unconditional probability of violations; see Table 8. In general, all the models show the tendency to underestimate the level of market risks at 5- and 20-day forecasting horizons but produce in most

cases independent VaR forecasts for all stocks. However, for some stocks, the duration-based RV estimators produce valid VaR forecasts at both forecasting horizons and at both coverage rates.

In Figure 6, we see that for both approaches, the VaR violations often coincide. Furthermore, the total numbers of “hits” (137 and 140 for RV^{5min} and $PDV_{FHMD-Burr,db}^{Spread}$, respectively) are close to the expected number of 137.55.

However, for a risk manager, the number of “hits” is an uninformative indicator because it does not reflect the trade-off between the costs of overpredictions and underpredictions (regulatory penalty vs. opportunity cost of capital). In order to account for these conflicting targets, we compare the forecasting accuracy on the basis of the asymmetric tick-loss function of Giacomini and Komunjer (2005), which is given by

$$TL_{\alpha}(r_{t+1}, VaR_{t+1|t}) = \left(\alpha - \mathbb{1}_{\{r_{t+1} < VaR_{t+1|t}\}} \right) (r_{t+1} - VaR_{t+1|t}^{\alpha}). \quad (21)$$

For 5%- as well as 1%-VaR forecasts, we present the average tick-loss function values for every stock in Tables 9–11. For each row, we highlight the best forecast in boldface numbers.

TABLE 11 | Out-of-sample VaR 20-day ahead forecasts evaluated by tick loss function.

		GARCH	RV^{5min}	$TSRV_{db}^{5sec}$	$BPV_{ss,db}^{5min}$	$PDV_{FHMD-Burr,db}^{Spread}$	$PDV_{ACD,db}^{Spread}$
AAPL	5%	1.0828*	0.9545	0.9612*	0.9580	0.9551	0.9545
	1%	0.1863 †††	0.2673	0.2710	0.2699	0.2710	0.2693
BAC	5%	1.3440	1.5885**	1.5835**	1.5768**	1.3440	1.3438
	1%	0.3305 ††	0.5139	0.5112	0.5101	0.5225	0.5217
DIS	5%	0.9522*	0.7775	0.7745	0.7798	0.7814	0.7744
	1%	0.1575 ***	0.2415	0.2407	0.2433	0.2434	0.2424
EVRG	5%	0.8822***	0.5778	0.5707	0.5804	0.5757	0.5706
	1%	0.1020 †††	0.1676	0.1643	0.1683	0.1656	0.1663
FB	5%	1.0054	0.9449	0.9466	0.9500	0.9313	0.9312
	1%	0.1370 †††	0.3027	0.3054**	0.3054**	0.2957*	0.2977
GE	5%	1.3261	1.3015	1.3077	1.3133	1.3239	1.3015
	1%	0.2976 †††	0.4751	0.4977	0.4969	0.4838	0.4925
IBM	5%	1.1180***	0.7728	0.7734	0.7720	0.7776	0.7719
	1%	0.1058 ††	0.2643	0.2634	0.2639	0.2645	0.2644
PFE	5%	0.9871***	0.6288	0.6269	0.6316	0.6269	0.6269
	1%	0.0979 ††	0.2228	0.2197	0.2247	0.2206	0.2214
TSLA	5%	1.5662	1.4662	1.4610	1.4679	1.4685	1.4609
	1%	0.3211	0.4135	0.4155	0.4152	0.4176	0.4180
WMT	5%	0.9700***	0.6042	0.6071	0.6054	0.6030	0.6030
	1%	0.0912 †††	0.1937	0.1923†	0.1961	0.1955**	0.1937

Note: We report the average tick loss function values for 5%- and 1%-VaR 20-day ahead forecast. Boldface numbers denote the lowest average tick loss values for every considered stock. We test for significant improvements (deteriorations) via the Diebold–Mariano test and report the significance levels 1%, 5%, and 10% by *** (†††), ** (††), and * (†).

We note that the asymmetric tick-loss function defined in (21) is not differentiable due to the presence of the indicator function. The nondifferentiability of the asymmetric tick-loss function may cause a problem in the implementation of the Diebold and Mariano (1995) test. However, as pointed out in Granger (1999), the nondifferentiability issue may be just a technicality due to the fact that it should always be possible to find a smooth function which is arbitrarily close to the non-smooth one. To test for the predictive power of the duration-based RV measure stemming from the ACD model, we conduct pairwise comparisons via the Diebold and Mariano (1995) test (in the framework of Giacomini and White 2006). We display significant outperformance (underperformance) by *** (†††), ** (††), and * (†) for the 1%, 5%, and 10% significance levels. Apart from our intraday-based RV measures, we opt for the classical location-scale family approach in combination with a GARCH(1, 1) model as described in Equation (17) as a benchmark.

According to Table 9, the predictions of 1-day VaR values based on the $PDV_{ACD,db}^{Spread}$ measure perform best in terms of the tick-loss function for almost every stock and confidence level. In comparison to other forecasts, we find that the $PDV_{ACD,db}^{Spread}$ -based forecasting approach is significantly better than the other forecasting methods in the vast majority of cases. In particular, compared to GARCH model-based forecasts, we find

highly significant outperformance in favor of the duration-based approach. This result holds for all other VaR forecasts that rely on high-frequency intraday data, as well. They also perform substantially better than predictions that only employ daily return data (GARCH). Among all competitors, VaR forecasts utilizing the other duration-based variance estimates, $PDV_{FHMD-Burr,db}^{Spread}$, seem to be a viable option. They often perform slightly better than return-based methods. However, in comparison to predictions on the basis of $PDV_{ACD,db}^{Spread}$ estimates, we often find significant inferior forecasting accuracy.

The tick-loss functions for 5- and 20-day VaR values based on both return- and duration-based estimators as well as for the GARCH(1,1) and for both coverage rates (5% and 1%) are reported in Tables 10 and 11. Again, the $PDV_{ACD,db}^{Spread}$ -based forecasting approach outperforms its competitors. However, unlike the return-based RV estimators, the GARCH(1,1) model becomes very competitive to the duration-based RV estimator based on the ACD and performs very well at the 20-day forecasting horizon and at the 1% coverage rate.

3.5 | Realized Utility Benefits

So far, we have shown that the price duration-based RV estimator can relatively outperform return-based ones via RMSE, QLIKE

TABLE 12 | Out-of-sample realized utility (RU) results.

	AAPL			BAC			DIS			EVRG		
	1 day	5 days	20 days	1 day	5 days	20 days	1 day	5 days	20 days	1 day	5 days	20 days
RV ^{5min}	0.0363	0.0374	0.0377	0.0272	0.0282	0.0290	0.0352	0.0361	0.0366	0.0365	0.0371	0.0375
RV ^{5min} _{ss}	0.0361	0.0373	0.0377	0.0273	0.0279	0.0290	0.0353	0.0361	0.0365	0.0365	0.0372	0.0376
BPV ^{5min} _{db}	0.0365	0.0375	0.0378	0.0272	0.0286	0.0292	0.0351	0.0360	0.0365	0.0366	0.0373	0.0375
BPV ^{5min} _{ss,db}	0.0361	0.0373	0.0377	0.0273	0.0279	0.0292	0.0352	0.0360	0.0365	0.0365	0.0372	0.0375
RK ^{5sec} _{bartlett}	0.0362	0.0373	0.0376	0.0272	0.0277	0.0290	0.0352	0.0361	0.0365	0.0366	0.0372	0.0376
RK ^{5sec} _{cubic}	0.0361	0.0373	0.0376	0.0273	0.0277	0.0290	0.0353	0.0361	0.0366	0.0365	0.0372	0.0376
RK ^{5sec} _{parzen}	0.0361	0.0373	0.0376	0.0273	0.0279	0.0290	0.0353	0.0362	0.0366	0.0365	0.0372	0.0376
RK ^{5sec} _{th2}	0.0361	0.0373	0.0376	0.0272	0.0277	0.0290	0.0352	0.0361	0.0365	0.0366	0.0372	0.0376
MSRV ^{5sec}	0.0363	0.0373	0.0377	0.0271	0.0276	0.0289	0.0353	0.0362	0.0366	0.0371	0.0375	0.0377
RV ^{5min} _{pa}	0.0364	0.0374	0.0379	0.0270	0.0274	0.0291	0.0354	0.0360	0.0366	0.0365	0.0372	0.0376
TSRV ^{5sec} _{db}	0.0363	0.0373	0.0377	0.0274	0.0278	0.0290	0.0353	0.0362	0.0366	0.0369	0.0375	0.0377
NPDV ^{Spread} _{db}	0.0360	0.0373	0.0377	0.0256	0.0259	0.0285	0.0348	0.0358	0.0364	0.0367	0.0373	0.0375
PDV ^{Spread} _{ACD-Burr,db}	0.0362	0.0373	0.0377	0.0255	0.0259	0.0285	0.0347	0.0358	0.0363	0.0367	0.0373	0.0375
PDV ^{Spread} _{FHMD-Burr,db}	0.0361	0.0372	0.0376	0.0256	0.0258	0.0285	0.0347	0.0358	0.0363	0.0368	0.0373	0.0375
PDV ^{Spread} _{FIACD-Burr,db}	0.0379	0.0383	0.0381	0.0265	0.0265	0.0290	0.0359	0.0366	0.0369	0.0372	0.0375	0.0376
PDV ^{Spread} _{Log-ACD-Burr,db}	0.0361	0.0373	0.0377	0.0256	0.0259	0.0285	0.0347	0.0357	0.0363	0.0367	0.0373	0.0375
PDV ^{Spread} _{MSMD-Burr,db}	0.0362	0.0373	0.0377	0.0256	0.0258	0.0285	0.0348	0.0358	0.0363	0.0368	0.0373	0.0375

Note: We report the realized utilities. The entries in the first column are as follows: The realized variance (RV^(Δ)), the subsampled RV (RV^(Δ)_{ss}), time scale RV (TSRV^(Δ)), and multiple time scale RV (MSRV), the realized kernel estimator (RK), the noise-robust realized variance estimator (RV^(Δ)_{pa}), the realized bipower variation estimator (BPV), parametric variance estimator (PDV), and nonparametric variance estimator (NPDV). Δ denotes the sampling interval.

TABLE 13 | Out-of-sample realized utility (RU) results.

	FB			GE			IBM		
	1 day	5 days	20 days	1 day	5 days	20 days	1 day	5 days	20 days
RV ^{5min}	0.0363	0.0379	0.0384	0.0307	0.0319	0.0333	0.0366	0.0376	0.0382
RV _{ss} ^{5min}	0.0362	0.0379	0.0384	0.0305	0.0318	0.0333	0.0366	0.0376	0.0382
BPV _{db} ^{5min}	0.0363	0.0380	0.0384	0.0300	0.0317	0.0333	0.0367	0.0377	0.0382
BPV _{ss, db} ^{5min}	0.0362	0.0379	0.0384	0.0301	0.0316	0.0333	0.0366	0.0376	0.0382
RK _{bartlett} ^{5sec}	0.0362	0.0379	0.0384	0.0305	0.0318	0.0333	0.0366	0.0376	0.0382
RK _{cubic} ^{5sec}	0.0362	0.0379	0.0384	0.0305	0.0318	0.0333	0.0366	0.0376	0.0382
RK _{parzen} ^{5sec}	0.0362	0.0379	0.0384	0.0305	0.0318	0.0333	0.0366	0.0376	0.0382
RK _{th2} ^{5sec}	0.0362	0.0379	0.0384	0.0304	0.0318	0.0333	0.0366	0.0376	0.0382
MSRV ^{5sec}	0.0365	0.0381	0.0384	0.0301	0.0315	0.0333	0.0369	0.0378	0.0383
RV _{pa} ^{5min}	0.0361	0.0377	0.0384	0.0303	0.0316	0.0334	0.0367	0.0376	0.0382
TSRV _{db} ^{5sec}	0.0365	0.0381	0.0384	0.0304	0.0318	0.0335	0.0368	0.0378	0.0383
NPDV _{db} ^{Spread}	0.0361	0.0379	0.0384	0.0291	0.0307	0.0329	0.0367	0.0377	0.0382
PDV _{ACD-Burr,db} ^{Spread}	0.0360	0.0378	0.0384	0.0291	0.0307	0.0329	0.0367	0.0377	0.0382
PDV _{FHMD-Burr,db} ^{Spread}	0.0360	0.0378	0.0384	0.0291	0.0307	0.0329	0.0367	0.0377	0.0382
PDV _{FIACD-Burr,db} ^{Spread}	0.0361	0.0378	0.0383	0.0294	0.0310	0.0331	0.0370	0.0378	0.0382
PDV _{Log-ACD-Burr,db} ^{Spread}	0.0360	0.0378	0.0384	0.0291	0.0307	0.0329	0.0366	0.0377	0.0382
PDV _{MSMD-Burr,db} ^{Spread}	0.0360	0.0378	0.0384	0.0290	0.0307	0.0330	0.0367	0.0377	0.0382
	PFE			TSLA			WMT		
	1 day	5 days	20 days	1 day	5 days	20 days	1 day	5 days	20 days
RV ^{5min}	0.0364	0.0372	0.0379	0.0371	0.0382	0.0389	0.0376	0.0383	0.0385
RV _{ss} ^{5min}	0.0365	0.0373	0.0378	0.0371	0.0382	0.0389	0.0376	0.0383	0.0385
BPV _{db} ^{5min}	0.0364	0.0372	0.0378	0.0370	0.0382	0.0389	0.0376	0.0383	0.0385
BPV _{ss, db} ^{5min}	0.0365	0.0372	0.0378	0.0370	0.0382	0.0389	0.0376	0.0383	0.0385
RK _{bartlett} ^{5sec}	0.0366	0.0373	0.0378	0.0372	0.0383	0.0389	0.0376	0.0383	0.0385
RK _{cubic} ^{5sec}	0.0366	0.0373	0.0378	0.0373	0.0383	0.0389	0.0376	0.0383	0.0385
RK _{parzen} ^{5sec}	0.0366	0.0373	0.0378	0.0372	0.0382	0.0389	0.0376	0.0383	0.0385
RK _{th2} ^{5sec}	0.0366	0.0373	0.0378	0.0373	0.0383	0.0389	0.0376	0.0383	0.0385
MSRV ^{5sec}	0.0365	0.0373	0.0378	0.0376	0.0385	0.0389	0.0377	0.0384	0.0386
RV _{pa} ^{5min}	0.0364	0.0372	0.0379	0.0374	0.0383	0.0389	0.0376	0.0382	0.0385
TSRV _{db} ^{5sec}	0.0367	0.0374	0.0379	0.0372	0.0383	0.0389	0.0376	0.0384	0.0386
NPDV _{db} ^{Spread}	0.0362	0.0370	0.0377	0.0370	0.0381	0.0389	0.0374	0.0382	0.0385
PDV _{ACD-Burr,db} ^{Spread}	0.0361	0.0369	0.0377	0.0369	0.0381	0.0389	0.0374	0.0382	0.0385
PDV _{FHMD-Burr,db} ^{Spread}	0.0362	0.0369	0.0377	0.0370	0.0381	0.0389	0.0374	0.0382	0.0385
PDV _{FIACD-Burr,db} ^{Spread}	0.0380	0.0382	0.0382	0.0372	0.0381	0.0388	0.0381	0.0384	0.0385

(Continues)

TABLE 13 | (Continued)

	PFE			TSLA			WMT		
	1 day	5 days	20 days	1 day	5 days	20 days	1 day	5 days	20 days
PDV ^{Spread} Log-ACD-Burr,db	0.0361	0.0369	0.0377	0.0369	0.0381	0.0389	0.0373	0.0381	0.0385
PDV ^{Spread} MSMD-Burr,db	0.0363	0.0370	0.0377	0.0370	0.0381	0.0389	0.0373	0.0382	0.0385

Note: We report the realized utilities. The entries in the first column are as follows: The realized variance (RV^{Δ}), the subsampled RV (RV_{ss}^{Δ}), time scale RV ($TSRV^{\Delta}$), and multiple time scale RV (MSRV), the realized kernel estimator (RK), the noise-robust realized variance estimator (RV_{pa}), the realized bipower variation estimator (BPV), parametric variance estimator (PDV), and non-parametric variance estimator (NPDV). Δ denotes the sampling interval.

and MCS. In this section, we want to evaluate the economic value of price duration-based RV forecasts. We adopt the framework introduced by Bollerslev et al. (2018) in which the expected utility of a mean-variance investor at t can be formalized as follows:

$$\mathbb{E}_t(u(W_{t+1})) = \mathbb{E}_t(W_{t+1}) - \frac{1}{2}\gamma^A \text{Var}_t(W_{t+1}), \quad (22)$$

where $\gamma^A \equiv -u''/u'$ is the absolute risk aversion of the investor. By assuming that the investor allocates at time t a fraction ρ_t of his current wealth (W) to a risky asset with return r_{t+1} and the rest to a risk-free money market account earning r_t^f . With this portfolio, his future wealth becomes $W_{t+1} = W_t(1 + \rho_t r_{t+1} + (1 - \rho_t)r_t^f) = W_t(1 + r_t^f) + W_t \rho_t r_{t+1}^e$, where $r_{t+1}^e \equiv r_{t+1} - r_t^f$ represents the excess return. After dropping constant terms, the expected utility in Equation (22) becomes the following:

$$U(x_t) = W_t \left(\rho_t \mathbb{E}_t \left(r_{t+1}^e - \frac{\gamma}{2} \rho_t^2 \mathbb{E}_t(RV_{t+1}) \right) \right), \quad (23)$$

where $\gamma \equiv \gamma^A W_t$ is the investor's relative risk aversion.

Under the assumption of a constant conditional sharpe ratio (SR), the expected utility can be rewritten as

$$U(x_t) = W_t \left(\rho_t SR \sqrt{\mathbb{E}_t(RV_{t+1})} - \frac{\gamma}{2} \rho_t^2 \mathbb{E}_t(RV_{t+1}) \right), \quad (24)$$

with $SR \equiv \mathbb{E}_t(r_{t+1}^e) / \sqrt{\mathbb{E}_t(RV_{t+1})}$ and it is maximized by investing $\rho_t^* = (SR/\gamma) / (\sqrt{\mathbb{E}_t(RV_{t+1})})$ of wealth to the risky asset and $1 - \rho_t^*$ to a risk-free money market. In order to compute the realized utility gains based on various RV estimators, we replace $\mathbb{E}_t(RV_{t+1})$ with the true and observed RV_{t+1} and we obtain

$$RU_t = \frac{SR^2}{\gamma} \frac{\sqrt{RV_{t+1}}}{\sqrt{\mathbb{E}_t^{\theta}(RV_{t+1})}} - \frac{SR^2}{2\gamma} \frac{RV_{t+1}}{\mathbb{E}_t^{\theta}(RV_{t+1})}, \quad (25)$$

where $\mathbb{E}_t^{\theta}(\cdot)$ denotes the conditional expectation under a specific RV estimator θ .

In line with Bollerslev et al. (2018), we assume that the sharpe ratio (SR) is 0.4 and the coefficient of risk aversion (γ) is 2 and the realized utility benefits are reported in Tables 12 and 13. In an ideal risk model, the predicted realized utility is given by $SR^2/2 = 8\%$. At 20-day forecasting horizon, all the price duration- and return-based RV estimators deliver the highest realized

utility (RU) for TSLA (RU=3.89%) and for FB (RU=3.84%). Overall, our out-of-sample realized utility results show that the price duration-based RV estimators yield higher utilities for many stocks than the return-based ones. The PDV^{Spread}_{FIACD-Burr,db} RV estimator seems to be the best model followed by $TSRV_{db}^{5sec}$.

4 | Conclusion

We compared in this paper the forecasting performance of various parametric price duration-based RV estimators constructed via the (FI)ACD, the Log-ACD, the MSMD and the FHMD models with the performance of the return-based RV estimators. Furthermore, we document the ability of the parametric price duration-based RV estimators to be more robust to microstructure noise and to outperform the GARCH(1,1) in terms of forecasting market risks. In an empirical study using time series data on price duration from January 2007 to December 2019 for 10 stocks on NYSE, we find that (i) the price duration-based approaches, in particular the ACD- and FHMD-Burr-based estimates, exhibit a high estimation precision under the RMSE criterion and (ii) the TSRV measure and the FHMD-based estimator perform best in terms of the QLike loss. The out-of-sample forecasting performances can be confirmed by the realized utilities that are higher for price-duration-based approaches. In addition, the results also show that the ACD(1,1) provides more accurate and valid VaR forecasts than the GARCH(1,1), suggesting that the information content in the irregularly-spaced price duration data can help improve the estimation and forecasting accuracy of the VaR forecasts. These empirical findings suggest that the ACD-Burr-based RV estimators are very robust to microstructure noise and produce more accurate volatility forecasts than return-based RV approaches. Therefore, they can be useful in risk management for high-frequency trading and in derivative pricing.

Acknowledgments

We would like to thank the editor and two anonymous reviewers for their constructive and detailed comments, which have significantly improved the paper. The usual disclaimer applies. Open Access funding enabled and organized by Projekt DEAL.

Ethics Statement

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Endnotes

- ¹ see Pacurar (2008) for a detailed literature review on ACD models.
- ² Andersen et al. (2008) showed that price intensity-based estimators seem to be robust to asset price jumps. Hong et al. (2023) derive the asymptotic theory for the non-parametric price duration-based estimators.
- ³ With jumps in Equation (1), the log-price Y_t can be decomposed into a continuous Brownian component and discontinuous component due to jumps. In this case, the true variance of process Y_t will be equal to the integrated volatility plus an additional term.
- ⁴ Δ denotes here the sampling interval.
- ⁵ We refer the reader to the Supporting Information for more details on the various frameworks.
- ⁶ The choice of the optimal threshold ι is subject to debate in the empirical finance research literature; see Andersen et al. (2008) and Hong et al. (2023) for more details.
- ⁷ Besides this so-called *time-of-the-day-effect* early studies based on duration data from the nineties or the start of the new millennium, like inter alia Engle and Russell (1998) and Bauwens et al. (2004), document a distinct *day-of-the-week-effect*, which does not seem to be present in our data.
- ⁸ We also apply a Nadaraya–Watson kernel regression and spline functions to remove the diurnal pattern. The logarithmic dummy variable regression seems to be the best approach for our data here.
- ⁹ We refrain to display the analogous figure for the spread data, since the seasonality pattern is qualitatively similar.
- ¹⁰ Given the documented bias in the estimated coefficients of the HAR model in the presence of measurement error in RV, we do not address this issue directly, as Bollerslev et al. (2016) do with their HARQ model but employ a weighted least squares (WLS) estimation technique proposed by Clements and Preve (2021) to overcome this deficiency.
- ¹¹ The results of the out-of-sample relative forecasting performance for all 10 stocks are available in the Supporting Information.
- ¹² Other options to characterize the dynamics of the log-RV than the above presented HAR model are also viable and are described by Maheu and McCurdy (2011).
- ¹³ For FB and TSLA the out-of-sample period is reduced.

References

- Ait-Sahalia, Y., I. Kalnina, and D. Xiu. 2020. “High-Frequency Factor Models and Regressions.” *Journal of Econometrics* 216, no. 1: 86–105. <https://doi.org/10.1016/j.jeconom.2020.01.007>.
- Andersen, T., T. Bollerslev, F. Diebold, and H. Ebens. 2001a. “The Distribution of Realized Stock Return Volatility.” *Journal of Financial Economics* 61, no. 1: 43–76. [https://doi.org/10.1016/S0304-405X\(01\)00055-1](https://doi.org/10.1016/S0304-405X(01)00055-1).
- Andersen, T., T. Bollerslev, F. Diebold, and P. Labys. 2001b. “The Distribution of Realized Exchange Rate Volatility.” *Journal of the American Statistical Association* 96, no. 453: 42–55. <https://doi.org/10.1198/016214501750332965>.
- Andersen, T. G., D. P. Dobrev, and E. Schaumburg. 2008. *Duration Based Volatility Estimation*. Manuscript Northwestern University.
- Bandi, F., and J. Russell. 2008. “Microstructure Noise, Realized Variance, and Optimal Sampling.” *Review of Economic Studies* 75, no. 2: 339–369. <https://www.jstor.org/stable/20185035>.
- 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, no. 6: 1481–1536. <https://doi.org/10.3982/ECTA6495>.
- Barndorff-Nielsen, O., and N. Shephard. 2002a. “Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models.” *Journal of the Royal Statistical Society Series B* 64, no. 2: 253–280. <https://www.jstor.org/stable/3088799>.
- Barndorff-Nielsen, O., and N. Shephard. 2002b. “Estimating Quadratic Variation Using Realized Variance.” *Journal of Applied Econometrics* 17, no. 5: 457–477. <https://www.jstor.org/stable/4129267>.
- Barndorff-Nielsen, O., and N. Shephard. 2006. “Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation.” *Journal of Financial Econometrics* 4, no. 1: 1–30. <https://doi.org/10.1093/jffinec/nbi022>.
- Bauwens, L., and P. Giot. 2000. “The Logarithmic ACD Model: An Application to the bid-Ask Quote Process of Three NYSE Stocks.” *Annales D'economie et Statistique* 60: 117–149.
- Bauwens, L., P. Giot, J. Grammig, and D. Veredas. 2004. “Comparison of Financial Duration Models via Density Forecasts.” *International Journal of Forecasting* 20: 589–609.
- Bollerslev, T., B. H. John Huss, and L. H. Pedersen. 2018. “Risk Everywhere: Modeling and Managing Volatility.” *Review of Financial Studies* 31: 2729–2773.
- Bollerslev, T., A. J. Patton, and R. Quaedvlieg. 2016. “Exploiting the Errors: A Simple Approach for Improved Volatility Forecasting.” *Journal of Econometrics* 192: 1–18.
- Brownlees, C., and G. Gallo. 2010. “Comparison of Volatility Measures: A Risk Management Perspective.” *Journal of Financial Econometrics* 8, no. 1: 29–56. <https://doi.org/10.1093/jffinec/nbp009>.
- Chen, F., F. Diebold, and F. Schorfheide. 2013. “A Markov Switching Multifactor Intertrade Duration Model, With Application to U.S. Equities.” *Journal of Econometrics* 177: 320–342. <https://doi.org/10.1016/j.jeconom.2013.04.016>.
- Christensen, K., R. C. Oomen, and M. Podolskij. 2014. “Fact or Friction: Jumps at Ultra High Frequency.” *Journal of Financial Economics* 114, no. 3: 576–599. <https://doi.org/10.1016/j.jfineco.2014.07.007>.
- Christoffersen, P. F. 1998. “Evaluating Interval Forecasts.” *International Economic Review* 39: 841–862.
- Clements, A., and D. P. Preve. 2021. “A Practical Guide to Harnessing the HAR Volatility Model.” *Journal of Banking & Finance* 133: 106285. <https://doi.org/10.1016/j.jbankfin.2021.106285>.
- Clements, M. P., A. B. Galvão, and J. H. Kim. 2008. “Quantile Forecasts of Daily Exchange Rate Returns From Forecasts of Realized Volatility.” *Journal of Empirical Finance* 15, no. 4: 729–750. <https://doi.org/10.1016/j.jempfin.2007.12.001>.
- Corsi, F. 2009. “A Simple Approximate Long-Memory Model of Realized Volatility.” *Journal of Financial Econometrics* 7, no. 2: 174–196. <https://doi.org/10.1093/jffinec/nbp001>.
- Diebold, F. X., and R. S. Mariano. 1995. “Comparing Predictive Accuracy.” *Journal of Business and Economic Statistics* 13: 253–263.
- Engle, R. 2000. “The Econometrics of Ultra High Frequency Data.” *Econometrica* 68: 1–22.
- Engle, R. F., and J. R. Russell. 1998. “Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data.” *Econometrica* 66: 1127–1162. <https://doi.org/10.2307/2999632>.
- Fuertes, A.-M., and J. Olmo. 2013. “Optimally Harnessing Inter-Day and Intra-Day Information for Daily Value-At-Risk Prediction.” *International Journal of Forecasting* 29, no. 1: 28–42. <https://doi.org/10.1016/j.ijforecast.2012.05.005>.

- Garman, M. B. 1976. "Market Microstructure." *Journal of Financial Economics* 3: 257–275.
- Ghysels, E., C. Gouriéroux, and J. Jasiak. 2004. "Stochastic Volatility Duration Models." *Journal of Econometrics* 119: 413–433.
- Giacomini, R., and I. Komunjer. 2005. "Evaluation and Combination of Conditional Quantile Forecasts." *Journal of Business and Economic Statistics* 23: 416–431.
- Giacomini, R., and H. White. 2006. "Tests of Conditional Predictive Ability." *Econometrica* 74: 1545–1578. <https://www.jstor.org/stable/4123083>.
- Giot, P., and S. Laurent. 2004. "Modelling Daily Value-at-Risk Using Realized Volatility and ARCH Type Models." *Journal of Empirical Finance* 11, no. 3: 379–398. <https://doi.org/10.1016/j.jempfin.2003.04.003>.
- Glosten, L. R., and P. R. Milgrom. 1985. "Bid, ask and Transaction Prices in a Specialist Market With Heterogeneously Informed Traders." *Journal of Financial Economics* 14: 71–100.
- Granger, C. W. 1999. "Outline of Forecast Theory Using Generalized Cost Functions." *Spanish Economic Review* 1: 161–173.
- Hansen, P. R., and A. Lunde. 2006. "Realized Variance and Market Microstructure Noise." *Journal of Business & Economic Statistics* 24, no. 2: 127–161. <https://www.jstor.org/stable/27638860>.
- Hansen, P. R., A. Lunde, and J. M. Nason. 2011. "The Model Confidence Set." *Econometrica* 79, no. 2: 453–497. <https://doi.org/10.3982/ECTA5771>.
- Hautsch, N. 2012. *Econometrics of Financial High-Frequency Data*. Springer Science Business Media.
- Hautsch, N., and M. Podolskij. 2013. "Preaveraging-Based Estimation of Quadratic Variation in the Presence of Noise and Jumps: Theory, Implementation, and Empirical Evidence." *Journal of Business & Economic Statistics* 31: 165–183.
- Hong, S. Y., I. Nolte, S. J. Taylor, and X. Zhao. 2023. "Volatility Estimation and Forecasts Based on Price Durations." *Journal of Financial Econometrics* 21: 106–144.
- Hull, J., and A. White. 1987. "The Pricing of Options on Assets With Stochastic Volatilities." *Journal of Finance* 42: 281–300.
- Jacod, J., Y. Li, P. A. Mykland, M. Pololskij, and M. Vetter. 2009. "Microstructure Noise in the Continuous Case: The Pre-Averaging Approach." *Stochastic Processes and Their Applications* 119, no. 7: 2249–2276. <https://doi.org/10.1016/j.spa.2008.11.004>.
- Jasiak, J. 1998. "Persistence in Intertrade Durations." *Journal of Finance* 19: 166–195.
- Liu, L., A. Patton, and K. Sheppard. 2015. "Does Anything Beat 5-Minute RV? A Comparison of Realized Measures Across Multiple Asset Classes." *Journal of Econometrics* 187, no. 1: 293–311. <https://doi.org/10.1016/j.jeconom.2015.02.008>.
- Liu, S., and Y.-K. Tse. 2015. "Intraday Value-at-Risk: An Asymmetric Autoregressive Conditional Duration Approach." *Journal of Econometrics* 189, no. 2: 437–446. <https://doi.org/10.1016/j.jeconom.2015.03.035>.
- Maheu, J., and T. McCurdy. 2011. "Do High-Frequency Measures of Volatility Improve Forecasts of Return Distributions?" *Journal of Econometrics* 160, no. 1: 69–76. <https://doi.org/10.1016/j.jeconom.2010.03.016>.
- Mukherjee, A., W. Peng, N. R. Swanson, and X. Yang. 2020. "Financial Econometrics and Big Data: A Survey of Volatility Estimators and Tests for the Presence Jumps and Co-Jumps." In *Financial, Macro and Micro Econometrics Using R*, edited by H. D. Vinod and C. Rao, vol. 42, 3–59. Elsevier.
- Pacurar, M. 2008. "Autoregressive Conditional Duration Models in Finance: A Survey of the Theoretical and Empirical Literature." *Journal of Economic Surveys* 22: 711–751. <https://doi.org/10.1111/j.1467-6419.2007.00547.x>.
- Patton, A. J. 2011a. "Data-Based Ranking of Realised Volatility Estimators." *Journal of Econometrics* 161, no. 2: 284–303. <https://doi.org/10.1016/j.jeconom.2010.12.010>.
- Patton, A. J. 2011b. "Volatility Forecast Comparison Using Imperfect Volatility Proxies." *Journal of Econometrics* 160, no. 1: 246–256. <https://doi.org/10.1016/j.jeconom.2010.03.034>.
- Patton, A., and K. Sheppard. 2009. "Optimal Combinations of Realised Volatility Estimators." *International Journal of Forecasting* 25, no. 2: 218–238. <https://doi.org/10.1016/j.ijforecast.2009.01.011>.
- Rodríguez-Poo, J. M., D. Veredas, and A. Espasa. 2008. "Semiparametric Estimation for Financial Durations." In *High Frequency Financial Econometrics: Recent Developments*, 225–251. Physica-Verlag HD.
- Segnon, M. and B. Schulte-Tillmann. 2025). "The Factorial Hidden Markov Duration Model With Application to Ultra-high Frequency Data." Working Paper.
- Tse, Y., and T. T. Yang. 2012. "Estimation of High-Frequency Volatility: An Autoregressive Conditional Duration Approach." *Journal of Business & Economic Statistics* 30, no. 4: 533–545. <https://doi.org/10.1080/0735015.2012.707582>.
- Tse, Y.-K., and Y. Dong. 2014. "Intraday Periodicity Adjustments of Transaction Duration and Their Effects on High-Frequency Volatility Estimation." *Journal of Empirical Finance* 28: 352–361. <https://doi.org/10.1016/j.jempfin.2014.04.004>.
- Zaharieva, M. D., M. Trede, and B. Wilfling. 2020. "Bayesian Semiparametric Multivariate Stochastic Volatility With Application." *Econometric Reviews* 39, no. 9: 947–970. <https://doi.org/10.1080/07474938.2020.1761152>.
- Zhang, L. 2006. "Efficient Estimation of Stochastic Volatility Using Noisy Observations: A Multi-Scale Approach." *Bernoulli* 12, no. 6: 1019–1043.
- Zhang, L., P. A. Mykland, and Y. Aït-Sahalia. 2005. "A Tale of Two Time Scales." *Journal of the American Statistical Association* 100, no. 472: 1394–1411.
- Žikeš, F., J. Baruník, and N. Shenai. 2017. "Modeling and Forecasting Persistent Financial Durations." *Econometric Reviews* 36: 1081–1110. <https://doi.org/10.1080/07474938.2014.977057>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. SchSegWied_FOR-24-0277_OnlineAppendix.pdf.

Appendix A

Data Cleaning

We retain only trades with positive trade prices and volumes and valid trade correction indicators "00" and "01". Following Aït-Sahalia et al. (2020), we further exclude trade sale conditions, "Z", "B", "U", "T", "L", "G", "W", "K", and "J", as well as an accompanying "I" for odd lot trades. We take the median price if multiple trades are recorded at the same timestamp which further reduces recording errors. This procedure however, does not yet guarantee fully reliable data. When merging the trade and quote observations within 1 μ s, we apply the following two filters: (i) the first is to remove all crossed quotes, that is, all quotes where the bid price exceeds the ask and (ii) the second is to exclude all transaction prices which are not part of the interval (Bid-Spread; Ask+Spread).