The Journal of Finance and Data Science 10 (2024) 100138

Contents lists available at ScienceDirect

The Journal of Finance and Data Science

journal homepage: www.keaipublishing.com/en/journals/jfds

Research article

Liquidity risk analysis via drawdown-based measures

Guglielmo D'Amico, Bice Di Basilio *, Filippo Petroni

Department of Economics, University of Chieti-Pescara, 65127 Pescara, Italy

ARTICLE INFO

Keywords:
Drawdown-based measures
Liquidity risk
High-frequency financial volumes
Semi-Markov model
Kullback–Leibler divergence

ABSTRACT

Trading volumes are key variables in determining the degree of an asset's liquidity. We examine the volume drawdown process and crash recovery measures in rolling-time windows to assess exposure to liquidity risk. The time-varying windows protect our financial indicators from the massive amount of volume transactions that characterize the opening and closing of the stock market. The empirical study is carried out for three Nasdaq-listed assets from April to September 2022. Firstly, we shape all of the volume time series using a weighted-indexed semi-Markov (WISMC) model, as well as the EGARCH and GJR models for comparisons. Next, we calculate drawdown-based risk measures on real and synthetic data, simulated from all the considered econometric models. Finally, we employ the Kullback-Leibler divergence to compare real and simulated risk indicators. Results reveal that the WISMC model reproduces all the drawdown-based risk measures better than the EGARCH and GJR models do for all the considered stocks.

1. Introduction

Financial liquidity refers to the ease with which an asset can be converted into cash without experiencing significant price movement. Every investor closely considers her/his exposure to liquidity risk when making trading decisions. Examining trading volumes, i.e., the number of asset transactions over a time period, is one way to assess this risk (see, e.g. Queirós, 2016; Johnson, 2008; Wattanatorn and Tansupswatdikul, 2019).

Over the years, several risk measures have been computed on financial volumes in order to quantify and manage liquidity risk. The most well-known are the quantile-based value-at-risk and expected shortfall (see e.g. Dionne et al., 2009; Fulga and Dedu, 2010). More recently, the drawdown-based risk indicators have attracted increased interest in academia (see e.g. Geboers et al., 2023). They take into account data evolution over time, which is a crucial aspect of financial time series, ignored by quantile-based risk indicators.

Among the most common drawdown-based risk measures, we mention maximum drawdown, average drawdown, and conditional drawdown-at-risk. Specifically, the maximum drawdown is the largest loss accrued with respect to the peak value recorded in a time interval (see e.g. Goldberg and Mouti, 2022). As a result, the average drawdown up to a specified time horizon is the time average of all drawdowns that have occurred up to time (see, e.g. Baghdadabad and Glabadanidis, 2013). Finally, conditional drawdown-at-risk is the average of all drawdowns that exceed a specific threshold, also known as drawdown-at-risk. It includes maximum and average drawdowns as its limiting cases, and is very similar to the definition of the expected shortfall (see, e.g. Krokhmal et al., 2005). Recently, a new family of drawdown-based risk measures has been advanced (see D'Amico et al., 2020; D'Amico et al., 2023; Zhang and Hadjiliadis, 2012; Hongzhong et al., 2018; Li et al., 2022; Masala and Petroni, 2023). These are crash and recovery measures, which are very useful

* Corresponding author.
E-mail addresses: g.damico@unich.it (G. D'Amico), bice.dibasilio@unich.it (B. Di Basilio), filippo.petroni@unich.it (F. Petroni).
Peer review under responsibility of KeAi Communications Co., Ltd.

https://doi.org/10.1016/j.jfds.2024.100138
Received 7 December 2023; Received in revised form 29 July 2024; Accepted 18 September 2024
Available online 24 September 2024

in portfolio decisions given their relation to market crashes and recoveries. In particular, the crash measures referred to as the drawdown of a fixed level, the time to crash, and the speed of the crash were first introduced in Zhang and Hadjiliadis (2012) and then extended in D'Amico et al., 2020; D'Amico et al., 2023. They provide crucial information to investors about the riskiness prior to a market crash event. The recovery measures known as the recovery time and the speed of the recovery analyze the steps that bring stocks from the financial collapse to the recovery; they were introduced and investigated in D'Amico et al., 2020.

In general, to reproduce financial time series, a stochastic model must be employed. Over time, several models have been proposed in the literature. Among these, we mind the classic econometric models such as the GARCH model family (see e.g. Gran et al., 2014; Mercuri, 2008; Arag et al., 2005) and the diffusion models (see e.g. Yuvan and Bier, 2018) used to reproduce both price and volumes series. Recently, effective alternatives based on semi-Markovian models have been offered to study both financial prices and volumes (see, e.g. D'Amico et al., 2012; D'Amico et al., 2021; D'Amico et al., 2017a; Barbu et al., 2021).

Semi-Markov processes (see e.g. Limnios and Oprisan, 2001; D'Amico et al., 2017b) are generalizations of Markov process (see e.g. Alexander et al., 2023; De Blasis, 2020) in which the time intervals between transitions are specified by any probability distribution. Thus, Markovian processes can be viewed as a subset of semi-Markovian processes that are obtained when the waiting time distributions in the system's states are memoryless distributions. Specifically, an ordinary semi-Markovian model considers the current state of the system and how long it has been in that state but not the history before. Since the past is very important in financial time series, this is a critical shortcoming for this model. To overcome this issue, an extension of the standard semi-Markov chain model named the weighted-indexed semi-Markov chain model (WISMC) was introduced (see, e.g. D'Amico et al., 2020; D'Amico et al., 2018; De Blasis, 2023; D'Amico et al., 2021; D'Amico et al., 2011). Its innovativeness and what distinguishes it from a standard semi-Markov model is the addition of an index process which allows the optimal accumulation of the past information contained in the financial time series.

In this work, we suggest to apply the previously mentioned crash and recovery drawdown-based risk indicators to high-frequency financial volumes, i.e. data recorded on a very small time scale such as the minute. All these risk indicators are mainly enforced to the price time series in the financial literature (see e.g. D'Amico et al., 2020; D'Amico et al., 2023). It is worth noting that their calculation on financial prices is used to measure market risk, while their use on financial volumes is useful for the assessment of an asset's liquidity risk sensitivity.

We shape the intraday volumes of the Nasdaq Stock Exchange-listed Tesla, Netflix, and Apple assets by different econometric models. Specifically, we reproduce high-frequency financial volumes using both a WISMC process and also the EGARCH and GJR models for comparisons. We chose EGARCH and GJR models since they are very popular in financial econometrics. By employing real and synthetic data, we compute drawdown-based risk indicators in rolling time windows by entering a variable starting time s. Next, for both real and simulated risk measures, we evaluate their best parametric distribution among the Lognormal, Weibull, Exponential, and Gamma laws selected through the AIC and BIC criteria (see D'Amico et al., 2020). Since our drawdown-based risk indicators are affected by Type-1 right censorship, all the estimation procedures are made considering this statistical data feature (see D'Amico et al., 2020). At the end, we compare real and synthetic financial indicators by means of the Kullback-Leibler divergence (see Kullback and Leibler, 1951). Globally, for all assets, the simulated financial risk indicators from the WISMC model are very close to the real ones. The synthetic risk measures from the EGARCH and GJR models, on the other hand, perform worse than those from the WISMC models. The encouraging findings demonstrate that the WISMC model can be a potent new econometric method for modeling financial volumes in addition to financial returns (see e.g. D'Amico et al., 2020; D'Amico et al., 2012; D'Amico et al., 2018; De Blasis, 2023). These characteristics point to the necessity for further research and application of the WISMC model in financial economics and its practical applications.

The rest of the paper is organized as follows. Section 2 provide a full discussion of the examined drawdown-based risk measures. Section 3 offers a brief overview of all the econometric models used in our investigation. Section 4 presents the findings of our analysis. Section 5 reports concluding remarks and future aims.

2. Drawdown-based risk indicators for financial volumes

When a financial asset or security may be quickly and easily converted into cash without significantly affecting its price, the stock owners have a liquid asset. Large volumes usually indicate a liquid asset, while low volumes denote an illiquid security that is difficult to monetize.

Generally, financial volumes are greatest at market opening and closing since traders take a position (see Graczyk and Duarte Queiros, 2016). To account for this behavior, we adopt risk indicators including a general starting time s that allows their computation in rolling time windows.

The risk of liquidity can be measured by observing drawdown events and requires the definition of specific indicators. To qualify the drawdown in the volume of an asset, we introduce the discrete time-varying volume process $X = \{X(t),\ t \in \mathbb{N}\}$ and its running maximum process $Y^s(t)$, defined $\forall s, t \in \mathbb{N}$ with $s < t$ as:

$$Y^s(t) := \max_{l \in [s,t] \cap \mathbb{N}} \{X(l)\}. \tag{1}$$

The addition of $s > 0$ reduces the impact of large initial volume transactions on our risk measures. Obviously, if time s equals zero, we include the effect of market opening in the measure.

Consequently, the drawdown process $D^s(t)$ is obtained by subtracting the volume process $X(t)$ from the running maximum process $Y^s(t)$:

$$D^s(t) := Y^s(t) - X(t) \quad \text{with} \quad t \geq s. \tag{2}$$

Equation (2) describes the correction of the security trading volume in relation to a prior relative maximum. Hence, large values of $D^s(t)$ denote that the market is far from the largest liquidity observed since time s.

To fully understand the previous definitions and the role played by time s, in Fig. 1 we propose a graphic illustration considering $s = 0$ and $s = 100$. The red, blue, and black lines represent the volume processes, the current maximum processes, and the drawdown processes for Apple stock within a trading day, respectively. The market's stock day is composed of 391 min when $s = 0$, and 291 min when $s = 100$. As it can be seen, Apple's volumes show a typical U-shape form (see e.g. Queirós, 2016) and higher initial fluctuations for $s = 0$ than for $s = 100$. These exceptionally large volume variations have significant implications for the computation of the drawdown-based risk measures since the running maximum process, and thus also the drawdown process, assume very high values from the first minutes of the trading day.

To investigate the impact of these abrupt volume fluctuations on our risk indicators, we compute them considering different starting times s according to the time from which the risk measures must be computed.

We propose crash and recovery measures related to market liquidity collapses. Let us assume we set a threshold $M > 0$ to denote the warning distance from the maximum observed volume transaction. Any displacement greater than M is considered as an alert of low volumes indicating an insufficient market interest towards the asset and hence low liquidity. The first time when this event happens is defined as the drawdown of a fixed level $M > 0$, i.e.:

$$\tau^s(M) := \min\{t \geq s | D^s(t) \geq M\}. \tag{3}$$

Now, if we consider the quantity $\rho^s(M)$ that is the last visit time of the maximum before the stopping time $\tau^s(M)$:

$$\rho^s(M) := \max\{t \in [s, \tau^s(M)] \cap \mathbb{N} | X(t) = Y^s(t)\}, \tag{4}$$

we identify the time to crash as:

$$T_c^s(M) := \tau^s(M) - \rho^s(M). \tag{5}$$

It is the time necessary to have the first M-change in the drawdown process.
Accordingly, the speed of crash, that is the velocity at which the first M-variation occurs, is defined as:

$$S_c^s(M) := \frac{M}{T_c^s(M)}. \tag{6}$$

The measures described above provide knowledge about an asset's riskiness before it reaches a M-intensity crash event. Likewise, it is interesting to study the successive stages that bring the financial asset from the crash to the recovery. To this end we introduce a second threshold $0 < M' < M$ denoting the level under which the drawdown process indicates the recovery of an acceptable level of stock liquidity. Precisely, we define the quantity $\gamma^s(M, M')$ as

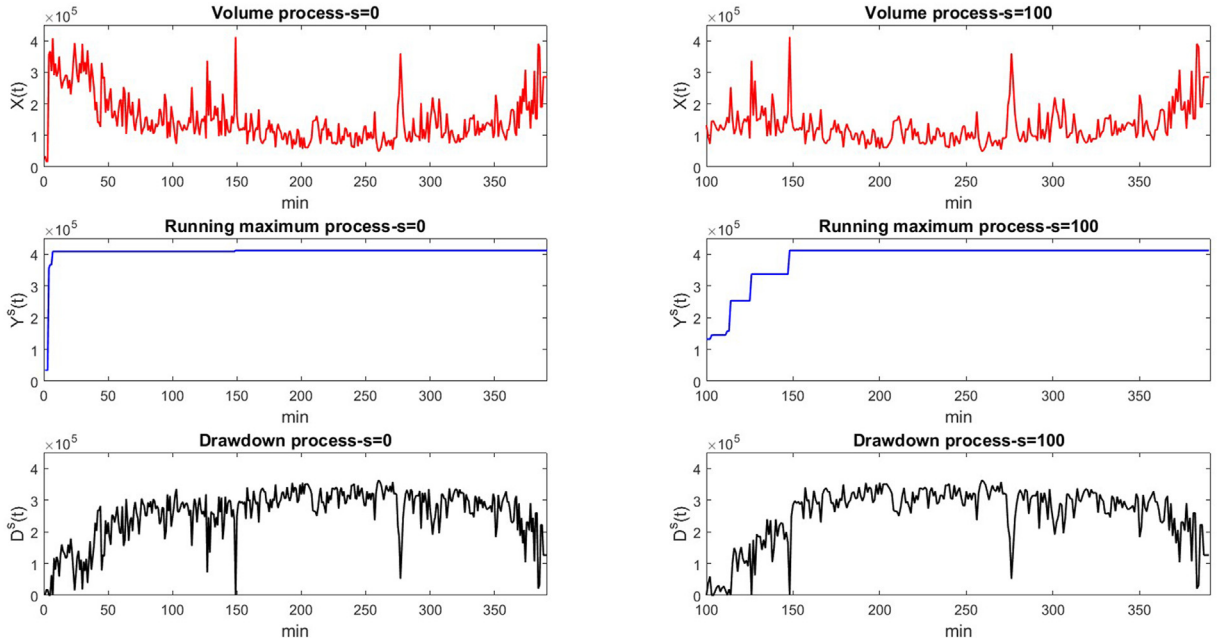

Fig. 1. Volume processes (red lines), running maximum processes (blue lines) and drawdown processes (black lines) for Apple asset, considering two different starting times ($s = 0$ and $s = 100$).

$$\gamma^s(M, M') := \min\{t > \tau^s(M) | D^s(t) \leq M'\}. \tag{7}$$

It denotes the first time in which the drawdown process declines below the threshold M', just after having crossed the threshold M for the first time.

The difference between $\gamma^s(M, M')$ and $\tau^s(M)$ is called the recovery time and measures the time necessary to resort to an acceptable liquidity level after having experienced a drawdown event:

$$R_t^s(M, M') := \gamma^s(M, M') - \tau^s(M). \tag{8}$$

Thence, the speed of recovery can be defined as the velocity at which the drawdown process attains the M' threshold after having exceeded the threshold M for the first time:

$$S_r^s(M, M') := \frac{M - M'}{R_t^s(M, M')}. \tag{9}$$

Fig. 2 illustrates graphically all the financial indicators just described, considering the Apple asset and fixing $s = 100$. The figure shows the time evolution of the drawdown process during a trading day. The left-side Y-axis measures the level of the drawdown process while the right-side Y-axis shows the values of the two thresholds M and M'.

The risk metrics mentioned above are all useful financial instruments for assessing the liquidity risk of a single security or a portfolio. They depend on the M and M' thresholds selected by investors based on their own risk tolerance.

Specifically, the crash measures τ^s, T_c^s, and S_c^s are M-dependent. Small M-values describe a situation where there is a particular intolerance about liquidity risk, and the alert is given as soon as this little displacement from the maximum liquidity is observed. Conversely, high M-values depict the opposite situation where there is low intolerance about liquidity risk, and the alert is given only when a strong displacement from the maximum liquidity is recorded. The risk indicator τ^s informs about the hazard of an asset in terms of liquidity risk, i.e. its proclivity to become illiquid. The risk measures T_c^s and S_c^s quantify the duration and speed at which a stock becomes illiquid.

Differently, the recovery measures R_t^s and S_r^s investigate the behavior of a stock following the achievement of a certain M-level in its drawdown, and thus, they show a dependence from both the M and the M'-thresholds. Precisely, the risk indicators R_t^s and S_r^s quantify how long it takes for an asset to move from an illiquid state to a more liquid one and how quickly this transition happens, respectively.

3. Mathematical models

This section provides a short overview of the econometric models employed to shape high-frequency financial volumes. Specifically, Subsection 3.1 describes the weighted-indexed semi-Markov chain (WISMC) model. The EGARCH and GJR models, used to make comparisons with the WISMC model, are discussed in Subsection 3.2.

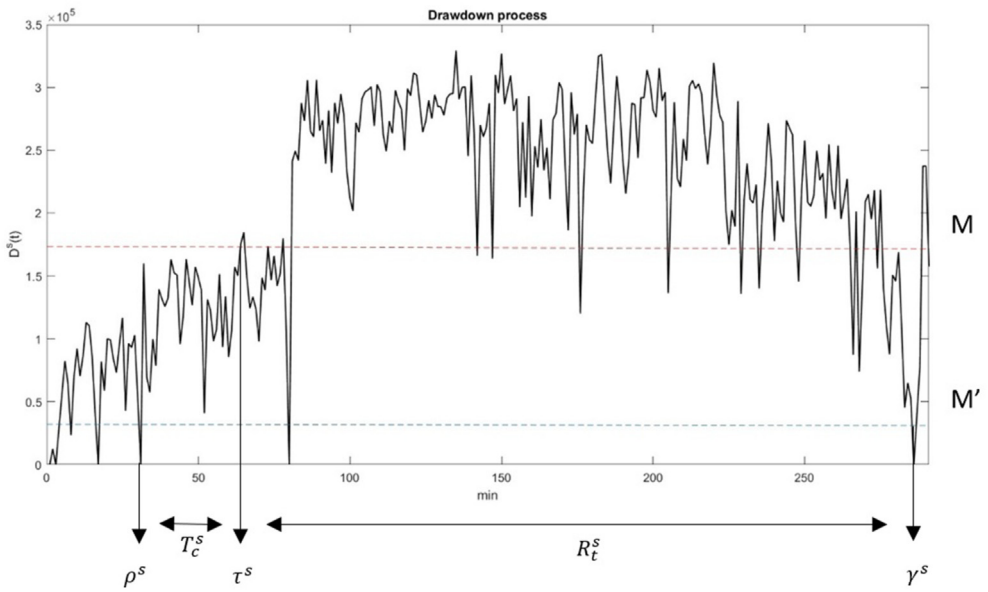

Fig. 2. Drawdown processes of Apple asset during a trading day, for $s = 100$. The orange and blue lines stand for the thresholds M and M'.

3.1. Weighted-indexed Semi-Markov chain model

This subsection provides an informal presentation of WISMC models. In order to demonstrate the idea generating this class of models, we first observe that very popular models like lattice one (binomial, trinomial, and similar, see e.g. Bhat and Kumar, 2010) are models where the probabilistic evolution of the system is independent of the past trajectory given the value of the process at the present time. This characteristic is known as the Markov property of these processes. The Markov property is useful for computational purposes but it poses strong restrictions that frequently are not respected in real financial data (see e.g. D'Amico et al., 2017b). A semi-Markov model would certainly be preferable since it considers not only the current state of the process, but also the time elapsed in this state since the last transition. Evidently, the semi-Markov framework is a first attempt to introduce memory effects inside the model dynamic. This direction of research has been further extended with the definition of WISMC models.

Basically, the WISMC model is described by three stochastic processes: $\{J_n\}_{n\in\mathbb{N}}$, $\{T_n\}_{n\in\mathbb{N}}$ and $\{U_n^\lambda\}_{n\in\mathbb{N}}$. In our framework, the random variables J_n and T_n denote the value of the volumes process at the n-th change (J_n) and its corresponding jump times (T_n), meaning the time in which the n-th change in the volumes process occurred, see e.g., D'Amico et al., 2012; D'Amico et al., 2018; D'Amico et al., 2021. Finally, the random variable U_n^λ describes the value of the index process in the n-th transition. It summarizes the information contained in the past trajectory of the volume process up to the nth transition and it is defined as:

$$U_n^\lambda = \sum_{k=0}^{n-1} \sum_{a=T_{n-1-k}}^{T_{n-k}-1} f^\lambda(J_{n-1-k}, T_n, a) + f^\lambda(J_n, T_n, T_n). \tag{10}$$

Hence, the index process makes use of a generic function f that provides a score to past information in the trajectory and of a memory parameter λ that should be calibrated on the data. Evidently, different choices of the score function f generate a variety of models within the WISMC class.

Previous applications on high-frequency financial data (see e.g. D'Amico et al., 2020; D'Amico et al., 2018; D'Amico et al., 2021) use an exponentially weighted moving average of the squared J as a score function f:

$$f^\lambda(J_{n-1-k}, T_n, a) = \frac{\lambda^{T_n-a} J_{n-1-k}^2}{\sum_{a=1}^{T_n} \lambda^a}. \tag{11}$$

The function f depends on both the past values of volumes J_{n-1-k} that occurred at times a, the current time $T_n > a$ and the parameter λ, which balances the past information.

According to previous literature focusing on financial returns (D'Amico et al., 2020; D'Amico et al., 2018; D'Amico et al., 2021), we now apply function (11) also to financial volumes, as squared volumes have a persistent autocorrelation over time, as confirmed in Fig. 4. This persistency can be responsible for trends and volatility clusters in the volume process, which can be identified by the index process using a suitable score function f and weight λ.

To build the WISMC model, we explicitly defined the dependency structure between the random variables J_n, T_n, and U_n^λ. To this aim, we assume that:

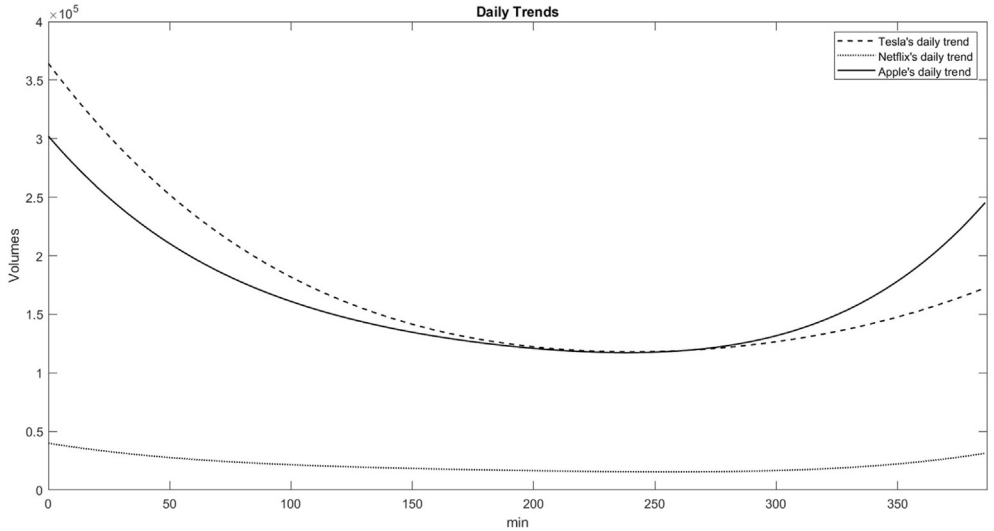

Fig. 3. Daily trend for Tesla, Netflix, and Apple assets.

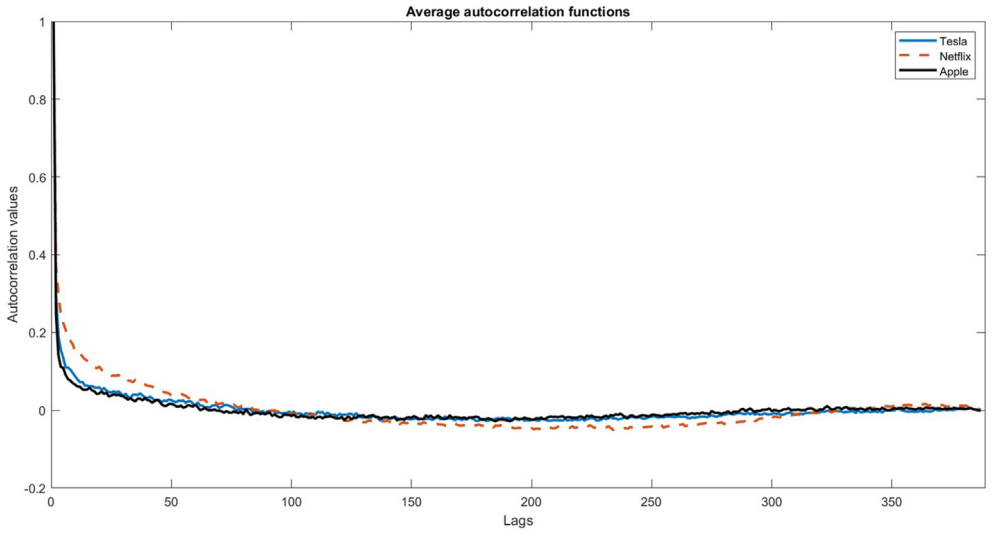

Fig. 4. Average autocorrelation functions for Tesla, Netflix and Apple.

$$\begin{aligned}\mathbb{P}[J_{n+1} &= j,\ T_{n+1} - T_n \le t \mid \sigma(J_h, T_h, U_h{}^{\lambda})_{h=0}^{n}, J_n = i, U_n{}^{\lambda} = u] \\ &= \mathbb{P}[J_{n+1} = j,\ T_{n+1} - T_n \le t \big| J_n = i, U_n^{\lambda} = u] =: Q_{ij}^{\lambda}(u,t),\end{aligned} \tag{12}$$

where $\sigma(J_h, T_h, U_h^{\lambda})$ is the natural filtration of the three-variate process $\{J_n, T_n, U_n^{\lambda}\}$.

In this mathematical model, the matrix of functions $\mathbf{Q}^{\lambda}(u,t) = (Q_{ij}^{\lambda}(u,t))_{i,j\in E}$, named the weighted-indexed semi-Markov kernel, is a crucial component. If $\mathbf{Q}^{\lambda}(u, t)$ is constant in u then, the weighted-indexed semi-Markov kernel degenerates into an ordinary semi-Markov kernel. Relation (12) affirms that the joint probability distribution of the volume variable J_{n+1} and the sojourn time length $T_{n+1} - T_n$ for having its next change from J_n to J_{n+1} depends on the current level of the volume J_n and on the index process value U_n^{λ}. The addition of the index process conveys important information as it allows for the identification of volatility clusters and the assessment of different transition probabilities according to its values. Obviously, if the transition probabilities $(Q_{ij}^{\lambda}(u,t))$ are not dependent on u, the WISMC model degrades into an ordinary semi-Markov chain.

For an exhaustive explanation of the estimation procedures for each quantity included in the model, see D'Amico et al., 2018.

3.2. EGARCH and GJR models

The most common extensions of the Generalized Auto-Regressive Conditional Heteroskedastic (GARCH) model (Bollerslev, 1986) are the Exponential GARCH (EGARCH) (Nelson, 1991) and the Glosten-Jagannathan-Runkle GARCH (GJR) models (Glosten et al., 1993).

These GARCH variants were created specifically to account for the asymmetry effect, while a standard GARCH model cannot account for it.

In particular, the EGARCH model employs a conditional variance equation in logarithmic form with an additional leverage term that enables the model to react differently to negative and positive shocks. Similarly, the GJR model uses an indicator function to respond more strongly to negative shocks.

Let X_t and μ_t be the volume process and the expected volume value, respectively. Let $\epsilon_t = X_t - \mu_t$ be the innovation process at time t. Then ϵ_t follows a general EGARCH (p, q) model if:

$$\epsilon_t = \sigma_t z_t, \tag{13}$$

$$log(\sigma_t^2) = k + \sum_{i=1}^{p} \gamma_i log(\sigma_{t-i}^2) + \sum_{j=1}^{q} \alpha_j \left[\frac{|\epsilon_{t-j}|}{\sigma_{t-j}} - \mathbb{E}\left\{ \frac{|\epsilon_{t-j}|}{\sigma_{t-j}} \right\} \right] + \sum_{j=1}^{q} \xi_j \left\{ \frac{\epsilon_{t-j}}{\sigma_{t-j}} \right\}. \tag{14}$$

The EGARCH (p, q) model has p GARCH components related to lagged log-variance terms, q ARCH components connected with lagged standardized innovation, and q leverage components associated with signed, lagged standardized innovation. The sequence of i.i.d. random variables $\{z_t\}$ is usually assumed to be Gaussian or Student-t distributed.

As for the GJR model, we say that ϵ_t follows a general GJR (p, q) model if:

$$\epsilon_t = \sigma_t z_t, \tag{15}$$

$$\sigma_t^2 = k + \sum_{i=1}^{p} \gamma_i \sigma_{t-i}^2 + \sum_{j=1}^{q} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{q} \xi_j \mathbb{1}_{\epsilon_{t-j}} \epsilon_{t-j}^2. \tag{16}$$

The GJR (p, q) has p GARCH components related to lagged variances, q ARCH components associated with lagged squared innovations, and q leverage components referred to the squared of negative lagged innovation. The indicator function $\mathbb{1}_{\epsilon_{t-j}}$ equals 1 if $\epsilon_{t-j} < 0$ and 0 otherwise. As well as for the EGARCH model, it assumes that the i.i.d. random variables $\{z_t\}$ are Gaussian or Student-t distributed. Finally, for stationarity and positively, it is supposed $k > 0$, $\gamma_i > 0$, $\alpha_i \geq 0$, $\alpha_i + \xi_j \geq 0$ and $\sum_{i=1}^{p} \gamma_i + \sum_{j=1}^{q} \alpha_j + \frac{1}{2} \sum_{j=1}^{q} \xi_j < 1$.

4. Results

This section is divided into two parts. Subsection 4.1 provides both an overview of the main statistical features of the financial data set used in our analyzes and information on how all the econometric models are implemented. Subsection 4.2 examines the risk measures on both real and simulated data from all the models considered and com pares the results.

4.1. Data analysis and preliminary model findings

Empirical research is carried out on intraday volumes of three stocks listed on the Nasdaq Stock Exchange. Data were downloaded from Thomson Reuters EikonTM (https://eikon.thomsonreuters.com/index.html) from April 2022 to September 2022. We chose Tesla, Netflix, and Apple assets, which belong to different industrial sectors. In fact, Tesla deals mainly with the Automotive and Renewable sector, while Netflix and Apple belong to the Communication Services, Technology & Entertainment, and Information Technology sectors, respectively. Globally, we examine 126 days of the stock market, each composed of 391 min.

Descriptive statistics of Tesla, Netflix, and Apple assets are reported in Table 1. It can be observed that all the considered stocks exhibit positive asymmetry and excess of kurtosis, hence they do not fit into the Gaussian setting.

Data above the 95-th percentile are considered as outliers, and are replaced with the semi-sum of the closest data. Furthermore, all our high-frequency trading volumes follow a daily trend whereby they climb at market open, fall during the trading day, and then rise again at market close. By averaging over each trading day, minute by minute, daily trend values are detected.

To accurately model daily volumes, we detrended all financial time series by shaping trends with a fourth-degree polynomial. Results of the polynomial fit with the corresponding standard errors are reported in Table 2. The estimated fourth-degree polynomials are plotted in Fig. 3 for each asset and denote the daily trend of volumes.

It is important to emphasize that fourth-degree polynomials are not able to overfit data because they only reproduce the daily trend, while the volume process exhibits a lot of noise around this trend. We can verify this statement by looking at the volume processes (red lines) in Fig. 1, where there is a strong irregularity in the volume processes that a fourth-degree polynomial would not be able to fully capture.

To shape volumes by a WISMC process, we discretize the values of both the volume process and the index process and we define the value of the parameter λ for each asset. The calibration of the number of states for the volume process and the quantity λ is realized by minimizing the mean percentage error (MPE) between real and synthetic autocorrelation functions, following the algorithm presented in D'Amico et al. (2018).

Fig. 4 reports the minute-by-minute average autocorrelation of the squared volumes, for each considered stocks. In particular, we first computed the autocorrelations day by day and then we estimated the average autocorrelations per minute considering all 126 trading days. It can be noted that all autocorrelation functions decrease slowly and this is especially visible for Netflix stock which has a larger autocorrelation function than the other two stocks.

Basically, the algorithm expects to build a trajectory by arbitrarily fixing a set of states and a value for λ, estimate the weighted-indexed semi-Markov kernel, and then perform a Monte Carlo simulation to generate a synthetic time series. Then, the autocorrelation functions for both real and simulated data are calculated and the MPE values are compared. All steps are repeated while changing the values of the states and of λ and we set the number of states and the value of λ that best describe the data by minimizing the MPE. Using this optimization procedure, we discover that, for all the considered stocks, the optimal number of states for the volume process is 5 and the optimal value of λ is around 0.93.

For comparison purposes, we also shape daily trading volumes using the EGARCH and GJR models. In particular, we focus on EGARCH(1,1), EGARCH(1,2), EGARCH(2,1), GJR(1,1), GJR(1,2), and GJR(2,1). We employ Gaussian innovations and the average values of the parameters derived by fitting the model and data day by day to simulate a financial volume series from each considered EGARCH and GJR model. The maximum likelihood estimation of the parameters is performed using Matlab software. Table 3 shows, for

Table 1
First quartile (Q1), second quartile (Q2), third quartile (Q3), mean (M), standard deviation (SD), skewness (SK), and kurtosis (KURT) of Tesla, Netflix, and Apple volumes.

Stock	Q1	Q2	Q3	M	SD	SK	KURT
TESLA	91,226	1.426e+05	2.335e+05	1.683e+05	1.182e+05	1.271	4.562
NETFLIX	9071	15,993	29,341	2.137e+04	1.972e+04	1.974	7.550
APPLE	98,346	144,172	221,609	1.631e+05	1.012e+05	1.105	4.321

Table 2
Estimated coefficient values of the fourth-degree polynomials and relative standard errors (S.E.).

TESLA		
	Estimated values	S.E.
$p1$	$0.319 \cdot 10^4$	$0.131 \cdot 10^4$
$p2$	$-0.883 \cdot 10^4$	$0.114 \cdot 10^4$
$p3$	$3.906 \cdot 10^4$	$0.351 \cdot 10^4$
$p4$	$-2.926 \cdot 10^4$	$0.224 \cdot 10^4$
$p5$	$12.390 \cdot 10^4$	$0.170 \cdot 10^4$
NETFLIX		
	Estimated values	S.E.
$p1$	$0.147 \cdot 10^4$	$0.021 \cdot 10^4$
$p2$	$0.039 \cdot 10^4$	$0.018 \cdot 10^4$
$p3$	$0.198 \cdot 10^4$	$0.057 \cdot 10^4$
$p4$	$-0.364 \cdot 10^4$	$0.036 \cdot 10^4$
$p5$	$1.670 \cdot 10^4$	$0.027 \cdot 10^4$
APPLE		
	Estimated values	S.E.
$p1$	$0.900 \cdot 10^4$	$0.128 \cdot 10^4$
$p2$	$0.218 \cdot 10^4$	$0.112 \cdot 10^4$
$p3$	$2.410 \cdot 10^4$	$0.334 \cdot 10^4$
$p4$	$-2.295 \cdot 10^4$	$0.219 \cdot 10^4$
$p5$	$12.220 \cdot 10^4$	$0.160 \cdot 10^4$

all stocks, the parameter values used to replicate daily volumes from the EGARCH and GJR models. Each parameter is associated with its standard error (in the bracket).

4.2. Results on drawdown-based risk indicators

Once the models are inferred, we can use them for comparison purposes with respect to the real data. We adapt the methodology advanced in D'Amico et al. (2020), which was applied to financial returns, to financial volumes.

First, we reproduce intraday high-frequency financial volumes on a 1-min time scale from all the considered econometric models. Then we compute the drawdown-based risk indicators on real and simulated data. Then, we fit real and synthetic risk measures using the Lognormal, Weibull, Exponential, and Gamma laws. To associate each real and simulated risk indicator with the corresponding best parametric law, we use the AIC and BIC criteria. Thus, we choose the parametric law with the smallest AIC and BIC value for every risk measures. In particular, the best parametric model selection for T_c^s is realized by fixing s and varying M. Likewise, the best model selection for R_t^s is done by focusing on s and M while M' changes. Since our risk measures are affected by Type-1 right censorship (see, e.g. Kaplan and Meier, 1958), following D'Amico et al. (2020), all estimation procedures are carried out with this statistical data condition in mind.

As for the selection of the best parametric model for real and simulated risk indicators using the WISMC model we directly choose the parametric law with the smallest AIC and BIC values. Regarding the EGARCH and GJR models, it is necessary to take a further step to identify only one model for each of these families of econometric models. To this end, after detecting the best parametric law for each considered EGARCH and GJR models, we choose the model with the smallest AIC and BIC values among the best. Practically, we detect the best parametric law for EGARCH (1,1), EGARCH (1,2), and EGARCH (2,1) according to the AIC and BIC criteria. Then, among the best-detected laws for each EGARCH, we select the EGARCH with the smallest AIC and BIC values. These steps are repeated also for GJR family, for both T_c^s and R_t^s measures.

Results of the parametric selection as well as the estimated parameter values for each best parametric law are reported as supplementary materials. Overall, we remark that the WISMC model always detects the Lognormal law for both T_c^s and R_t^s risk measures, and this is the same behavior observed in real data. On the contrary, the EGARCH and GJR models change their best parametric law choice time by time revealing a kind of model instability when used to identify drawdown based risk measures of the volume variable.

The best parametric distributions of S_c^s and S_r^s are directly derived from the best T_c^s and R_c^s's parametric distributions since S_c^s and S_r^s are a non-linear transformations of them.

Once we identified the best parametric law for each model and for the real data, we need to measure how far are the models' laws with respect to the real data distribution. This is done using the Kullback-Leibler divergence (Kullback and Leibler, 1951) which is a non-symmetric measure that quantifies the divergence between two given probability distributions P and Q. In particular, the larger the Kullback-Leibler divergence, the more dissimilar the two distributions will be. Obviously, if it takes the value zero, the two distributions will be identical.

Table 3
Average and standard deviation (in brackets) values of the parameters of the EGARCH and GJR for Tesla, Netflix and Apple.

EGARCH(1,1)	TESLA	NETFLIX	APPLE
Constant	5.157 (5.144)	2.796 (3.953)	5.236 (5.494)
Arch1	0.278 (0.155)	0.278 (0.221)	0.272 (0.155)
Garch1	0.770 (0.230)	0.848 (0.221)	0.763 (0.249)
Leverage1	0.056 (0.10)	0.116 (0.239)	0.044 (0.105)
EGARCH(1,2)	TESLA	NETFLIX	APPLE
Constant	3.711 (6.594)	2.848 (5.818)	4.195 (6.344)
Arch1	0.318 (0.224)	0.241 (0.491)	0.328 (0.573)
Arch2	−0.116 (0.270)	−0.023 (0.338)	−0.130 (0.214)
Garch1	0.834 (0.298)	0.842 (0.329)	0.809 (0.293)
Leverage1	0.055 (0.212)	0.143 (0.340)	0.050 (0.224)
Leverage2	−0.027 (0.212)	−0.020 (0.338)	−0.027 (0.197)
EGARCH(2,1)	TESLA	NETFLIX	APPLE
Constant	4.249 (5.338)	2.880 (4.099)	5.647 (7.843)
Arch1	0.307 (0.158)	0.305 (0.228)	0.298 (0.167)
Garch1	0.515 (0.346)	0.616 (0.394)	0.446 (0.456)
Garch2	0.296 (0.387)	0.225 (0.370)	0.294 (0.400)
Leverage1	0.040 (0.100)	0.116 (0.249)	0.035 (0.197)
GJR(1,1)	TESLA	NETFLIX	APPLE
Constant	1.675e+09 (1.280e+09)	3.321e+07 (3.468e+07)	1.143e+09 (9.835e+08)
Arch1	0.221 (0.170)	0.279 (0.528)	0.205 (0.184)
Garch1	0.527 (0.274)	0.589 (0.308)	0.554 (0.298)
Leverage1	−0.046 (0.533)	−0.132 (0.241)	−0.044 (0.202)
GJR(1,2)	TESLA	NETFLIX	APPLE
Constant	6.839e+08 (9.130e+08)	2.161e+07 (2.743e+07)	4.400e+08 (5.851e+08)
Arch1	0.134 (0.138)	0.195 (0.205)	0.113 (0.109)
Arch2	0.0164 (0.045)	0.06 (0.114)	0.020 (0.055)
Garch1	0.742 (0.221)	0.647 (0.274)	0.754 (0.224)
Leverage1	−0.045 (0.161)	−0.095 (0.230)	−0.012 (0.114)
Leverage2	0.054 (0.118)	0.024 (0.202)	0.029 (0.114)
GJR(2,1)	TESLA	NETFLIX	APPLE
Constant	3.582e+08 (3.535e+08)	7.687e+06 (9.021e+06)	2.387e+08 (2.389e+08)
Arch1	0.141 (0.100)	0.183 (0.148)	0.129 (0.105)
Garch1	0.341 (0.281)	0.395 (0.313)	0.327 (0.293)
Garch2	0.450 (0.279)	0.374 (0.297)	0.476 (0.288)
Leverage1	0.03 (0.100)	−0.035 (0.170)	0.024 (0.300)

Formally, the Kullback-Leibler divergence $D(P\|Q)$ in its continuous version is defined as follows:

$$D(P\|Q) = \int_{-\infty}^{+\infty} p(x) log_2 \left(\frac{p(x)}{q(x)}\right) dx, \tag{17}$$

where p and q are the probability densities of P and Q. It expresses the measure of the information lost when the distribution Q is employed to approximate the distribution P.

Small values of $D(P\|Q)$ denote that the considered model closely approximates the real data. Large values, on the other hand, suggest that the considered model is unsuitable for describing real data.

In detail, we compute the Kullback-Leibler divergence for the distribution of the risk measure T_c evaluated on the simulated data compared to the real ones in order to estimate the goodness of the proposed models. We repeat the same computations also for the risk indicator R_t. Because the Kullback-Leibler divergence is invariant to parameter transformations, its computation is not necessary for S_c^s and S_r^s.

Just to give an example, in Table 4 and 5 we display the Kullback-Leibler values for Apple asset with $s = 100$, considering both T_c^s and R_t^s. We can note that, for all the fixed thresholds, the WISMC model has the smallest Kullback-Leibler divergence for both T_c^s and R_t^s. Full results for all assets and risk measures are reported as supplementary materials. Globally, the WISMC model reproduces the drawdown-based risk measures better than the other econometric models.

The values of the KL-divergence are of interest not only to assess the divergence of a model (distribution) with respect to the true one but have important additional implications in statistics that can be worth discussing in relation to our financial indicators right now. For example, let us focus on the probability distribution of the time to crash T_c^s. The application of our methodology ends up with a pdf estimated on real data p(x) and three pdf estimated according to the different stochastic models considered $q^W(x)$, $q^E(x)$ and $q^G(x)$, where the apex distinguishes between the generating models according to their initial letters. According to Martino et al. (2017) and Li et al. (2020), let us suppose we are interested in the estimation of a generic quantity

Table 4
Kullback-Leibler divergence computed on the risk measure T_c^s as a function of M, fixing $s = 100$ for Apple asset. The smallest distances are in bold.

Kullback-Leibler divergence for T_c^s-APPLE	
$s = 100 - M = 30$ %	
WISMC	**0.0339**
EGARCH(1,1)	0.9307
GJR(2,1)	0.2777
$s = 100 - M = 40$ %	
WISMC	**0.0223**
EGARCH(1,1)	0.5999
GJR(2,1)	0.2580
$s = 100 - M = 80$ %	
WISMC	**0.0220**
EGARCH(1,2)	0.1070
GJR(1,1)	0.0836

$$I = \int_0^\infty h(x)p(x)dx \tag{18}$$

with $h(x)$ a square-integrable function with respect to p(x). Clearly, we can do it using Monte Carlo simulations drawing x_n, $n = 1, \ldots; N$ independent realizations from p(x) and using the sample mean estimator

$$\widehat{I} = \frac{1}{N}\sum_{n=1}^{N} h(x_n). \tag{19}$$

In general, we don't know the real distribution, then we apply estimator (18) on data generated by an alternative pdf $q^i(x)$, $i \in \{W, E, G\}$. The resulting estimators, namely $\widehat{I}^W, \widehat{I}^E, \widehat{I}^G$ will be less efficient than I as they are generated by proxy distributions of the unknown real one. To quantify the loss of efficiency, it is very common the use of an indicator called Effective Sample Size (ESS) which informally is the number of samples drawn from the true distribution required in order to have the same efficiency in the estimation of I as those obtained with N samples drawn from the approximating distribution $q^i(x)$. There are different formal definition of the ESS, one of them is of particular interest for us because involves the KL-divergence. The relation is the following:

$$ESS = N \cdot 2^{-D(P\|Q)}. \tag{20}$$

Based on relation (20) we can interpret the results of our analysis in terms of efficiency of the estimation of an hypothetical functional I of the time to crash T_c^s. Using the values of the KL divergence reported in Table 4, for example consider the case $s = 100 - M = 30$ %, we can compute the corresponding values of ESS for an arbitrary choice of $N = 100$. The values are the following:

$$\begin{aligned} ESS^W &= 100 \cdot 2^{-0.0339} \approx 98, \\ ESS^E &= 100 \cdot 2^{-0.9307} \approx 52, \\ ESS^G &= 100 \cdot 2^{-0.2777} \approx 82. \end{aligned} \tag{21}$$

Therefore, to achieve the same efficiency as that obtained with the pdf obtained from the WISMC model using 100 drawn it is necessary to execute 98 drawn from the real distribution. This shows that WISMC is very efficient. On the contrary the efficiency decreases for GJR model being necessary only 82 simulations and is worst for EGARCH whose efficiency is easily matched with only 52 drawn from the real distribution. These implications are amplified when dealing with the estimation of functionals of the recovery time R_t^s using the values of the KL divergence reported in Table 5 (case $s = 100 - M' = 30$ %). In this case a computation of the ESS produces the following result:

$$\begin{aligned} ESS^W &= 100 \cdot 2^{-0.0049} \approx 100, \\ ESS^E &= 100 \cdot 2^{-0.7289} \approx 60, \\ ESS^G &= 100 \cdot 2^{-1.1749} \approx 44. \end{aligned} \tag{22}$$

Similar conclusions hold true for the other cases discussed in Tables 4 and 5.

In addition, in order to have a more omprehensive evaluation of the models, in Tables 6 and 7 we provide results for the two-sample Kolmogorov–Smirnov test for both T_c^s and R_t^s, respectively (see e.g. Pratt et al., 1981). The test are executed at the significance level $\alpha = 1$ %. As it can be seen from Tables 6 and 7, the statistics for WISMC are always the lowest while the p-values for WISMC are always the highest, compared with the EGARCH and GJR ones. The WISMC model is the only one accepting the equalitarian hypothesis between the probability distribution functions from real data and simulated one. This is perfectly in line with what was already found with

Table 5
Kullback-Leibler divergence computed on the risk measure R_t^s as a function of M', fixing $M = 80$ % and $s = 100$ for Apple asset. The smallest distances are in bold.

Kullback-Leibler divergence for R_t^s with $M = 80$ % - APPLE	
$s = 100 - M' = 30$ %	
WISMC	**0.0049**
EGARCH(2,1)	0.7289
GJR(1,1)	1.1749
$s = 100 - M' = 40$ %	
WISMC	**0.0042**
EGARCH(2,1)	1.1032
GJR(1,1)	1.3359
$s = 100 - M' = 50$ %	
WISMC	**0.0117**
EGARCH(2,1)	1.7378
GJR(1,1)	1.6302

the Kullback-Leibler distances in Tables 4 and 5 Furthermore, for all the models the test rejects the null hypothesis of equality of the distribution from the real data and those from the model.

Tables 8 and 9 summarize the results of the model comparisons for T_c^s and R_t^s, respectively. The first column of Tables 8 and 9 shows, in percentage terms, how many times a model is chosen within its family. Note that the value related to the WISMC model is always 100 % because we consider only one score function f. If we had chosen additional f we would have probably identified other, better-performing sub-models on a case-by-case basis. For the EGARCH and GJR models, we can have all scores between 0 % and 100 %. Focusing on the first column of Table 8 for Apple stock, we note that the EGARCH (1,2) and GJR (1,1) models are the most selected models within their respective classes with 55 % and 78 % times.

The second column of Tables 8 and 9 displays the percentage each model is considered superior compared to the others in terms of Kullback-Leibler distance. As an example, let us consider the Apple case. In Table 8, we observe that WISMC has the smallest distance 78 % times. The remaining 22 % times the GJR(1,1) model, which is selected 78 % times as the best model within its class, has the shortest distance. Finally, the EGARCH class never has the smallest distance. Overall, the WISMC model is the best-performing model in terms of Kullback-Leibler divergence the most times for all stocks, both for T_c^s and R_t^s measures. Hence, our analysis results indicate that the WISMC model is an appropriate choice in modeling and simulating high-frequency financial volumes and, as already demonstrated, also for financial returns (see, e.g. D'Amico et al., 2020; D'Amico et al., 2012; D'Amico et al., 2018; De Blasis, 2023).The analysis also offers a broad perspective on the debate between employing EGARCH or GJR and WISMC. Actually, if we choose to use the GJR model class and consider Apple as an example in Table 8, we will find that 78 % of the time we choose GJR(1,1), which is only the 22 % of the time better than WISMC. The remaining 22 % of the time we would instead select a GJR (2,1) which however is always worse than the WISMC having in 0 % of cases a smaller KL divergence. This consideration holds also for Tesla and Netflix and for the other risk measures. This reflection also applied to Tesla and Netflix stocks for both T_c and R_t.

Table 6
Results of Two-samples Kolmogorov–Smirnov test for T_c^s at $\alpha = 1$ %.

Two-sample Kolmogorov–Smirnov test			
T_c^s-APPLE $s = 100 - M = 30$ %			
	p-values	Statistics	Test results
WISMC	0.0279	0.0650	Accept H_0
EGARCH(1,1)	2.3418e-42	0.3090	Reject H_0
GJR(2,1)	1.1809e-14	0.1800	Reject H_0
T_c^s-APPLE $s = 100 - M = 40$ %			
	p-values	Statistics	Test results
WISMC	0.0592	0.0590	Accept H_0
EGARCH(1,1)	1.3422e-36	0.2870	Reject H_0
GJR(2,1)	1.8654e-15	0.1850	Reject H_0
T_c^s-APPLE $s = 100 - M = 80$ %			
	p-values	Statistics	Test results
WISMC	0.0106	0.0720	Accept H_0
EGARCH(1,2)	1.4799e-44	0.3170	Reject H_0
GJR(1,1)	1.2116e-05	0.1090	Reject H_0

Table 7
Results of Two-samples Kolmogorov–Smirnov test for R_t^s at $\alpha = 1$ %.

Two-samples Kolmogorov–Smirnov test			
R_t^s-APPLE $s = 100$, $M = 80$ %, $M' = 30$ %			
	p-values	Statistics	Test results
WISMC	0.0186	0.0680	Accept H_0
EGARCH(1,1)	2.7120e-58	0.3630	Reject H_0
GJR(2,1)	3.1981e-86	0.4420	Reject H_0
R_t^s-APPLE $s = 100$, $M = 80$ %, $M' = 40$ %			
	p-values	Statistics	Test results
WISMC	0.7888	0.0290	Accept H_0
EGARCH(1,1)	1.5888e-61	0.3730	Reject H_0
GJR(2,1)	1.4447e-88	0.4480	Reject H_0
R_t^s-APPLE $s = 100$, $M = 80$ %, $M' = 50$ %			
	p-values	Statistics	Test results
WISMC	0.0666	0.0580	Accept H_0
EGARCH(1,2)	1.3413e-68	0.3940	Reject H_0
GJR(1,1)	2.1426e-71	0.4020	Reject H_0

The reason of why the WISMC model outperforms the EGARCH and GJR models can find an explanation in physical terms. In fact, the WISMC model has a very general structure (no parametric assumption is advanced) and very flexible in nature as it is able to change transition probability of the kernel (12) according to clustering behavior of the variable captured by the value u of the index process. All these empirical evidences encourage further theoretical and applied investigation of this class of models in finance.

Table 8
The number of times each model was chosen as the best model for the measure T_c^s and the number of times each model has the smallest Kullback-Leibler distance. Results refer to all the selected s and M values.

Scores for T_c^s-TESLA		
	Selected model (%)	Smallest KL (%)
WISMC	100	78
EGARCH(1,1)	22	0
EGARCH(1,2)	12	0
EGARCH(2,1)	66	0
GJR(1,1)	34	0
GJR(1,2)	44	22
GJR(2,1)	22	0
Scores for T_c^s-NETFLIX		
	Selected model (%)	Smallest KL (%)
WISMC	100	67
EGARCH(1,1)	55	0
EGARCH(1,2)	33	11
EGARCH(2,1)	12	0
GJR(1,1)	55	11
GJR(1,2)	12	0
GJR(2,1)	33	11
Scores for T_c^s-APPLE		
	Selected model (%)	Smallest KL (%)
WISMC	100	78
EGARCH(1,1)	45	0
EGARCH(1,2)	55	0
EGARCH(2,1)	0	0
GJR(1,1)	78	22
GJR(1,2)	0	0
GJR(2,1)	22	0

Table 9
The number of times each model was chosen as the best model for the measure R_t^s and the number of times each model has the smallest Kullback-Leibler distance. Results refer to all the selected s, M and M' values.

Scores for R_t^s-TESLA		
	Selected model (%)	Smallest KL (%)
WISMC	100	56
EGARCH(1,1)	89	11
EGARCH(1,2)	0	0
EGARCH(2,1)	11	0
GJR(1,1)	67	33
GJR(1,2)	22	0
GJR(2,1)	11	0
Scores for R_t^s-NETFLIX		
	Selected model (%)	Smallest KL (%)
WISMC	100	100
EGARCH(1,1)	44	0
EGARCH(1,2)	56	0
EGARCH(2,1)	0	0
GJR(1,1)	56	0
GJR(1,2)	33	0
GJR(2,1)	11	0
Scores for R_t^s-APPLE		
	Selected model (%)	Smallest KL (%)
WISMC	100	89
EGARCH(1,1)	44	0
EGARCH(1,2)	12	0
EGARCH(2,1)	44	0
GJR(1,1)	44	11
GJR(1,2)	44	0
GJR(2,1)	12	0

5. Conclusion

We investigate several risk indicators related to liquidity crises in high-frequency financial volumes of Tesla, Netflix, and Apple assets.

We generate synthetic volume time series by using the weighted indexed semi-Markov (WISMC) model and the EGARCH and GJR models. Then, we compute all drawdown-based risk measures on both real and artificial data and we explore them by means of parametric models whose estimation procedures are carried out considering the right censorship. Finally, we quantify the distance between real and simulated risk measures using the Kullback-Leibler divergence. Globally, the WISMC model performs better than the EGARCH and GJR models.

Since our results are essential for trading decisions based on liquidity risk, we can consider extending the investigation in several directions. Firstly, a joint analysis of volumes and returns can be performed using the multivariate WISMC model presented in D'Amico et al. (2021) and then also a forecasting analysis (see, e.g. Brownlees et al., 2011; Wilinski, 2019; Ghysels et al., 2000; Chuang et al., 2012). Secondly, the results can be applied for financial portfolio selection (see e.g. Brezina and Brezina, 2018). Finally, these risk measures could also be used in the valuation of real options (see, e.g. Bufalo et al., 2022).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jfds.2024.100138.

References

Alexander, N., Scherer, W., Thompson, J., 2023. Asset allocation using a markov process of clustered efficient frontier coefficients states. The Journal of Finance and Data Science, 100110.

Aragó, V., Nieto, L., 2005. Heteroskedasticity in the returns of the main world stock exchange indices: volume versus garch effects. J. Int. Financ. Mark. Inst. Money 15 (3), 271–284.

Baghdadabad, M.R.T., Glabadanidis, P., 2013. Average drawdown risk and capital asset pricing. Rev. Pac. Basin Financ. Mark. Policies 16 (4), 1350028.

Barbu, V.S., Karagrigoriou, A., Makrides, A., 2021. Semi-markov processes for earthquake forecast. Statistical Methods and Modeling of Seismogenesis 299–308.

Bhat, H.S., Kumar, N., 2010. Markov tree options pricing. In: Proceedings of the 2009 SIAM Conference on "Mathematics for Industry" the Art of "Mathematics for Industry". SIAM, pp. 162–173.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econom. 31 (3), 307–327.

Brezina, I., Brezina Jr., I., 2018. Drawdown as method of portfolio risk selection. In: Production Management and Business Development, pp. 43–46.

Brownlees, C.T., Cipollini, F., Gallo, G.M., 2011. Intra-daily volume modeling and prediction for algorithmic trading. J. Financ. Econom. 9 (3), 489–518.

Bufalo, M., Di Bari, A., Villani, G., 2022. Multi-stage real option evaluation with double barrier under stochastic volatility and interest rate. Ann. Finance 1–20.

Chuang, W.-I., Liu, H.-H., Susmel, R., 2012. The bivariate garch approach to investigating the relation between stock returns, trading volume, and return volatility. Global Finance J. 23 (1), 1–15.

De Blasis, R., 2020. The price leadership share: a new measure of price discovery in financial markets. Ann. Finance 16 (3), 381–405.

De Blasis, R., 2023. Weighted-indexed semi-markov model: calibration and application to financial modeling. Financial Innovation 9 (1), 1–16.

Dionne, G., Duchesne, P., Pacurar, M., 2009. Intraday value at risk (ivar) using tick-by-tick data with application to the toronto stock exchange. J. Empir. Finance 16 (5), 777–792.

D'Amico, G., Petroni, F., 2011. A semi-markov model with memory for price changes. J. Stat. Mech. Theor. Exp. 2011 (12), P12009.

D'Amico, G., Petroni, F., 2012. Weighted-indexed semi-markov models for modeling financial returns. J. Stat. Mech. Theor. Exp. 2012 (7), P07015.

D'Amico, G., Gismondi, F., Petroni, F., 2017a. A new approach to the modeling of financial volumes. In: Stochastic Processes and Applications: SPAS2017, Västerås and Stockholm, Sweden, pp. 363–373.

D'Amico, G., Di Biase, G., Janssen, J., Manca, R., 2017b. Semi-Markov Migration Models for Credit Risk.

D'Amico, G., Petroni, F., 2018. Copula based multivariate semi-markov models with applications in high-frequency finance. Eur. J. Oper. Res. 267 (2), 765–777.

D'Amico, G., Di Basilio, B., Petroni, F., 2020. A semi-markovian approach to drawdown-based measures. Adv. Complex Syst. 23 (8), 2050020.

D'Amico, G., Petroni, F., 2021. A micro-to-macro approach to returns, volumes and waiting times. Appl. Stoch Model Bus. Ind. 37 (4), 767–789.

D'Amico, G., Di Basilio, B., Petroni, F., Gismondi, F., September 30–October 2, 2023. An econometric analysis of drawdown based measures. In: Stochastic Processes, Statistical Methods, and Engineering Mathematics: SPAS 2019, Västerås, pp. 489–510. Sweden.

Fulga, C., Dedu, S., 2010. A new approach in multi-objective portfolio optimization using value-at-risk based risk measure. In: 2010 2nd IEEE International Conference on Information and Financial Engineering. IEEE, pp. 765–769.

Geboers, H., Depaire, B., Annaert, J., 2023. A review on drawdown risk measures and their implications for risk management. J. Econ. Surv. 37 (3), 865–889.

Ghysels, E., Gourieroux, C., Jasiak, J., 2000. Causality between returns and traded volumes. Ann. Econ. Stat. 189–206.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. J. Finance 48 (5), 1779–1801.

Goldberg, L.R., Mouti, S., 2022. Sustainable investing and the cross-section of returns and maximum drawdown. The Journal of Finance and Data Science 8, 353–387.

Graczyk, M.B., Duarte Queiros, S.M., 2016. Intraday seasonalities and nonstationarity of trading volume in financial markets: individual and cross-sectional features. PLoS One 11 (11), e0165057.

Grané, A., Veiga, H., 2014. Outliers, garch-type models and risk measures: a comparison of several approaches. J. Empir. Finance 26, 26–40.

Hongzhong, Z., et al., 2018. Drawdowns preceding drawups in a finite time-horizon. World Scientific Book Chapters 17–40.

Johnson, T.C., 2008. Volume, liquidity, and liquidity risk. J. Financ. Econ. 87 (2), 388–417.

Kaplan, E.L., Meier, P., 1958. Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc. 53 (282), 457–481.

Krokhmal, P., Uryasev, S., Zrazhevsky, G., 2005. Numerical comparison of conditional value-at-risk and conditional drawdown-at-risk approaches: application to hedge funds. In: Applications of Stochastic Programming, pp. 609–631.

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Ann. Math. Stat. 22 (1), 79–86.

Li, Z., Fan, P., Dong, Y., 2020. Flexible effective sample size based on the message importance measure. IEEE Open Journal of Signal Processing 1, 216–229.

Li, L., Zeng, P., Zhang, G., 2022. Speed and Duration of Drawdown under General Markov Models. Available at: SSRN 4222362.

Limnios, N., Oprisan, G., 2001. Semi-Markov Processes and Reliability.

Martino, L., Elvira, V., Louzada, F., 2017. Effective sample size for importance sampling based on discrepancy measures. Signal Process. 131, 386–401.

Masala, G., Petroni, F., 2023. Drawdown risk measures for asset portfolios with high frequency data. Ann. Finance 19 (2), 265–289.

Mercuri, L., 2008. Option pricing in a garch model with tempered stable innovations. Finance Res. Lett. 5 (3), 172–182.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica: J. Econom. Soc. 347–370.

Pratt, J.W., Gibbons, J.D., Pratt, J.W., Gibbons, J.D., 1981. Kolmogorov-smirnov two-sample tests. Concepts of nonparametric theory 318–344.

Queirós, S.M.D., 2016. Trading volume in financial markets: an introductory review, Chaos. Solitons & Fractals 88, 24–37.

Wattanatorn, W., Tansupswatdikul, P., 2019. An ability to forecast market liquidity–evidence from south east asia mutual fund industry. The Journal of Finance and Data Science 5 (1), 22–32.

Wilinski, A., 2019. Time series modeling and forecasting based on a Markov chain with changing transition matrices. Expert Syst. Appl. 133, 163–172.

Yuvan, S., Bier, M., 2018. A reaction–diffusion model for market fluctuations–a relation between price change and traded volumes. Phys. Lett. 382 (6), 367–371.

Zhang, H., Hadjiliadis, O., 2012. Drawdowns and the speed of market crash. Methodol. Comput. Appl. Probab. 14, 739–752.