2011

The impact of transaction duration, volume and direction on price dynamics and volatility

Anthony S. TAY
Singapore Management University, anthonytay@smu.edu.sg

Christopher TING
Singapore Management University, christophert@smu.edu.sg

Yiu Kuen TSE
Singapore Management University, yktse@smu.edu.sg

Mitchell WARACHKA
Singapore Management University, mitchell@smu.edu.sg

Published in Quantitative Finance, 2011, Volume 11, Issue 3 , Pages 447-457
http://doi.org/10.1080/14697680903405742

The impact of transaction duration, volume and direction on price dynamics and volatility

ANTHONY S. TAY*†, CHRISTOPHER TING‡, YIU KUEN TSE†
and MITCH WARACHKA‡

†School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903
‡Lee Kong Chian School of Business, Singapore Management University, 50 Stamford Road, Singapore 178899

We explore the role of trade volume, trade direction, and the duration between trades in explaining price dynamics and volatility using an Asymmetric Autoregressive Conditional Duration model applied to intraday transactions data. Our results suggest that volume, direction and duration are important determinants of price dynamics, while duration is also an important determinant of volatility. However, the impact of volume and direction on volatility is marginal after controlling for duration, and the impact of volume on volatility appears to be confined to periods of infrequent trading.

Keywords: Econometric theory; Applied econometrics; Econometrics of financial markets; Forecasting ability

1. Introduction

We study the role of trade volume (trade size of each transaction), trade direction (buy versus sell order), and the duration between trades, in explaining price dynamics and volatility. We explore these relationships using the Asymmetric Autoregressive Conditional Duration (AACD) model of Bauwens and Giot (2003) applied to intraday transactions data. Two versions of this model are used: the first uses only lagged conditional expected duration and lagged realized duration as explanatory variables for the conditional expected duration between trades. The second version adds trade volume, trade direction, as well as the interaction of these two variables with lagged duration, to the list of explanatory variables. From these models of conditional expected duration, we derive implications for price dynamics and volatility.

We compare our results with both the theoretical literature on market microstructure as well as the empirical literature relating volume and duration to price dynamics and volatility. One of the two main market microstructure theories that we address is that of Easley and O'Hara (1992), who argue that both the presence and absence of trade can provide useful information to participants regarding the presence and absence of information. Short durations and large orders indicate the presence of information, whereas an absence of trade indicates no valuable information, and in either case market makers adjust prices accordingly. One outcome is that duration and volume should be correlated with price variance. We also relate our results to Diamond and Verrecchia (1987), where short sale constraints reduce the adjustment speed of prices to bad news in particular, resulting in longer durations causing downward biases in returns.

Our empirical results suggest that trade duration, size and direction are important determinants of price dynamics. For instance, down-ticks are more likely than up-ticks after long durations, which is in line with the prediction of Diamond and Verrecchia (1987). We also find that trade duration is an important determinant of volatility, which is consistent with the prediction of Easley and O'Hara (1992), and reaffirms the empirical findings of other papers that also investigate this relationship, for example Manganelli (2005) and Hautsch (2008). However, our results concerning volume and volatility are different than that in other studies. For example, Hautsch (2008) finds strong evidence of a common unobservable component driving volatility and volume, and Manganelli (2005) finds evidence that higher volume leads to higher volatility for frequently traded stocks. We find that the size of a transaction in general has an

*Corresponding author. Email: anthonytay@smu.edu.sg

insignificant effect on volatility in active markets, which supports the empirical findings of Jones *et al.* (1994), but is in contrast to the predictions of Easley and O'Hara (1992). In inactive markets, however, we find that larger transaction size leads to lower volatility.

The AACD model is an extension of the Autoregressive Conditional Duration (ACD) model, introduced by Engle and Russell (1998) and Engle (2000). Such models and their extensions have proven very useful for analysing irregularly spaced data, and the literature on ACD models has expanded rapidly, with recent contributions by Grammig and Maurer (2000), Zhang *et al.* (2001), Engle and Lunde (2003), Ghysels *et al.* (2004), and Fernandes and Grammig (2005), among others. However, most of these focus on the dynamic properties of the durations between trades, and do not consider the interrelationship between duration and other trade-related variables. An exception is Bauwens and Veredas (2004), who relate price durations to spread, trade intensity, and volume. Our paper is a contribution following, and extending, this line of research.

In the next section, we summarize the AACD model as applied to tick changes in transactions, and discuss the price dynamics of the implemented AACD model. Section 3 describes our data. Section 4 interprets our empirical results and highlights their salient market microstructure implications. Our conclusions are summarized in section 5.

2. The model and its implications for price dynamics and volatility

We consider a three-state AACD model with possible price movements of one tick down, no tick change and one tick up. While some trades may occur beyond one tick size, they are very infrequent in our sample and a three-state model appears to be adequate. We assume that the sequence of trades is determined in the following way. With the occurrence of the $(i-1)$th trade, there are three possibilities: the ith trade may be a trade at one tick down, no tick change, or one tick up. These three potential tick movements each follow a latent stochastic point process whose inter-arrival times have independent exponential distributions. The observed tick movement is the outcome of a competition among the three underlying point processes to be the first arrival.

A full discussion of the model, including properties and estimation, can be found in the appendix. Here we focus on the main features of the model. Let the index $i = 1, 2, 3, \ldots$ denote the order of trades, t_i denote the time of the ith trade, and $x_i = t_i - t_{i-1}$ denote the duration of the ith transaction. Let w_i denote the tick movement of the ith trade, where w_i may take values $j = -1, 0, 1$, representing one tick down, no tick change and one tick up, respectively. The information set after the $(i-1)$th trade is denoted by Φ_{i-1}. This may consist of past tick movements, volumes of transactions and lagged durations. The random duration for the ith (potential) tick movement is denoted by T_{ji}, where $j = -1, 0, 1$. We assume that, conditional on Φ_{i-1}, the three T_{ji} are independently distributed as exponential variables with mean ψ_{ji}. The reciprocal of ψ_{ji} is called the intensity, and is denoted by λ_{ji}.

The AACD model centers on the conditional expected duration ψ_{ji}. Our basic model is

$$\ln \psi_{ji} = \sum_{k=-1}^{1} v_{jk} D_k(w_{i-1}) + \alpha_j \ln \psi_{j,i-1} + \beta_j \ln x_{i-1}, \quad j = -1, 0, 1, \tag{1}$$

where $D_k(z) = 1$, if $z = k$ and 0 otherwise. Note that an increase (decrease) in the conditional expected duration ψ_{ji} implies a smaller (larger) intensity λ_{ji}, which in turn reduces (increases) the probability that the next transaction is of type j.

The intercepts v_{jk} in equation (1) represent the sensitivity of the next price movement to the prior transaction. When the previous tick movement is of type k, the intercept for $\ln \psi_{ji}$ is v_{jk}. A larger (smaller) v_{jk} implies that tick movement k induces a lower (higher) intensity of the next tick being of type j. However, the resulting probability distribution for price movements depends on the relative magnitudes of v_{jk}.

To incorporate the effects of trade direction and trade volume as well as their interactions with realized lagged durations, we also implement an augmented AACD model with the conditional expected durations given by

$$\begin{aligned}\ln \psi_{ji} = \sum_{k=-1}^{1} v_{jk} D_k(w_{i-1}) + \alpha_j \ln \psi_{j,i-1} + \beta_j \ln x_{i-1} + \gamma_j \ln s_{i-1} \\ + \varphi_j y_{i-1} + \theta_j (y_{i-1} \ln s_{i-1}) + \eta_j (y_{i-1} \ln x_{i-1}) \\ + \xi_j (y_{i-1} \ln x_{i-1} \ln s_{i-1}),\end{aligned} \tag{2}$$

for $j = -1, 0, 1$. The s_{i-1} terms denote the trade size (volume in lots) of the last transaction whose trade direction is represented as $y_{i-1} = \pm 1$ according to the usual convention of 1 for buy-initiated trades and -1 for sell-initiated trades. Therefore, the conditional information set is extended to $\Phi_{i-1} = \{t_h, w_h, s_h, y_h;\ h = 1, \ldots, i-1\}$ with additional parameters measuring the sensitivity of the conditional durations to various trade variables and their interactions.

For example, a positive θ_j implies that large buy orders increase the conditional duration of tick j. Once again, whether a large buy order actually reduces the probability of the next tick being j depends on the relative magnitudes of θ_j. Similarly, if tick movement j has the largest η, the probability of tick j decreases after a buy trade following a long duration. The ξ coefficients capture the interaction between trade direction, trade size, and trade frequency. Overall, the augmented model allows the influence of trade variables on price dynamics to be examined individually as well as jointly.

We now consider the price dynamics implied by the AACD model, as well as the corresponding conditional return and return volatility. Suppressing the time index, we denote the expected duration of a price decrease, no price change, and a price increase as ψ_{-1}, ψ_0, and ψ_1, respectively. Under the AACD framework, the stock price between time t and T ($>t$) evolves as

$$P(T) = P(t) + \delta \sum_{i=1}^{n} w_i, \tag{3}$$

where n is the random number of trades between time t and T while δ denotes one tick size. Note that the number of trades n between time t and T may equal zero, implying $P(t+s) = P(t)$ for $0 \leq s \leq T-t$.

We next examine the instantaneous expected return and return variance implied by the AACD price dynamics. Following standard practice, these quantities are defined respectively as

$$R(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathrm{E}\left[\frac{P(t+\Delta t) - P(t)}{P(t)}\right] \tag{4}$$

and

$$\sigma^2(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathrm{Var}\left[\frac{P(t+\Delta t) - P(t)}{P(t)}\right]. \tag{5}$$

For a small time interval Δt, the probability of a trade occurring is $(\lambda_{-1} + \lambda_0 + \lambda_1)\,\Delta t$. The probability of more than one trade occurring is of an order higher than Δt. Hence, dropping terms of $O((\Delta t)^2)$ and higher, we have

$$\mathrm{E}[P(t+\Delta t) - P(t)] = \delta\,(\lambda_1 - \lambda_{-1})\,\Delta t\,, \tag{6}$$

from which we obtain

$$R(t_{i-1}) = \frac{\delta\,(\lambda_{1i} - \lambda_{-1,i})}{P(t_{i-1})} \tag{7}$$

after including the trading time index. Note that $\lambda_{-1,i}$ and λ_{1i} are based on information available at time t_{i-1}. Similarly, up to $O(\Delta t)$ the variance of a price change is given by

$$\mathrm{Var}[P(t+\Delta t) - P(t)] = \delta^2\,(\lambda_1 + \lambda_{-1})\,\Delta t\,, \tag{8}$$

implying the instantaneous return variance, with the time index specified, equals

$$\sigma^2(t_{i-1}) = \left(\frac{\delta}{P(t_{i-1})}\right)^2 (\lambda_{1i} + \lambda_{-1,i}). \tag{9}$$

The instantaneous variance in equation (9) parallels the result of Engle and Russell (1998) for unmarked durations which signify price changes irrespective of their direction.

3. Data

We apply the AACD model to intraday data on five NYSE companies: Boeing (BA), General Electric (GE), International Business Machines (IBM), Altria Group (formerly Philip Morris) (MO), and AT&T (T). These five firms are a subset of those studied by Dufour and Engle (2000). The data is obtained from the TAQ database for July 1, 1994 to June 30, 1995. This sample period is chosen since there are no changes in the minimum tick size and average durations are stable. The NYSE traded with a tick size of one-eighth during our sample period. This tick size subsequently changed to one-sixteenth before decimalization. In later trading regimes, price movements beyond one tick became more common. However, the AACD model may be extended to incorporate additional up and down ticks. In addition, tick movements may be categorized into groups with the average tick size within each group defining a mark. This approach compromises the precision of the tick size, but reduces the number of parameters requiring estimation. Our sample period is chosen due to the stability of the tick size, and is probably less susceptible to parametric instability. The optimal selection of groups to categorize tick movements in later regimes is an important topic for future research.

We extract three variables on each stock: time of trade, transaction price, and signed volume inferred using the Lee and Ready (1991) algorithm. We also correct for the opening auction and for time-of-day effects, using procedures similar to those of Engle and Russell (1998). In particular, opening effects require the transactions occurring in the first 20 minutes of each day to be removed. The average duration for transactions over the following 10 minutes serves as the waiting time for the first trade after 10:00 a.m. (E.S.T.). All transactions recorded after 4:00 p.m. are deleted. In some cases, the opening transaction occurred after the first 20 minutes. Also, on a few days there are insufficient transactions between 9:50 a.m. and 10:00 a.m. to obtain a meaningful average starting duration. Therefore, days with opening transactions after 9:50 a.m. and with less than three transactions over the next 10 minutes are removed, along with November 25, 1994 due to an early 'day after Thanksgiving' closing. Even after these deletions, a tremendous number of observations for each company remain, as documented in table 1.

We estimate diurnal factors by applying a smoothing spline to the average duration at each time point with available data. We used the MATLAB function *csaps.m* to compute the smoothing spline. The diurnal factor is adjusted to ensure the sample mean of the diurnally adjusted durations is equal to the sample mean of the non-diurnally adjusted data. The diurnally adjusted durations are then formed by dividing each duration by the corresponding diurnal factor. For the remainder of this paper, durations x_i refer to mean-diurnally adjusted durations. The diurnal factors for all five duration series are similar to those of Engle and Russell (1998). In particular, the diurnal factors initially increase, with the largest diurnal factor occurring in the middle of the day, before decreasing.

Relevant summary statistics for our data are given in table 1. The number of observations available for BA is substantially lower than the other stocks due to less frequent trading as indicated by its average duration. For each stock, the distribution of price changes is fairly symmetric with the majority of trades occurring at the prevailing price. Price movements of more than one tick occur less than 0.5% of the time. Therefore, although the AACD model is easily extended to incorporate more than three marks, this extension would only complicate the

Table 1. Summary statistics for price movements, durations and order flow.

Statistics	BA	GE	IBM	MO	T
Frequency (%) of price movements					
2 ticks up or more	0.19	0.15	0.16	0.23	0.11
1 tick up	14.70	15.17	8.79	14.91	13.36
0 tick—no price change	70.35	69.18	82.10	69.84	72.94
1 tick down	14.59	15.38	8.82	14.81	13.54
2 ticks down or more	0.16	0.07	0.13	0.20	0.05
Average diurnally adjusted duration (in seconds)					
All trades $\bar{x}$	88.78	31.83	41.42	48.88	39.29
Trades at uptick $\hat{\psi}_1$	596.10	207.20	462.80	322.60	291.60
Trades at prevailing price $\hat{\psi}_0$	126.20	46.02	50.46	69.98	53.88
Trades at downtick $\hat{\psi}_{-1}$	601.80	206.00	462.70	325.70	289.20
Order flow statistics (volume in lots)					
Frequency of buys (%)	44.87	57.63	47.99	44.32	54.35
Frequency of sells (%)	55.13	42.37	52.01	55.68	45.65
Serial correlation of trade direction	0.35	0.32	0.52	0.32	0.40
Runs test of trade direction	−81.32	−132.56	−186.27	−105.77	−146.61
Average volume (lot size)	27.80	19.91	30.83	31.48	25.31
Average log volume	1.97	1.70	2.36	2.13	1.61
Average daily number of trades	243.30	677.90	521.10	442.30	549.10
Number of observations in sample	54,500	170,157	129,239	110,120	135,087

estimation process without any significant effect on our results.

Table 1 also records the average durations between consecutive price increases, decreases, and trades executed at the prevailing price. Since the conditional expected duration for mark j equals

$$E[\psi_{ji}|\Phi_{i-1}] = \frac{1}{\lambda_{ji}} = \frac{1/(\lambda_{-1,i}+\lambda_{0i}+\lambda_{1i})}{\lambda_{ji}/(\lambda_{-1,i}+\lambda_{0i}+\lambda_{1i})} = \frac{E[x_i|\Phi_{i-1}]}{f_{ji}}, \quad (10)$$

we estimate the average durations as $\bar{x}/\widehat{f_j}$ to initialize the ACMD process when estimating the model using MLE, where f_{ji} is defined in equation (3) and $\widehat{f_j}$ is simply the unconditional probability of mark j occurring. The estimation is performed using the CML program in GAUSS with standard errors computed using the robust QMLE method.

4. Empirical market microstructure implications

Following Engle and Russell (1998), diagnostics such as the Box–Ljung statistics for the standardized durations (i.e. duration divided by the conditional expected duration) and their squared counterparts are calculated, along with the excess dispersion statistic. These are not reported here, but we note that our results are comparable to Engle and Russell (1998) as well as Engle (2000) with a steep decline in the statistics after raw durations are transformed into standardized durations using the AACD model.

4.1. Results for price dynamics

Estimates of the basic model for each of the five companies are summarized in table 2. With the exception of one parameter, all estimates are statistically significant at the 5% level, attesting to the autoregressive role of trade frequency and its importance to price dynamics. Two other conclusions are also apparent. First, the estimated coefficients exhibit remarkable resemblance over the five stocks. Second, $\alpha_0+\beta_0$ is smaller than $\alpha_{-1}+\beta_{-1}$ and $\alpha_1+\beta_1$ for all stocks, suggesting higher persistence in the conditional durations of transactions executed at the prevailing prices.

Empirical results from the augmented model are recorded in table 3. Once again, the estimated coefficients exhibit remarkable resemblance among the five firms. Furthermore, the following regularities are observed for each firm.

(1) $\nu_{-1,-1} > \nu_{1,-1}$ and $\nu_{11} > \nu_{-1,1}$, implying that a downtick induces a greater probability of an uptick and *vice versa*. Thus, returns exhibit negative serial correlation as in Dufour and Engle (2000). This property likely results from bid–ask bounce.

(2) $\beta_1 > \beta_{-1}$, providing evidence that long durations induce greater probabilities for downticks relative to upticks. This supports the hypothesis of Diamond and Verrecchia (1987) with no news meaning bad news. This property is also exhibited in the basic model.

(3) $\eta_1 > \eta_0 > \eta_{-1}$, with $\eta_1 > 0$ and $\eta_{-1} < 0$. Thus, a buy trade after a long duration induces a greater probability of a downtick. This negative return is consistent with Dufour and Engle (2000). We also observe that $\varphi_1 > \varphi_0 > \varphi_{-1}$, indicating that a buy trade *per se* implies a greater probability of a subsequent downtick. This finding contrasts with Dufour and Engle (2000), although the disparity likely reflects the fact that φ_j measures the effect of

Table 2. Parameter estimates of the basic AACD model.

Price movements and previous trade variables	Parameter	BA	GE	IBM	MO	T
Downtick, $j=-1$						
Downtick	$v_{-1,-1}$	3.8836	3.1717	3.4067	3.4321	2.7866
No tick	$v_{-1,0}$	2.3687	1.4542	2.4796	1.6648	0.8215
Uptick	$v_{-1,1}$	0.4857	−0.4684	0.6118	−0.3523	−1.6580
Conditional distribution	α_{-1}	0.5919	0.7296	0.5587	0.6756	0.8389
Lagged duration	β_{-1}	0.1146	0.0406	0.1273	0.1176	0.0812
No Change, $j=0$						
Downtick	$v_{0,-1}$	0.1325	0.7714	0.1908	0.2067	0.2055
No tick	$v_{0,0}$	0.0913	0.5132	0.1332	0.1032	0.0772
Uptick	$v_{0,1}$	0.1030	0.5162	0.0748	0.0983	0.0479
Conditional duration	α_0	0.9488	0.8011	0.9205	0.9318	0.9466
Lagged duration	β_0	0.0440	0.0757	0.0576	0.0528	0.0413
Uptick, $j=1$						
Downtick	$v_{1,-1}$	0.7675	−0.9531	0.8017	**−0.0674**	−1.5746
No tick	$v_{1,0}$	2.3242	1.6428	2.7918	1.6939	1.1352
Uptick	$v_{1,1}$	3.8195	3.8559	3.9710	3.3834	3.2728
Conditional duration	α_1	0.5830	0.6984	0.5023	0.6552	0.7844
Lagged duration	β_1	0.1199	0.0791	0.1458	0.1308	0.1037

Notes. The basic model is given in equation (5). All coefficients for the five firms are significant at the 1% level, with the exception of one parameter highlighted in bold whose p-value is 0.19.

Table 3. Parameter estimates of the augmented AACD model.

Price movements and previous trade variables	Parameter	BA	GE	IBM	MO	T
Downtick, $j=-1$						
Downtick	$v_{-1,-1}$	5.8477	5.1988	4.6592	5.1329	3.6559
No tick	$v_{-1,0}$	4.7032	3.7769	4.0497	3.7231	1.4422
Uptick	$v_{-1,1}$	3.3392	2.2638	2.3623	2.1743	−1.3977
Conditional duration	α_{-1}	0.2488	0.3574	0.3102	0.3582	0.7416
Lagged duration	β_{-1}	0.1844	0.0933	0.1960	0.1851	0.0994
Volume	γ_{-1}	−0.0983	−0.0253	−0.0414	−0.1204	0.0298
Direction	φ_{-1}	−0.4757	−0.6167	−0.3969	−0.5396	−0.3815
Volume-direction	θ_{-1}	0.1331	0.0612	0.1147	0.1076	0.0736
Duration-direction	η_{-1}	−0.1405	−0.0986	−0.1740	−0.1201	−0.0138
Volume-duration-direction	ξ_{-1}	0.0169	0.0191	0.0297	0.0162	**0.0041**
No change, $j=0$						
Downtick	$v_{0,-1}$	0.2380	2.0816	0.3026	0.9567	0.3759
No tick	$v_{0,0}$	0.2013	1.8067	0.2533	0.8460	0.2270
Uptick	$v_{0,1}$	0.2075	1.8313	0.1793	0.9166	0.1912
Conditional duration	α_0	0.9230	0.5200	0.8946	0.7486	0.9027
Lagged duration	β_0	0.0522	0.0772	0.0661	0.1097	0.0598
Volume	γ_0	−0.0165	−0.1294	−0.0180	−0.0762	−0.0194
Direction	φ_0	0.0427	−0.0493	**0.0083**	**0.0172**	−0.0356
Volume-direction	θ_0	−0.0150	**−0.0005**	**−0.0039**	−0.0143	0.0087
Duration-direction	η_0	**−0.0042**	**−0.0023**	**0.0035**	0.0086	0.0061
Volume-duration-direction	ξ_0	**0.0019**	**0.0016**	**−0.0003**	**0.0011**	**−0.0012**
Uptick, $j=1$						
Downtick	$v_{1,-1}$	3.1884	2.4102	2.3678	1.5281	**−0.2315**
No tick	$v_{1,0}$	4.3806	4.4190	4.1433	3.1296	2.5050
Uptick	$v_{1,1}$	5.6828	6.2825	5.1680	4.6729	4.5850
Conditional duration	α_1	0.2789	0.2716	0.2781	0.4399	0.5634
Lagged duration	β_1	0.1869	0.1697	0.2514	0.2086	0.1694
Volume	γ_1	−0.0317	−0.1019	**−0.0090**	−0.0887	**0.0023**
Direction	φ_1	0.4009	0.8715	0.2785	0.3577	0.6152
Volume-direction	θ_1	−0.1315	−0.0832	−0.0793	−0.1084	−0.1092
Duration-direction	η_1	0.1154	0.1823	0.2128	0.1196	0.0476
Volume-duration-direction	ξ_1	**−0.0038**	−0.0450	−0.0347	−0.0118	**−0.0033**

Notes: The augmented model is given in equation (6). The majority of coefficients in the augmented model are significant at the 1% level for all five firms. Coefficients that are not significant at the 5% level are denoted in boldface.

Table 4. Distribution of instantaneous volatilities.

Ticker	5th	Mean	95th
BA	3.09	21.71	51.93
GE	5.04	33.84	85.38
IBM	3.37	16.83	43.02
MO	4.06	25.34	61.24
T	3.28	26.79	71.93

trade direction net of its interaction with other variables such as volume and duration.

(4) $\theta_{-1} > \theta_0 > \theta_1$, implying that a large buy trade increases the probability of an uptick relative to a downtick. A similar effect is produced by a large buy trade after a long duration since $\xi_{-1} > \xi_0 > \xi_1$. Together with the previous items, these results demonstrate that the price impact of a trade has to be studied in conjunction with its direction, size, and frequency.

Regarding volume, we observe that $\gamma_j < 0$ for all tick movements and stocks, except for γ_{-1} and γ_1 for stock T. Therefore, when volume increases, trading activity becomes more intense, although the relative magnitudes of γ are ambiguous.

4.2. *Results for instantaneous volatility*

We examine the instantaneous volatility at each in-sample transaction time using equation (9) with $\delta = 1/8$ and $P(t_{i-1})$ equaling realized transaction prices. For ease of comparison, we scale the instantaneous volatilities by $\sqrt{260 \times 6.5 \times 60 \times 60}$ to obtain annualized percentages. We summarize their fifth percentile, mean, and 95th percentile in table 4. The mean volatilities (percent per year) reported here are comparable to figures computed from lower frequency daily return data, offering additional empirical support for the AACD model specification. Observe that GE is the most volatile stock, while IBM is the least volatile, a fact consistent with the number of zero-tick movements in table 1.

We compute the instantaneous volatility in different scenarios categorized by trade direction, size, and frequency, as well as associated tick movements and market conditions. We did this for all five stocks, but we report only results for IBM; details on the other stocks are available upon request. We consider scenarios with long and short durations and conditional durations, and for large and small volumes. These are defined precisely in table 5, e.g. long duration refers to the 90th percentile, which for IBM is 100.98.

Table 6 summarizes the instantaneous volatility, in percent per annum. As the results for all five stocks are similar, we only present the results for IBM. In table 6, the column 'short duration/long conditional durations' corresponds to a trade arriving unexpectedly soon in an inactive market. Conversely, the column 'long duration/short conditional durations' characterizes a trade arriving unexpectedly late in an active market. The rows signify three aspects of the scenarios, corresponding to trade direction (purchase or sale), price movement (uptick, downtick, or no tick), and trade volume (small or large). As sale uptick and purchase downtick are exceptionally infrequent, their results are not presented.

Although certain volatilities in table 6 appear large, they coincide with artificially constructed scenarios that do not prevail over long periods. The majority of trades are executed with neutral tick movements and have associated volatilities that agree with our intuition. Several observations emerge from table 6.

(1) Volatility increases with shorter conditional durations. Thus, volatility is higher during active markets with frequent trading. This is similar to the result found by Manganelli (2005).

(2) Volatility increases with shorter lagged durations. Consequently, shorter arrival times induce higher volatility, irrespective of the market's conditional duration (active or inactive).

(3) A purchase uptick with small volume induces a higher volatility than with large volume, with sales exhibiting a similar pattern for downticks. Although this phenomenon may seem counter-intuitive, purchases coinciding with upticks (or sales coinciding with downticks) may signal price revisions upwards (downwards). If higher volume creates a stronger signal, then the price has a higher probability of remaining at the revised level, without reversing in the short term. Thus, higher volume increases the probability that the next trade corresponds to a neutral tick movement. This result supports the empirical findings of Jones *et al*. (1994) that volume has an insignificant role in determining volatility, and is in contrast to empirical studies such as those of Manganelli (2005) and Hautsch (2008). Indeed, the AACD model reveals a potential explanation for why higher volume actually reduces volatility for certain transactions.

(4) The volatilities for long duration/short conditional durations are larger than those of short duration/long conditional durations in all cases. Thus, an unexpectedly late trade arrival in an active market induces a higher volatility than an unexpected trade in an inactive market.

Overall, consistent with Easley and O'Hara (1992), high volatility coincides with market conditions characterized by short durations. In addition, the effect of trade direction on volatility appears ambiguous.

4.3. *Impulse response analysis*

We now consider the impacts of transaction duration and volume on volatility over longer horizons (of up to 5 minutes). We simulate the price paths of the impulse response functions of the augmented AACD model. These simulations analyse the effects of trade direction, size, and frequency on volatility dynamics. As starting values, the previous realized duration is set equal to its unconditional sample mean. The starting values for the conditional durations of the three processes are then

Table 5. Summary statistics of volatility scenarios (IBM).

	Volume	Duration	Conditional duration (downtick)	Conditional duration (no change)	Conditional duration (uptick)
Small (10th percentile)	1	4.19	222.27	31.66	227.12
Medium (50th percentile)	10	22.28	645.60	48.19	650.14
Large (90th percentile)	75	100.98	2668.50	72.48	3011.71

Notes: The values define small, medium, and large trades, short and long durations, as well and short and long conditional expected durations for each of the three possible tick movements.

Table 6. Instantaneous volatilities of BA and IBM in various scenarios.

	Duration/conditional durations			
Scenario	Short/short	Short/long	Long/short	Long/long
BA				
Purchase-no tick-small volume	41.14	27.40	35.66	23.60
Purchase-no tick-large volume	41.60	28.00	30.62	20.46
Purchase-uptick-small volume	73.18	48.24	67.36	44.38
Purchase-uptick-large volume	64.62	42.66	52.20	34.42
Sale-no tick-small volume	38.76	26.60	32.02	22.10
Sale-no tick-large volume	41.12	27.74	30.64	20.82
Sale-downtick-small volume	63.50	43.98	55.56	38.52
Sale-downtick-large volume	52.96	36.54	44.32	30.64
IBM				
Purchase-no tick-small volume	30.02	20.54	26.40	17.98
Purchase-no tick-large volume	28.14	19.38	19.94	13.68
Purchase-uptick-small volume	62.00	42.20	59.44	40.44
Purchase-uptick-large volume	49.18	33.50	38.30	26.08
Sale-no tick-small volume	29.04	20.16	25.20	17.56
Sale-no tick-large volume	28.98	19.96	20.56	14.16
Sale-downtick-small volume	62.86	43.86	58.60	40.90
Sale-downtick-large volume	50.38	35.08	37.04	25.80

Notes. The figures are the instantaneous volatility (annualized) in various scenarios. The columns state the percentile of the lagged duration and the conditional expected duration (e.g., 'short/long' means short lagged duration and long conditional duration) while the rows state three aspects of the scenario corresponding to trade direction, price movement, and volume.

computed by dividing this value by the unconditional probabilities $\widehat{f_j}$ of the three marks (sample proportions). Thus, the conditional durations are $\psi_j = \bar{x}/\widehat{f_j}$ for $j = -1, 0, 1$, where $\bar{x}$ is the sample mean of the transaction durations.

Simulation of the augmented model requires the exogenous trade direction and size variables, denoted y_i and s_i respectively. To generate these variables from the sample data, we adopt a re-sampling procedure. Based on the entire sample of data $\{y_i, s_i\}$, indexed by $i = 1, \ldots, N$, we randomly select an integer m from 1 to $N-B+1$, where B represents a specified block size. Once m is selected, the series $\{y_i, s_i\}$, for $i = m, \ldots, m+B-1$, is drawn as the trade direction and volume for the next B tick movements. The conditional expected durations are then computed to yield (w_i, t_i). After B transactions, another integer m is randomly selected as the starting point of another block of trade direction and volume data. We choose B to be 50, approximately 10% of IBM's daily transactions.

A total of 10,000 trials are conducted for each scenario, with the first 600 seconds deleted to ensure a 'natural state' has been reached. After this period, a transaction of a particular direction and size is introduced whose realized lagged duration is also an element of the scenario. We then study the dynamics of the instantaneous volatility over the subsequent five-minute period.

Of the scenarios we investigate, two are characterized by whether the trade is a purchase or sale. Both of these scenarios are conducted for small and large transactions, with further refinements corresponding to whether the previous realized durations are short or long, as defined in table 5.

Figures 1 and 2 present the average instantaneous volatility over 10,000 simulated paths using estimates of the augmented model for IBM. The figures indicate that a transaction occurring after a short duration increases volatility, irrespective of its size and direction. Indeed, the marginal contribution of trade size and direction to volatility appears very limited in active markets. The relationship between high volatility and short durations confirms the predictions of Easley and O'Hara (1992).

However, for a trade occurring after a long duration, volatility tends to decline after high volume transactions regardless of trade direction. Thus, with short durations (high trade frequency), volume has a secondary role in volatility dynamics. Conversely, in inactive markets (low trade frequency), volume cannot be ignored. This phenomenon also applies to the other four stocks (additional figures available on request). An explanation for this

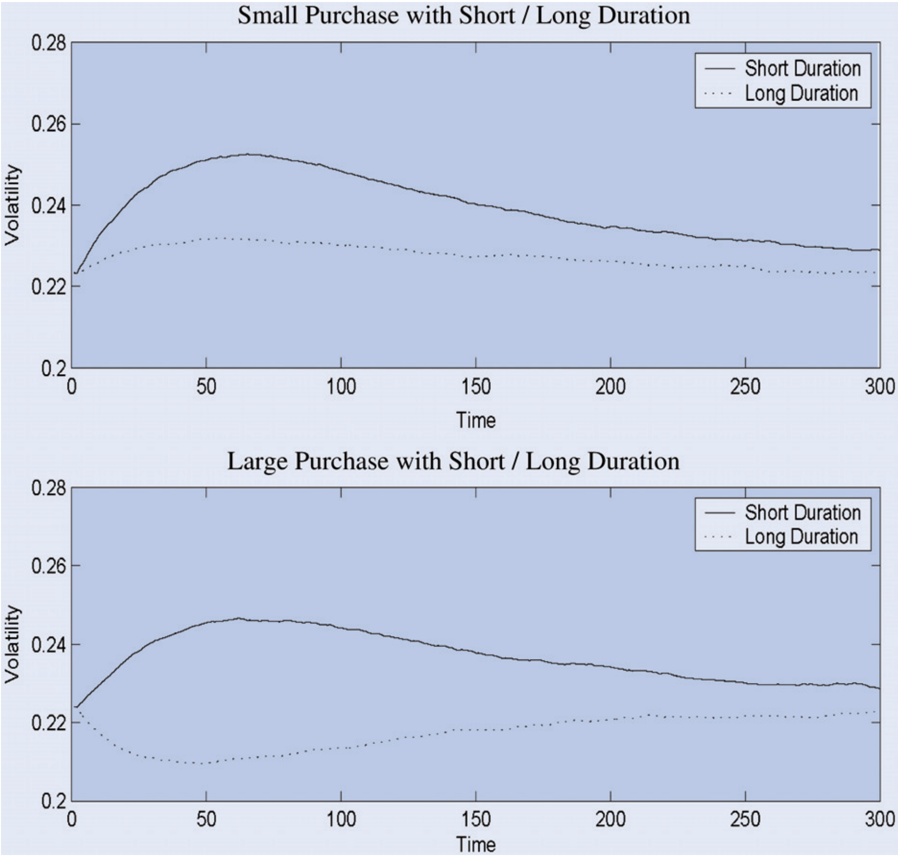

Figure 1. Plots of annualized volatility for IBM purchase scenarios.
Notes: Plots of annualized volatility for IBM purchase scenarios in transactions time. Small and large purchases have similar volatility impacts, while different durations have distinct impacts on volatility.

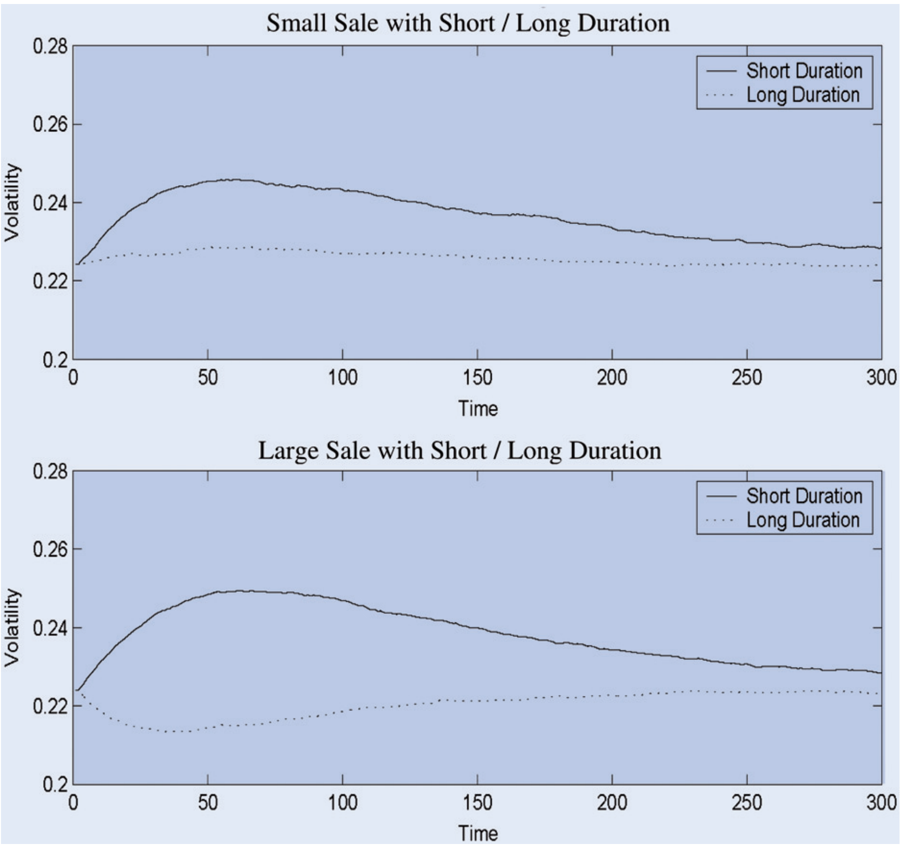

Figure 2. Plots of annualized volatility for IBM sale scenarios.
Notes: Plots of annualized volatility for IBM sale scenarios in transactions time. Small and large sales have similar volatility impacts, while different durations have distinct impacts on volatility.

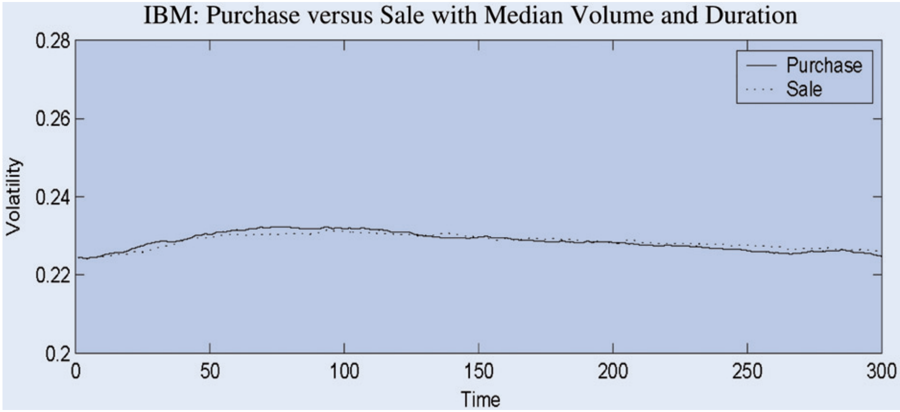

Figure 3. Plots of BA and IBM volatility after purchase and sale. Notes: Plots of IBM volatility after purchase and sale at median volume and median durations in transactions time. For both firms, trade direction does not appear to exert a large influence on volatility with purchases and sales provoking similar reactions in the instantaneous volatility.

result is that high volume implies greater certainty that the stock's price reflects its true value, reducing the likelihood of an immediate price reversal. In contrast, small volume transactions after long durations may signify a greater possibility of a price reversal. To our knowledge, the interaction between duration and volume has not been previously documented.

Figure 3 displays the results for trade direction in market conditions that are characterized by median trade intensity and size. Observe that volatility's response to a sale or purchase for IBM is identical. This is generally true for all five stocks: at most there is only a very slight difference between sale and purchase scenarios.

5. Conclusions

We use an AACD model to explore the impact of the characteristics of past trades, such as direction, volume, duration, and their interactions, on subsequent arrival times and price movements. We find evidence in support of Diamond and Verrecchia (1987) that infrequent trading is consistent with bad news. We find that trade frequency has a significant influence on volatility dynamics. Consistent with Easley and O'Hara (1992), we find that active markets defined by shorter durations experience higher volatility. In contrast, the role of trade direction and size is less salient. In particular, the impact of volume on volatility dynamics is confined to periods of infrequent trading. Thus, examining the importance of volume to volatility requires a transactions level analysis as aggregating volume over longer periods obscures its contribution in periods of inactive trading.

The AACD approach that we employ in this paper is one of a number of approaches that jointly model the duration with the price process. Others models that do so include those of Russell (1999), who proposes the Autoregressive Conditional Intensity (ACI) model, which is a multivariate dependent point process focusing on the conditional intensity rather than conditional durations, and Russell and Engle (2004), who propose an autoregressive conditional multinomial model for the price process. In this paper we use the AACD to explore the relationship between trading volume, trading intensity, and trade direction. A better understanding of these market microstructure issues may help ultimately in better market design. However, the applicability of this approach is wider. Our model can be used, for instance, to help understand and predict changes in intraday prices and volatility, and elsewhere (Tay *et al.* 2009) we employ the same approach to trade direction to compute intraday Probabilities of Informed Trading (PINs).

Acknowledgements

This paper, and its companion paper 'Modeling transaction data of trade direction and estimation of probability of informed trading', supersedes a previously circulated paper entitled 'An autoregressive conditional marked duration model for transaction data', which was presented at the First Symposium on Econometric Theory and Application (SETA), Taipei, 2005. We thank all who have commented on these papers, and, in particular, two anonymous referees for their very helpful comments.

References

Bauwens, L. and Giot, P., Asymmetric ACD models: Introducing price information in ACD models. *Empir. Econ.*, 2003, **28**, 709–731.

Bauwens, L., Giot, P., Grammig, J. and Veredas, D., A comparison of financial duration models via density forecast. *Int. J. Forecast.*, 2004, **20**, 589–604.

Bauwens, L. and Veredas, D., The stochastic conditional duration model: A latent factor model for the analysis of financial durations. *J. Econometr.*, 2004, **119**, 381–412.

Diamond, D.W. and Verrecchia, R.E., Constraints on short-selling and asset price adjustment to private information. *J. Financial Econ.*, 1987, **18**, 277–311.

Dufour, A. and Engle, R., Time and the price impact of a trade. *J. Finance*, 2000, **55**, 2467–2498.

Easley, D. and O'Hara, M., Time and the process of security price adjustment. *J. Finance*, 1992, **47**, 577–605.

Engle, R., The econometrics of ultra-high frequency data. *Econometrica*, 2000, **68**, 1–22.

Engle, R. and Russell, J., Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, 1998, **66**, 1127–1162.
Engle, R. and Lunde, A., Trades and quotes: A bivariate point process. *J. Financial Econometr.*, 2003, **1**, 159–188.
Fernandes, M. and Grammig, J., Non-parametric specification tests for conditional duration models. *J. Econometr.*, 2005, **127**, 35–68.
Ghysels, E., Gourieroux, C. and Jasiak, J., Stochastic volatility duration models. *J. Econometr.*, 2004, **119**, 413–435.
Grammig, J. and Maurer, K.O., Non-monotone hazard functions and the autoregressive conditional duration model. *Econometr. J.*, 2000, **3**, 16–38.
Hautsch, N., Capturing common components in high-frequency financial time series: A multivariate stochastic multiplicative error model. *J. Econ. Dynam. Control*, 2008, **32**, 3978–4015.
Jones, C., Kaul, G. and Lipson, M., Transactions, volume, and volatility. *Rev. Financial Stud.*, 1994, **7**, 631–651.
Lee, C. and Ready, M., Inferring trade direction from intradaily data. *J. Finance*, 1991, **46**, 733–746.
Manganelli, S., Duration, volume and volatility impact of trades. *J. Financial Mkts*, 2005, **8**, 377–399.
Russell, J., Econometric modeling of multivariate irregularly-spaced high-frequency data. Manuscript, University of Chicago, 1999.
Russell, J. and Engle, R., A discrete-state continuous-time model of financial transactions prices and times: The autoregressive conditional multinomial-autoregressive conditional duration model. *J. Business Econ. Statist.*, 2004, **23**, 166–180.
Tay, A., Ting, C., Tse, Y.K. and Warachka, M., Using high-frequency transaction data to estimate the probability of informed trading. *J. Financial Econometr.*, 2009, **7**, 288–311.
Zhang, M.Y., Russell, J.R. and Tsay, R.S., A nonlinear autoregressive conditional duration model with applications to financial transaction data. *J. Econometr.*, 2001, **104**, 179–201.

Appendix A: The AACD model

We assume the duration of the underlying tick processes of the ith trade, denoted by T_{ji} for $j=-1,0,1$, to be exponentially distributed with expected value ψ_{ji} conditional on the information set Φ_{i-1}. Thus, the probability density function of T_{ji} given Φ_{i-1} is

$$f_{T_{ji}}(t) = \frac{1}{\psi_{ji}} \exp\left(-\frac{t}{\psi_{ji}}\right), \tag{A1}$$

with cumulative distribution function

$$F_{T_{ji}}(t) = \Pr(T_{ji} \le t) = 1 - \exp\left(-\frac{t}{\psi_{ji}}\right), \tag{A2}$$

and survival function

$$S_{T_{ji}}(t) = 1 - \Pr(T_{ji} \le t) = \exp\left(-\frac{t}{\psi_{ji}}\right). \tag{A3}$$

Note that $f_{T_{ji}}(t)$ depends only on ψ_{ji}, the expectation of the inter-arrival time T_{ji}. However, the ψ_{ji} values are conditioned on the same information set Φ_{i-1}. Consequently, intensity functions for the various price movements are inter-related.

We denote the joint density of (w_i, t_i), conditional on Φ_{i-1}, as $p_i(w_i, t_i|\Phi_{i-1})$. Applying the rule of conditional probability, we have the following conditional joint density for (w_i,t_i):

$$p_i(k, t_i|\Phi_{i-1}) = \Pr\left(\bigcap_{j=-1}^{1} \{T_{ji} \ge x_i\}\right) f_{T_{ki}}(x_i|T_{ki} \ge x_i), \quad k = -1, 0, 1. \tag{A4}$$

The first term on the right-hand side of the above equation represents the probability of no transaction in the (open) interval (t_{i-1}, t_i). The second term is the conditional density of tick movement k occurring at time t_i, given no transaction in the interval (t_{i-1}, t_i). Invoking the independent exponential distribution assumption for T_{ji}, the second component of equation (A4) becomes

$$f_{T_{ki}}(t|T_{ki} \ge t) = \frac{f_{T_{ki}}(t)}{S_{T_{ki}}(t)} = \frac{1}{\psi_{ki}}. \tag{A5}$$

Substituting equation (A5) into equation (A4), and making use of the independence assumption of T_{ji}, equation (A4) can be written as

$$p_i(k, t_i|\Phi_{i-1}) = \left(\prod_{j=-1}^{1} S_{T_{ji}}(t_i - t_{i-1})\right)\lambda_{ki}. \tag{A6}$$

Substituting equation (A3) into equation (A6) yields the following expression for the conditional joint density of (w_i, t_i):

$$p_i(k, t_i|\Phi_{i-1}) = \lambda_{ki} \exp\{-(\lambda_{-1,i} + \lambda_{0i} + \lambda_{1i})x_i\}. \tag{A7}$$

Summing over the possible tick movements k in equation (A7) produces the conditional marginal density of x_i

$$\begin{aligned} f_{x_i}(x|\Phi_{i-1}) &= \sum_{k=-1}^{1} p_i(k, x + t_{i-1}|\Phi_{i-1}) \\ &= (\lambda_{-1,i} + \lambda_{0i} + \lambda_{1i}) \exp\{-(\lambda_{-1,i} + \lambda_{0i} + \lambda_{1i})x\}, \end{aligned} \tag{A8}$$

which follows an exponential distribution with mean $1/(\lambda_{-1,i} + \lambda_{0i} + \lambda_{1i})$. Conversely, after integrating over the duration x, the conditional marginal density of w_i is given by

$$\begin{aligned} f_{w_i}(k|\Phi_{i-1}) &= \int_0^\infty \lambda_{ki} \exp\{-(\lambda_{-1,i} + \lambda_{0i} + \lambda_{1i})x\}\, \mathrm{d}x \\ &= \frac{\lambda_{ki}}{\lambda_{-1,i} + \lambda_{0i} + \lambda_{1i}}, \quad k = -1, 0, 1. \end{aligned} \tag{A9}$$

As the joint density in equation (A7) is the product of the marginal densities for x_i and w_i, written in equations (A8) and (A9) respectively, they are independent conditional on the information set Φ_{i-1}. This result is a consequence of the underlying processes for tick movements being independent Poisson processes.

After each transaction, regardless of its outcome, the conditional expected duration of each latent process is updated and the competition restarts again. Thus, given a sample of observations $\{w_i,t_i\}$ for $i = 1, \ldots, N$, the log-likelihood function is

$$\sum_{i=1}^{N} \ln p_i(w_i, t_i) = \sum_{i=1}^{N} \left(\left(\sum_{j=-1}^{1} \ln S_{ji}(x_i) \right) + \ln\left(\sum_{j=-1}^{1} \lambda_{ji} D_{w_i}(j) \right) \right)$$
$$= -\sum_{i=1}^{N} \left(\sum_{j=-1}^{1} \frac{x_i}{\psi_{ji}} - \ln\left(\sum_{j=-1}^{1} \frac{D_{w_i}(j)}{\psi_{ji}} \right) \right), \tag{A10}$$

where $D_{w_i}(j) = 1$, if $j = w_i$ and 0 otherwise. The model parameters may be estimated using maximum likelihood estimation (MLE) once the functional forms of the conditional expected durations ψ_{ji} are specified.

The AACD model can be generalized to incorporate latent inter-arrival times that follow the Weibull distribution. The density function of a random variable X following the two-parameter Weibull distribution is

$$f_X(x) = \frac{\phi}{\psi}\left(\frac{x}{\psi}\right)^{\phi-1} \exp\left[-\left(\frac{x}{\psi}\right)^{\phi} \right], \tag{A11}$$

where ψ is the scale parameter and $\phi\ (>0)$ is the shape parameter. The survival function associated with the Weibull distribution is

$$S_X(x) = \exp\left[-\left(\frac{x}{\psi}\right)^{\phi} \right], \tag{A12}$$

while its mean equals

$$E(X) = \psi\, \Gamma\left(\frac{1}{\phi} + 1\right), \tag{A13}$$

where $\Gamma(\cdot)$ is the gamma function. Repeating the derivation in section 2, we obtain the conditional joint density function of w_i and x_i as

$$p_i(k, t_i | \Phi_{i-1}) = \frac{\phi}{\psi_{ki}}\left(\frac{x}{\psi_{ki}}\right)^{\phi-1} \exp\left[-\sum_{j=-1}^{1} \left(\frac{x}{\psi_{ji}}\right)^{\phi} \right],$$
$$k = -1, 0, 1. \tag{A14}$$

Denoting $\lambda_i^{\phi} = \lambda_{-1,i}^{\phi} + \lambda_{1i}^{\phi}$, we obtain the conditional marginal density of x_i as

$$f_{x_i}(x|\Phi_{i-1}) = \phi x^{\phi-1} \lambda_i^{\phi} \exp[-\lambda_i^{\phi} x^{\phi}], \tag{A15}$$

and the conditional marginal density of w_i as

$$f_{w_i}(k|\Phi_{i-1}) = \left(\frac{\lambda_{ki}}{\lambda_i}\right)^{\phi}, \quad k = -1, 0, 1. \tag{A16}$$

Thus, w_i and x_i are independent, conditional on Φ_{i-1}. More generally, Bauwens and Giot (2003) show that w_i and t_i are conditionally independent if the inter-arrival times of price movements are distributed according to a Weibull distribution with identical shape parameters. When these shape parameters take the value one, the Weibull distribution reduces to an exponential distribution. Bauwens and Giot (2003) provide empirical support for the conditional independence between tick movements w_i and durations x_i given the information set Φ_{i-1}. Note that the conditional independence of w_i and x_i given Φ_{i-1} is valid for only one trade. Due to the dependence of the conditional expected duration (intensity) parameters ψ_{ki} (λ_{ki}), the random variables w_i and x_i are statistically dependent over time, in particular $\{w_{i+1}, w_{i+2}, \ldots\}$ and $\{x_{i+1}, x_{i+2}, \ldots\}$ are statistically dependent conditional on Φ_{i-1}.

Finally, the log-likelihood function of the data is given by

$$\sum_{i=1}^{N} \ln p_i(w_i = j, t_i)$$
$$= -\sum_{i=1}^{N} \left(\sum_{j=-1}^{1} \left(\frac{x_i}{\psi_{ji}}\right)^{\phi} - \ln\left(\sum_{j=-1}^{1} D_{w_i}(j)\, \frac{\phi}{\psi_{ji}} \left(\frac{x_i}{\psi_{ji}}\right)^{\phi-1} \right) \right). \tag{A17}$$

Our decision to adopt the exponential distribution assumption is supported both by our data, as well as results in applications by other researchers. Bauwens and Giot (2003) report empirical evidence in support of the exponential inter-arrival times. Bauwens *et al.* (2004) compare the predictive performance of various conditional distributions of duration beyond the exponential. Of the many distributions considered there, none are clearly preferred over the exponential distributions. We have also estimated our AACD models assuming the Weibull distribution, which includes the exponential distribution as a special case. The results are very similar to those using the exponential assumption. Diagnostics also show no improvement in the Weibull implementation over the exponential implementation.