High-frequency transaction data: a comparison between two asymmetric models

APPLIED FINANCE LETTERS

VOLUME 12, 2023

HIGH-FREQUENCY TRANSACTION DATA: A COMPARISON BETWEEN TWO ASYMMETRIC MODELS

MICHAEL KUNKLER[1*]

1. University of Reading, Berkshire, United Kingdom

* Corresponding Author: Michael Kunkler, Department of Economics, University of Reading, Berkshire, United Kingdom * m.kunkler@pgr.reading.ac.uk

Abstract

This paper compares two asymmetric models for high-frequency transaction data in financial markets, namely, the three-state Asymmetric Autoregressive Conditional Duration (AACD) model and the Activity Direction Size (ADS) model. It is shown that the two asymmetric models measure different aspects of the same underlying asymmetric nature of high-frequency transaction data. It is also shown that by extending the AACD model to include two size variables and adjusting for partial durations, each model's parameter estimates can be used to estimate the other model's parameters exactly. Thus, the two asymmetric models are equivalent, and measure the durations and price changes jointly.

Keywords: High-frequency transaction data

JEL: G15

1. Introduction

High-frequency transaction data in financial markets occurs in continuous time and is irregularly spaced. Models of high-frequency transaction data measure the durations and/or price changes of security. Some models focus on the inter-arrival times (durations) between high-frequency transactions and other models focus on the changes in the price of the security of interest. For example, by far the most common class of models for measuring the inter-arrival times between transactions is the class of Autoregressive Conditional Duration (ACD) models (Bauwens & Hautsch, 2009). More generally, it has been shown that durations and price changes can be modelled jointly as a marked point process (see Engle, 2000). As the number of models of high-frequency transaction data increases, it is important to reflect on the similarities of current models. The motivation of this paper is to highlight the underlying commonalities between two asymmetric models of high-frequency transaction data, namely, the Activity Direction Size (ADS) model (Rydberg and Shephard, 2003) and the three-state Asymmetric Autoregressive Conditional Duration (AACD) model (Tay et al., 2011). The three-state asymmetric autoregressive conditional duration is a generalisation of the Autoregressive Conditional Duration (ACD) model of Engle and Russell (1998). Thus, by comparing two generalised asymmetric models, such as the ADS model and the three-state AACD, further comparisons can be made to more specific models of high-frequency transaction data.

The Activity Direction Size (ADS) model decomposes high-frequency scaled price movements of a security into three main variables, namely, activity, direction, and size

(Rydberg and Shephard, 2003). For each transaction, *activity* measures whether the price moved (active) or was flat (inactive). Conditional on the price moving, *direction* measures whether the price moved up or down. Also conditional on the price moving, *size* measures the magnitude of the move in multiples of tick sizes. The ADS model captures the asymmetry between up and down price movements through both direction and size.

Inter-arrival times (durations) between transactions are usually modelled with conditional duration models, with the most well-known model being the Autoregressive Conditional Duration (ACD) model of Engle and Russell (1998). A two-state Asymmetric Autoregressive Conditional Duration (AACD) model of Bauwens and Giot (2003) extended the ACD model by including two durations, namely, up durations and down durations. In addition, a three-state Asymmetric Autoregressive Conditional Duration (AACD) of Tay et al. (2011) extended the two-state AACD model by including three durations, namely, flat durations, up durations, and down durations. Extensive surveys on conditional duration models demonstrate that it is an active area of research (see Pacurar, 2008; Bhogal and Ramanathan, 2019).

In this paper, the three-state AACD model is extended to mirror the two size variables of the ADS model, namely, an *up size* and a *down size*. In addition, by incorporating partial durations, this paper contributes to the literature by showing that each model's parameter estimates can be used to estimate the other model's parameters exactly. Thus, the two asymmetric models are equivalent and measure different aspects of the same asymmetric model. Ultimately, the two asymmetric models measure the durations and price changes jointly, which demonstrates that both models are more general than originally expected.

2. Material and Methods

2.1 Preliminaries

High-frequency transaction data in financial markets occurs in continuous time and is irregularly spaced. For example, when a transaction occurs, a price, a volume, and a time are recorded. Models of high-frequency transaction data measure both the durations and price movements of a security. More specifically, the price of a security within a specified time period can be written as:

$$P_{t_e} = P_{t_i} + \sum_{n=1}^{N} \Delta P_{t_n} \tag{1}$$

where P_{t_e} is the price of the security at the *end time* t_e, P_{t_i} is the price of the security at the *initial time* t_i, ΔP_{t_n} is the change in the price of the security between times t_n and t_{n-1}, t_n is the transaction's arrival time, and N is the number of transactions, with $t_i \leq t_1 \leq \cdots \leq t_n \leq \cdots \leq t_N \leq t_e$.

The prices are discrete and live on a lattice of prices driven by the tick size of the security (Rydberg and Shephard, 2003). The scaled price of a security at time t_n can be written as:

$$Z_{t_n} = P_{t_n}/\kappa \tag{2}$$

where Z_{t_n} is the *scaled price*, P_{t_n} is the price, and κ is the *tick size*. The tick size scales the price to be an integer lattice of scaled prices. In addition, the change in the scaled price of a security at time t_n can be written as:

$$\begin{aligned} \Delta Z_{t_n} &= Z_{t_n} - Z_{t_{n-1}} \\ &= P_{t_n}/\kappa - P_{t_{n-1}}/\kappa \\ &= \Delta P_{t_n}/\kappa \end{aligned} \tag{3}$$

where ΔZ_{t_n} is the change in the scaled price, ΔP_{t_n} is the change in the price with $\Delta P_{t_n} = P_{t_n} - P_{t_{n-1}}$, and κ is the tick size.

2.2 Partial Durations

Partial duration represents the time-interval (duration) between the end time and the last transaction time. For example, the total elapsed time can be written as:

$$T = t_e - t_i \tag{4}$$

where T is the *total time*, t_i is the initial time, and t_e is the end time. The *duration* (inter-arrival time) between two transactions is given by:

$$\Delta t_n = t_n - t_{n-1} \tag{5}$$

where Δt_n is the duration, t_n and t_{n-1} are transaction arrival times. The *expected unadjusted duration* can be estimated by:

$$\begin{aligned} \psi &= \tfrac{1}{N}\textstyle\sum_{n=1}^{N} \Delta t_n \\ &= \tfrac{1}{N}(t_N - t_i) \\ &= \tfrac{T_N}{N} \end{aligned} \tag{6}$$

where ψ is the expected unadjusted duration, Δt_n is the duration in Equation (5), $T_N = (t_N - t_i)$ and N is the number of transactions.

However, if a transaction doesn't occur at the end time, there exists a *partial duration*, which is the time since the last transaction at time t_N. The partial duration is given by:

$$\delta = t_e - t_N \tag{7}$$

where δ is the *partial duration*, t_N is the last transaction time, and t_e is the end time. Incorporating the partial duration into the expected duration estimation in Equation (6) produces:

$$\begin{aligned} \psi^* &= \tfrac{1}{N}\delta + \psi \\ &= \tfrac{1}{N}\delta + \tfrac{T_N}{N} \\ &= \tfrac{T_N}{N} \end{aligned} \tag{8}$$

where ψ^* is the expected *adjusted duration*, δ is the partial duration in Equation (7), ψ is the expected *unadjusted duration* in Equation (6), Δt_n is the duration in Equation (5), and T is the total time in Equation (3), with:

$$T = \delta + \textstyle\sum_{n=1}^{N} \Delta t_n$$

$$= t_e - t_i \tag{9}$$

When N is large, the partial duration δ plays a small role in the calculation of the expected duration in Equation (8), since:

$$\lim_{N \to \infty} \frac{1}{N} \delta = 0 \tag{10}$$

However, the partial duration plays an important role when comparing the two asymmetric models.

Assuming that the arrival times of transactions are exponentially distributed, the expected intensity is estimated by:

$$\begin{aligned} \lambda &= \frac{1}{\psi^*} \\ &= \frac{N}{T} \end{aligned} \tag{11}$$

where λ is the expected adjusted intensity, ψ^* is the expected adjusted duration in Equation (8), N is the number of transactions, and T is the total time in Equation (9).

2.3 Activity Direction Size (ADS) Model

The Activity Direction Size (ADS) decomposition of the change in the *scaled* price of a security at time t_n can be written as:

$$\Delta Z_{t_n} = A_{t_n} D_{t_n} S_{t_n} \tag{12}$$

where ΔZ_{t_n} is the change in the scaled price in Equation (3), A_{t_n} is the *activity*, D_{t_n} is the *direction*, and S_{t_n} is the *size* (Rydberg and Shephard, 2003).

The *activity* of a security represents whether a transaction moves the price or not. The probability that transactions are active (moves the price) can be estimated by:

$$\begin{aligned} p(A = 1) &= \frac{1}{N} \sum_{n=1}^{N} I(A_{t_n} = 1) \\ &= \frac{N_A}{N} \end{aligned} \tag{13}$$

where $p(A = 1)$ is the probability that transactions are active, A_{t_n} is the *activity* of the transaction price at time t_n, $I(A_{t_n} = 1)$ an indicator variable that is one when a transaction moves the price and zero otherwise; and N_A is the number of transactions where the price moved, with $N_A = \sum_{n=1}^{N} I(A_{t_n} = 1)$. In contrast, the probability that transactions are *flat* (not active) can be estimated by:

$$\begin{aligned} p(A = 0) &= 1 - p(A = 1) \\ &= 1 - \frac{N_A}{N} \\ &= \frac{N_F}{N} \end{aligned} \tag{14}$$

where $p(A = 0)$ is the probability that transactions are flat, $p(A = 1)$ is the probability that transactions are active in Equation (13), N_F is the number of flat transactions, N_A is the number of active transactions, and N is the total number of transactions, with $N = N_F + N_A$.

The *up direction* of a security represents whether a transaction moved the price up or not. The conditional probability for the up direction is given by:

$$\begin{aligned} p(D = 1|A = 1) &= \frac{1}{N_A}\sum_{a=1}^{N_A} I(D_{t_a} = 1) \\ &= \frac{N_U}{N_A} \end{aligned} \tag{15}$$

where $p(D = 1|A = 1)$ is the conditional probability that active transactions move the price up, D_{t_a} is the *direction* of the transaction price at time t_n, $I(D_{t_n} = 1)$ an indicator variable that is one when the price moves up and zero otherwise. Similarly, the *down direction* of a security represents whether a transaction moved the price up or not. The conditional probability for the down direction is given by:

$$\begin{aligned} p(D = -1|A = 1) &= \frac{1}{N_A}\sum_{a=1}^{N_A} I(D_{t_a} = -1) \\ &= \frac{N_D}{N_A} \end{aligned} \tag{16}$$

where $p(D = -1|A = 1)$ is the conditional probability that active transactions move the price down, D_{t_a} is the *direction* of the transaction price at time t_n, $I(D_{t_n} = -1)$ an indicator variable that is one when the price moves down and zero otherwise, and $N_A = N_U + N_D$ with:

$$p(D = 1|A = 1) + p(D = -1|A = 1) = 1 \tag{17}$$

Finally, the *size* of a security represents the magnitude of the price movement in multiples of tick sizes. Assuming information asymmetry between the magnitude of the up and down movements, *size* is usually separated into *up size* and *down size*. Typically, there is no *flat size*, as the conditional probability for flat size is given by:

$$p(S_F = 0|A = 0) = 1 \tag{18}$$

where $p(S_F = 0|A = 0)$ is the conditional probability that flat transactions do not move the price, which is one. The conditional probabilities for *up size* and *down size* are given by:

$$p(S_U = z|D = 1, A = 1) \tag{19}$$

$$p(S_D = z|D = -1, A = 1) \tag{20}$$

where $p(S_U = z|D = 1, A = 1)$ is the conditional probability that up transactions move the price up by z tick sizes, and $p(S_D = z|D = -1, A = 1)$ is the conditional probability that down transactions move the price down by z tick sizes.

Table 1 reports the ADS model for different values (z) of the scaled price change. The expected value of the change in the scaled price is given by:

$$\begin{aligned} E(\Delta Z) &= p(A=0)E(S_F) + p(A=1)p(D=1|A=1)E(S_U) - p(A=1)p(D=-1|A=1)E(S_D) \\ &= w_U E(S_U) - w_D E(S_D) \end{aligned} \quad (21)$$

where $E(\Delta Z)$ is the expected change in the scaled price of the security, $E(S_F)$ is the expected *flat size*, $E(S_U)$ is the expected *up size*, $E(S_D)$ is the expected *down size*, $w_U = p(A=1)p(D=1|A=1)$ is the up weight, and $w_D = p(A=1)p(D=-1|A=1)$ is the down weight. The $E(S_F)=0$ since by definition there is no price movement for flat transactions.

Table 1: Activity Direction Size (ADS) Model

States	z	$p(\Delta Z_{t_n} = z)$	Weights
Flat	$z=0$	w_F	$w_F = p(A=0)$
Up	$z = 1,2,..$	$w_U p(S_U = z)$	$w_U = p(A=1)p(D=1\|A=1)$
Down	$z=-1,-2,..$	$w_D p(S_D = z)$	$w_D = p(A=1)p(D=-1\|A=1)$

Notes: Table 1 reports the ADS model for different values (z) of the scaled price change with: $p(A=0)$ is the probability of *flat* transactions, $p(A=1)$ is the probability of *active* transactions, $p(D=1|A=1)$ is the conditional probability that *active* transactions move the price up, $p(D=-1|A=1)$ is the conditional probability that *active* transactions move the price down, w_F is the *flat* weight, w_U is the *up* weight, w_D is the *down* weight, $p(S_U=z)$ is the probability of an *up size* transaction equalling z, and $p(S_D=z)$ is the probability of a *down size* transaction equalling z.

2.4 Asymmetric Autoregressive Conditional Duration with Size (AACDS)

Inter-arrival times (durations) between transactions are typically modelled with conditional duration models, such as the Autoregressive Conditional Duration (ACD) model of Engle and Russell (1998). The three-state AACD model of Tay et al. (2011) extended the two-state AACD of Bauwens and Giot (2003) to include three durations, namely, *flat durations* (*F*), *up durations* (*U*), and *down durations* (*D*).

A size variable was not included in the original three-state AACD model, as less than 0.5% of the transactions moved more than one tick (Tay et al., 2011). However, in this paper, two size variables are included in the three-state AACD model to mirror the size variables of the ADS model, namely, an *up size* and a *down size*. The extended three-state AACD with size model will be referred to as the *three-state AACDS model*.

The standard ACD model can be written in terms of a marked point process (see Engle, 2000). In this context, the *three-state AACDS model* can be written as three marked point processes by:

$$\begin{aligned} p(\psi_F^*, S_F) &= p(\psi_F^*)p(S_F|\psi_F^*) = p(\psi_F^*) \\ p(\psi_U^*, S_U) &= p(\psi_U^*)p(S_U|\psi_U^*) \\ p(\psi_D^*, S_D) &= p(\psi_D^*)p(S_D|\psi_D^*) \end{aligned} \quad (22)$$

where $p(\psi_F^*, S_F)$, $p(\psi_U^*, S_U)$, and $p(\psi_D^*, S_D)$ are joint probability distributions of the associated adjusted durations and scaled price changes for flat transactions, up transactions, down transactions, respectively. All flat size movements are equal to zero, so that $p(S_F = 0|\psi_F^*) = 1$.

Using Equation (8), the expected adjusted duration for the three states can be written as:

$$\begin{aligned}\psi_F^* &= \frac{T}{N_F}\\ \psi_U^* &= \frac{T}{N_U}\\ \psi_D^* &= \frac{T}{N_D}\end{aligned} \tag{23}$$

where ψ_F^* is the expected adjusted *flat* duration for N_F flat durations, ψ_U^* is the expected adjusted *up* duration for N_U up durations, ψ_D^* is the expected adjusted *down* duration for N_D down durations, and T is the total time in Equation (9).

Assuming that the arrival times of the transactions of the three states are exponentially distributed and using Equation (11), the expected intensities are estimated by:

$$\begin{aligned}\lambda_F &= \frac{1}{\psi_F^*} = \frac{N_F}{T}\\ \lambda_U &= \frac{1}{\psi_U^*} = \frac{N_U}{T}\\ \lambda_D &= \frac{1}{\psi_D^*} = \frac{N_D}{T}\end{aligned} \tag{24}$$

where λ_F is the expected *flat intensity*, λ_U is the expected *up intensity*, λ_D is the expected *down intensity*, and the other terms are from Equation (23). The expected intensity for all transactions can be estimated by:

$$\begin{aligned}\lambda &= \lambda_F + \lambda_U + \lambda_D\\ &= \frac{1}{T}(N_F + N_U + N_D)\\ &= \frac{N}{T}\\ &= \frac{1}{\psi^*}\end{aligned} \tag{25}$$

where λ is the expected intensity of all N transactions, with $N = N_F + N_U + N_D$, ψ^* is the expected *adjusted duration*, and the other terms are from Equation (23) and Equation (24).

Table 2 reports the AACDS model for different values (z) of the change in the scaled price. The expected value of the scaled price change can be written as:

$$\begin{aligned}E(\Delta Z) &= \frac{\lambda_F}{\lambda_F+\lambda_U+\lambda_D}E(S_F) + \frac{\lambda_U}{\lambda_F+\lambda_U+\lambda_D}E(S_U) - \frac{\lambda_D}{\lambda_F+\lambda_U+\lambda_D}E(S_D)\\ &= w_F E(S_F) + w_U E(S_U) - w_D E(S_D)\\ &= w_U E(S_U) - w_D E(S_D)\end{aligned} \tag{26}$$

where $E(\Delta Z)$ is the expected change in the scaled price of the security, $E(S_F)$ is the expected *flat size*, $E(S_U)$ is the expected *up size*, $E(S_D)$ is the expected *down size*, $w_F = \frac{\lambda_F}{\lambda_F+\lambda_U+\lambda_D}$ is the flat weight, $w_U = \frac{\lambda_U}{\lambda_F+\lambda_U+\lambda_D}$ is the up weight, and $w_D = \frac{\lambda_D}{\lambda_F+\lambda_U+\lambda_D}$ is the down weight. The $E(S_F) = 0$ since by definition there is no price movement for flat transactions.

Table 2: Three-state AACDS Model

States	z	$p(\Delta Z_{t_n} = z)$	Weights
Flat	$z = 0$	w_F	$w_F = \frac{\lambda_F}{\lambda_F+\lambda_U+\lambda_D}$
Up	$z = 1,2,..$	$w_U p(S_U = z)$	$w_U = \frac{\lambda_U}{\lambda_F+\lambda_U+\lambda_D}$
Down	$z = -1,-2,..$	$w_D p(S_D = z)$	$w_D = \frac{\lambda_D}{\lambda_F+\lambda_U+\lambda_D}$

Notes: Table 2 reports the AACDS model for different values (z) of the change in the scaled price with: λ_F is the *flat* (or *inactive*) intensity, λ_A is the total number of *active* intensity, λ_U is the *up* intensity, λ_D is the *down* intensity, w_F is the *flat* weight, w_U is the *up* weight, w_D is the *down* weight, $p(S_U = z)$ is the probability of an *up size* transaction equalling z, and $p(S_D = z)$ is the probability of a *down size* transaction equalling z.

2.5 Comparison

In this section, it is shown that the estimated parameters of each asymmetric model can be used to estimate the parameters of the other model. It is also shown that if the partial durations are used, the models are identical. The three-state Asymmetric Autoregressive Conditional Duration with Size (AACDS) model consists of three durations (flat, up, and down) and two size variables (up and down). The ADS model consists of two indicator variables (activity and direction) and two size variables (up and down).

The probabilities associated with the activity and direction of the ADS model can be written in terms of both the expected adjusted durations and the expected intensities of the three-state AACDS model by:

$$p(A = 0) = \frac{N_F}{N} = \frac{N_F}{T}\frac{T}{N} = \frac{\psi^*}{\psi_F^*} = \frac{\lambda_F}{\lambda} \tag{27}$$

$$p(A = 1) = \frac{N_A}{N} = \frac{N_A}{T}\frac{T}{N} = \frac{\psi^*}{\psi_A^*} = \frac{\lambda_A}{\lambda} \tag{28}$$

$$p(D = 1|A = 1) = \frac{N_U}{N_A} = \frac{N_U}{T}\frac{T}{N_A} = \frac{\psi_A^*}{\psi_U^*} = \frac{\lambda_U}{\lambda_A} \tag{29}$$

$$p(D = -1|A = 1) = \frac{N_D}{N_A} = \frac{N_D}{T}\frac{T}{N_A} = \frac{\psi_A^*}{\psi_D^*} = \frac{\lambda_D}{\lambda_A} \tag{30}$$

where $\psi_A^* = \frac{T}{N_A}$ is the expected adjusted *active* duration for $N_A = N_U + N_D$ active durations, $\lambda_A = \frac{N_A}{T} = \frac{1}{\psi_A^*}$ is the expected *active intensity*, and all other terms have been previously described. Thus, the estimated parameters of the three-state AACDS model can be used to estimate the parameters of the ADS model.

Similarly, the three expected intensities of the three-state AACDS can be written in terms of the estimated probabilities of the ADS model by:

$$\lambda_F = \frac{N_F}{T} = \frac{N_F}{N}\frac{N}{T} = \frac{N}{T}p(A = 0) \tag{31}$$

$$\lambda_U = \frac{N_U}{T} = \frac{N_U}{N}\frac{N}{T} = \frac{N}{T}p(A = 1)p(D = 1|A = 1) \tag{32}$$

$$\lambda_D = \frac{N_D}{T} = \frac{N_D}{N}\frac{N}{T} = \frac{N}{T}p(A = 1)p(D = -1|A = 1) \tag{33}$$

where all terms have been previously described. Thus, the estimated parameters of the ADS model can be used to estimate the parameters of the three-state AACDS model. It should be noted that because all duration models account for partial durations, the common total time (T) makes the two asymmetric models equivalent.

Table 3: Comparison of the two asymmetric models

States	z	$p(\Delta Z_{t_n} = z)$	ADS Weights	AACDS Weights
Flat	$z = 0$	w_F	$w_F = p(A = 0)$	$w_F = \frac{\lambda_F}{\lambda_F+\lambda_U+\lambda_D}$
Up	$z = 1,2,..$	$w_U p(S_U = z)$	$w_U = p(A = 1)p(D = 1\|A = 1)$	$w_U = \frac{\lambda_U}{\lambda_F+\lambda_U+\lambda_D}$
Down	$z = -1,-2,..$	$w_D p(S_D = z)$	$w_D = p(A = 1)p(D = -1\|A = 1)$	$w_D = \frac{\lambda_D}{\lambda_F+\lambda_U+\lambda_D}$

Notes: Table 3 reports both the ADS and the AACDS model for different values (z) of the scaled price change with: $p(A = 0)$ is the probability of *flat* transactions, $p(A = 1)$ is the probability of *active* transactions, $p(D = 1|A = 1)$ is the conditional probability that *active* transactions move the price up, $p(D = -1|A = 1)$ is the conditional probability that *active* transactions move the price down, λ_F is the *flat* (or *inactive*) intensity, λ_A is the total number of *active* intensity, λ_U is the *up* intensity, λ_D is the *down* intensity, w_F is the *flat* weight, w_U is the *up* weight, w_D is the *down* weight, $p(S_U = z)$ is the probability of an *up size* transaction equalling z, and $p(S_D = z)$ is the probability of a *down size* transaction equalling z.

In general, the expectation of the scaled change in price for both models can be written as:

$$\begin{aligned} E(\Delta Z) &= w_F E(S_F) + w_U E(S_U) - w_D E(S_D) \\ &= w_U E(S_U) - w_D E(S_D) \end{aligned} \tag{34}$$

where $E(\Delta Z)$ is the expected change in the scaled price of the security, $E(S_F)$ is the expected *flat size*, $E(S_U)$ is the expected *up size*, and $E(S_D)$ is the expected *down size*. The $E(S_F) = 0$ since by definition there is no price movement for flat transactions.

2.6 Predictive (conditional) models

It should be noted that more sophisticated dynamic models are typically used to estimate both asymmetric models. In this paper, both asymmetric models have been described as simple in-sample models. Models can be estimated *contemporaneously* (in-sample) or *predictively* (out-of-sample). In addition, most applications of both asymmetric markets have been used for predictions. For example, the joint probability function at time t_n for a are *predictive* model can be written as:

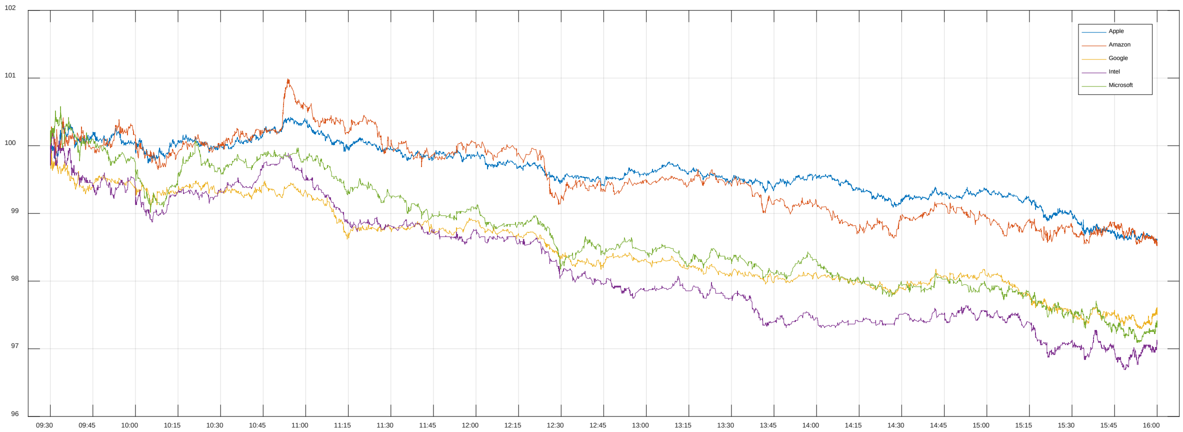
$$p(\psi_{t_n}, S_{t_n} | \mathcal{F}_{t_{n-1}}) \tag{35}$$

where ψ_n is the duration, S_{t_n} is the size, and all explanatory variables used in the model exist in the *filtration* $\mathcal{F}_{t_{n-1}}$: past *information*. However, the motivation of this paper is to compare the variables of the two asymmetric models, rather than focus on any particular version of the models.

3. Results

The data sample of transaction prices is a single day (21st June 2012) sourced from Lobster for five US securities, namely, Apple, Amazon, Google, Intel, and Microsoft. Figure 1 displays the prices rebased at 100. All securities depreciated over the day, where the depreciations were -1.38% for Apple, -1.48% for Amazon, -2.42% for Google, -2.94% for Intel, and -2.66% for Microsoft.

Figure 1: Security prices rebased at 100



Notes: Figure 1 displays the transaction prices for the 21st of June, 2012 for the five securities, namely, Apple, Amazon, Google, Intel, and Microsoft.

Table 4 reports summary statistics of the high-frequency transaction data for the prices of the five securities. The tick size κ for all five securities is 0.005. The total time T for all five securities is 23,400 seconds (6.5 hours). The securities with the three highest number of transactions are: Apple at 34,990, Microsoft at 33,414, and Intel at 32,483. In contrast, the securities with the two lowest number of transactions are: Amazon at 11,419 and Google at 11,678. Interesting, both Intel and Microsoft have a large number of transactions that do not move the price with 29,693 from 32,483 and 30,218 from 33,414, respectively.

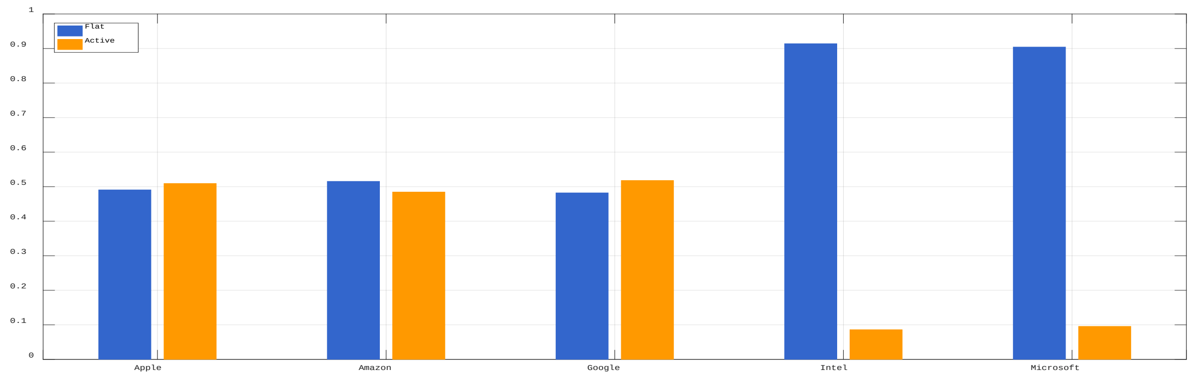
Table 4: Summary Statistics of the Security Prices

	Apple	Amazon	Google	Intel	Microsoft
κ	0.005	0.005	0.005	0.005	0.005
T	23,400	23,400	23,400	23,400	23,400
N	34,990	11,419	11,678	32,483	33,414
N_F	17,175	5,886	5,631	29,693	30,218
N_A	17,815	5,533	6,047	2,790	3,196
N_U	8,948	2,921	2,973	1,397	1,545
N_D	8,867	2,612	3,074	1,393	1,651

Notes: Table 4 reports the summary statistics of the high-frequency transaction data for the prices of the five securities, which consists of: κ is the tick size, T is the total time, N is the total number of transactions, N_F is the total number of *flat* (or *inactive*) transactions, N_A is the total number of *active* transactions, N_U is the total number of *up* transactions, and N_D is the total number of *down* transactions.

Table 5 reports the parameter estimates of the Activity Direction Size (ADS) model. Figure 2 displays the activity of the transaction prices, displaying both the probability of flat transaction prices ($p(A = 0)$) and the probability of active transaction prices ($p(A = 1)$). Intel and Microsoft both have high probabilities for flat transaction prices with values of 0.914 and 0.904. Thus, over 90% of the transactions associated with Intel and Microsoft do not move the price. In contrast, the conditional probability for both directions are all close to 0.500. For example, the largest conditional probability difference of 0.056 is for Amazon, which has a conditional probability of an up direction of 0.528 compared to a conditional probability of a downward direction of 0.472. The size variables are common to both asymmetric models and will be discussed later.

Figure 2: Activity Probabilities



Notes: Figure 2 displays the activity of the transaction prices for the 21st of June, 2012 for the five securities, namely, Apple, Amazon, Google, Intel, and Microsoft. *Flat* represents the probability of flat transaction prices and *Active* represents the probability active transaction prices.

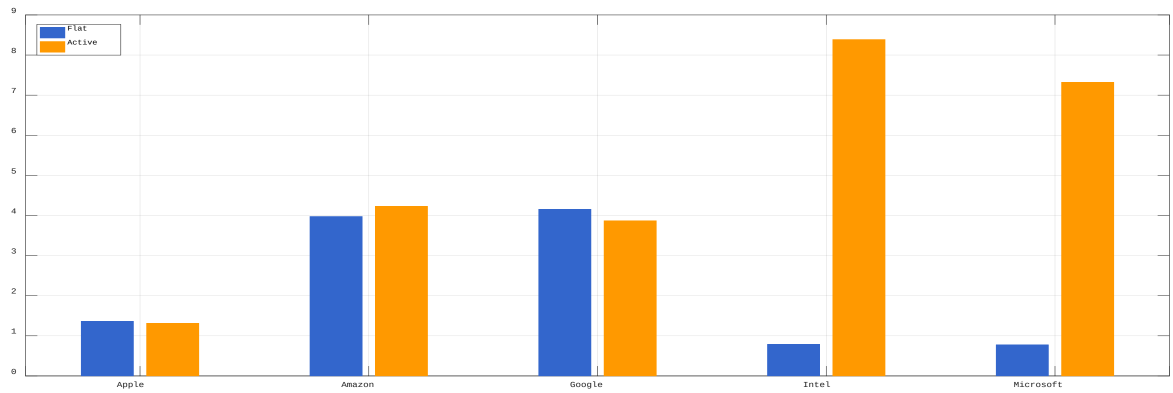
Table 5: Activity Direction Size (ADS) model

	Apple	Amazon	Google	Intel	Microsoft
$p(A = 0)$	0.491	0.516	0.482	0.914	0.904
$p(A = 1)$	0.509	0.485	0.518	0.086	0.096
$p(D = 1\|A = 1)$	0.502	0.528	0.492	0.501	0.483
$p(D = -1\|A = 1)$	0.498	0.472	0.508	0.499	0.517
S_U	6.537	6.045	10.747	1.395	1.548
S_D	6.780	7.013	11.308	1.515	1.548

Notes: Table 5 reports the parameter estimates of the Activity Direction Size (ADS) model, which consists of: $p(A = 0)$ is the probability of *flat* transactions, $p(A = 1)$ is the probability of *active* transactions, $p(D = 1|A = 1)$ is the conditional probability that *active* transactions move the price up, $p(D = -1|A = 1)$ is the conditional probability that *active* transactions move the price down, S_U is the expected up size, and S_D is the expected *down* size.

Table 6 reports the parameter estimates of the Asymmetric Autoregressive Conditional Duration with Size (AACDS) model. Figure 3 displays the expected durations of the transaction prices, displaying both the expected adjusted flat durations (ψ_F^*) and the expected adjusted active durations (ψ_A^*). The smallest expected adjusted flat durations are 0.788 seconds for Intel and are 0.774 seconds for Microsoft. The same two securities also have the largest expected adjusted active durations, which are 8.387 seconds for Intel and are 7.322 seconds for Microsoft. In contrast, the expected adjusted duration for both directions are very similar for all securities. For example, the largest expected duration difference of 0.973 seconds is for Microsoft, which has an expected adjusted up duration of 15.146 seconds compared to an expected adjusted down duration of 14.173. Note that all of the intensities are the inverse of their associated expected adjusted durations. For example, Apple has the highest intensity for all transactions of $\lambda = 1.495$, which is the inverse of the expected adjusted duration for all transactions of $\psi^* = 1/\lambda = 0.669$. One again, the size variables are common to both asymmetric models and will be discussed later.

Figure 3: Expected Durations in Seconds



Notes: Figure 3 displays the expected durations of the transaction prices for the 21st of June 2012 for the five securities, namely, Apple, Amazon, Google, Intel, and Microsoft.

Table 6: Asymmetric Autoregressive Conditional Duration with Size (AACDS) model

	Apple	Amazon	Google	Intel	Microsoft
ψ^*	0.669	2.049	2.004	0.720	0.700
ψ_F^*	1.362	3.976	4.156	0.788	0.774
ψ_A^*	1.314	4.229	3.870	8.387	7.322
ψ_U^*	2.615	8.011	7.871	16.750	15.146
ψ_D^*	2.639	8.959	7.612	16.798	14.173
λ	1.495	0.488	0.499	1.388	1.428
λ_F	0.734	0.252	0.241	1.269	1.291
λ_A	0.761	0.237	0.258	0.119	0.137
λ_U	0.382	0.125	0.127	0.060	0.066
λ_D	0.379	0.112	0.131	0.060	0.071
S_U	6.537	6.045	10.747	1.395	1.548
S_D	6.780	7.013	11.308	1.515	1.548

Notes: Table 6 reports the parameter estimates for the Asymmetric Autoregressive Conditional Duration with Size (AACDS) model, which consists of: ψ^* is the expected *adjusted* duration, ψ_F^* is the expected *adjusted flat* duration, ψ_F^* is the expected *adjusted active* duration, ψ_A^* is the expected *adjusted up* duration, ψ_U^* is the expected duration, ψ_D^* is the expected *adjusted down* duration, λ is the intensity, λ_F is the *flat* (or *inactive*) intensity, λ_A is the total number of *active* intensity, λ_U is the *up* intensity, λ_D is the *down* intensity, S_U is the expected up size, and S_D is the expected *down* size.

The asymmetric models are equivalent and measure different aspects of the same asymmetric nature of high-frequency transaction data. Each model's parameter estimates can be used to estimate the other model's parameters exactly. For example, the expected adjusted flat duration for the AACDS model is 0.788 seconds for Intel and can be calculated from the parameter estimates of the ADS model by:

$$\psi_F^* = \frac{T}{N}\frac{1}{p(A=0)} = \frac{0.720}{0.914} = 0.788 \tag{36}$$

where ψ_F^* is the expected adjusted *flat* duration, ψ^* is the expected adjusted duration for all transaction prices, and $p(A = 0)$ is the probability that transactions are flat. Similarly, the probability that transactions are active for the ADS model is 0.086 for Intel and can be calculated from the parameter estimates of the AACDS model by:

$$p(A = 1) = \frac{\psi^*}{\psi_A^*} = \frac{0.720}{8.387} = 0.086 \tag{37}$$

where $p(A = 1)$ is the probability that transactions are active, ψ^* is the expected adjusted duration for all transaction prices, and ψ_A^* is the expected adjusted *active* duration.

Table 7: Common parameter estimates of the two asymmetric models

	Apple	Amazon	Google	Intel	Microsoft
w_F	0.491	0.516	0.482	0.914	0.904
w_U	0.256	0.256	0.255	0.043	0.046
w_D	0.253	0.229	0.263	0.043	0.049
S_U	6.537	6.045	10.747	1.395	1.548
S_D	6.780	7.013	11.308	1.515	1.548
$\boldsymbol{E(\Delta Z)}$	**-0.046**	**-0.058**	**-0.241**	**-0.005**	**-0.005**

Notes: Table 7 reports the common parameter estimates of the two asymmetric models consisting of: w_F is the *flat* weight, w_U is the *up* weight, w_D is the *down* weight, S_U is the expected *up size*, S_D is the expected *down size*, and $E(\Delta Z)$ is the expected change in the scaled price.

Table 7 reports the common parameter estimates of both asymmetric models. The expected values for both size variables are larger than one tick size for all securities, which justifies the inclusion of a size variable in the three-state Asymmetric Autoregressive Conditional Duration (AACD) model. Google has the highest expected tick sizes of 10.747 for up moves and 11.308 for down moves. In contrast, Intel has the lowest expected tick sizes of 1.395 for up moves and 1.515 for down moves. In addition, Amazon has the largest difference of 0.968 between the up size variable and the down size variable.

The expected values of $E(\Delta Z)$ for each security is negative, where the values are -0.046 for Apple, -0.058 for Amazon, -0.241 for Google, -0.005 for Intel, and -0.005 for Microsoft. Google has the largest expected value which can be seen by using Equation (34):

$$\begin{aligned} E(\Delta Z) &= w_U E(S_U) + w_D E(S_D) \\ &= 0.255 \times 10.747 - 0.263 \times 11.308 \\ &= -0.241 \end{aligned} \tag{38}$$

where $E(\Delta Z)$ is the expected change in the scaled price of the security, $E(S_F)$ is the expected *flat size*, $E(S_U)$ is the expected *up size*, and $E(S_D)$ is the expected *down size*.

Conclusion

This paper compared two asymmetric models of high-frequency transaction data in financial markets, namely, the Activity Direction Size (ADS) model and the three-state Asymmetric Autoregressive Conditional Duration with Size (AACDS) model. It was shown that both asymmetric models are equivalent and measure different aspects of the same asymmetric nature of high-frequency transaction data. The size variables plays an integral part of both asymmetric models, as the magnitude of price changes occur in multiples of the underlying tick size. Thus, the inclusion of two size variables in the AACD model extends it to model durations and price changes jointly: creating a more general model. The implication of this paper is that researchers can compare the parameter estimates of one model with the parameter estimates of the other model, especially when more sophisticated dynamic models are used: one model provides a yardstick for the other.

References

Bauwens, L., Giot, P., 2003. Asymmetric ACD Models: Introducing Price Information in ACD Models. Empirical Economics, 28, 709–731.

Bauwens, L., Hautsch, N., 2009. Modelling financial high frequency data using point processes. In Handbook of financial time series (pp 953-979). Berlin, Heidelberg: Springer Berlin Heidelberg.

Bhogal, S.K., Ramanathan, T.V., 2019. Conditional duration models for high-frequency data: A review on recent developments. Journal of Economic Surveys, 33, 1, 252–273.

Engle, R.F., Russell, J.R., 1998. Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. Econometrica, 66, 5, 1127–1162.

Engle, R.F., 2000. The Econometrics of Ultra-High-Frequency Data. Econometrica, 68, 1, 1–22.

Pacurar, M., 2008. Autoregressive conditional duration models in finance: a survey of the theoretical and empirical literature. Journal of Economic Surveys, 22, 711–751.

Rydberg, T.H., Shephard, N., 2003. Dynamics of trade-by-trade price movements: decomposition and models. Journal of Financial Econometrics, 1, 1, 2–25.

Saart, P.W., Gao, J. Allen, D.E., 2015. Semiparametric autoregressive conditional duration model: Theory and practice. Econometric Reviews, 34, 6–10, 849–881.

Tay, A.S., Ting, C., Tse, Y.K., Warachka, M., 2011. The impact of transaction duration, volume and direction on price dynamics and volatility. Quantitative Finance, 11, 3, 447-457.